Advances in formal Slavic linguistics 2021

Edited by

Petr Biskup

Marcel Börner

Olav Mueller-Reichau

Iuliia Shcherbina

Open Slavic Linguistics

language
science
press

Open Slavic Linguistics

Editors: Berit Gehrke, Denisa Lenertová, Roland Meyer, Radek Šimík, Luka Szucsich & Joanna Zaleska

In this series:

1. Lenertová, Denisa, Roland Meyer, Radek Šimík & Luka Szucsich (Eds.). Advances in formal Slavic linguistics 2016.

2. Wiland, Bartosz. The spell-out algorithm and lexicalization patterns: Slavic verbs and complementizers.

3. Marušič, Franc, Petra Mišmaš & Rok Žaucer (eds.). Advances in formal Slavic linguistics 2017.

4. Blümel, Andreas, Jovana Gajić, Ljudmila Geist, Uwe Junghanns & Hagen Pitsch (eds.). Advances in formal Slavic linguistics 2018.

5. Dočekal, Mojmír & Marcin Wągiel (eds.). Formal approaches to number in Slavic and beyond.

6. Wągiel, Marcin. Subatomic quantification.

7. Kolaković, Zrinka, Edyta Jurkiewicz-Rohrbacher, Björn Hansen, Dušica Filipović Đurđević & Nataša Fritz. Clitics in the wild: Empirical studies on the microvariation of the pronominal, reflexive and verbal clitics in Bosnian, Croatian and Serbian.

Advances in formal Slavic linguistics 2021

Edited by

Petr Biskup

Marcel Börner

Olav Mueller-Reichau

Iuliia Shcherbina

language
science
press

Freie Universität Berlin

Contents

Contents

Preface

Advances in formal Slavic linguistics 2021 offers a collection of thirteen high quality articles on Slavic linguistics. The volume covers all branches of Slavic languages and features synchronic as well as diachronic analyses. It contains both empirically oriented work, underpinned by experimental methods or corpora analyses, and more theoretically based contributions. It comprises a wide array of topics, such as degree achievements, clitic climbing in Czech and Polish, typology of Slavic l-participles, aspectual markers in Russian and Czech, doubling in South Slavic relative clauses, congruence and case-agreement in close apposition in Russian, cataphora in Slovenian, Russian and Polish participles, prefixation and telicity in Serbo-Croatian, Bulgarian adjectives, negative questions in Russian and German and imperfectivity in discourse.

Early versions of the papers included in this volume were presented at the conference on Formal Description of Slavic Languages 14 or at the satellite Workshop on Secondary Imperfectives in Slavic, which were held in Leipzig on June 2–5, 2021 – the year referred to in the title of the volume. Originally, the conference was set to December 2020 but due to the Covid pandemic it had to be postponed and could only take place in the hybrid format in June 2021.

Three quarters of the submitted abstracts made it into the 36 presentations of the conference. Each article underwent an extensive reviewing process in line with the usual standards (double-blind peer reviewing). The conference also featured 5 invited talks. The 13 papers in the present volume were developed from these contributions in the course of a further thorough reviewing process. Neither the original conference nor the present volume would have been possible without the readiness of so many experts to devote their time and thoughts to the critical evaluation and helpful commenting of their colleagues' research papers. We would like to thank both the 38 anonymous reviewers for the present volume, and the more than 80 reviewers of the original conference abstracts.

This book would have also been impossible without our student assistants, Anastasiya Koretskykh and Julius Lambert. We also wish to acknowledge the extensive technical support of the whole Language Science Press editorial team, particularly Radek Šimík and Berit Gehrke.

Preface

Petr Biskup, Marcel Börner, Olav Mueller-Reichau & Iuliia Shcherbina
Leipzig, 21 July 2023

Chapter 1

Specification of telicity in Serbo-Croatian, without null prefixes

Boban Arsenijević

University of Graz

The paper reconsiders the claim that null prefixes must be posited in order to maintain the generalization that telicity is necessarily marked by an affix in Slavic languages (Łazorczyk 2010). Two classes of verbs apparently showing telic behavior without overt aspectual affixes are investigated on the empirical material from Serbo-Croatian: simple telic perfectives, and simple imperfectives compatible with the *za*-phrase (SC modifiers with the preposition *za* 'for' are equivalent to the English *in-X-time* expression, and SC modifiers without a preposition to the English *for-X-time* expression). It is argued that the former are indeed telic verbs without an aspectual affix, but that these verbs are idiomatically stored rather than being compositionally interpreted, and hence are irrelevant for the generalization. The latter are argued to be genuinely atelic. Their compatibility with the *za*-phrase is not evidence for telicity: *za*-phrase is not exclusively compatible with telic eventualities. This view is supported by a number of semantic and morphological similarities and differences between the verb classes involved, and quantitative evidence from corpus research. At least for Serbo-Croatian, then, Łazorczyk's (2010) generalization that telicity never occurs without affixes can be maintained without postulating null prefixes.

Keywords: null prefixes, telicity, aspectual pairs, simple verbs, Serbo-Croatian

1 Introduction

Slavic verbal aspectual morphology is a hallmark of both Slavic linguistics and general research of aspect, and probably needs no introduction – especially in a

Boban Arsenijević

volume from a Slavic conference encompassing a workshop on secondary imperfectives. I therefore give only a very superficial introduction to Slavic aspectual morphology, and then skip to the narrow topics of the article.

The prototypical morphologically simple Slavic verb (inflection morphology aside) is broadly assumed to be imperfective and atelic, as in (1a).[1,2] It derives a perfective telic verb by taking a lexical prefix – one that corresponds to the predicate of result, as in (1b), or the semelfactive suffix that imposes arbitrary bounds, as in (1c). The verb emerging as a perfectivized version of a simple imperfective can be imperfectivized again by a suffix, resulting in what is traditionally referred to as secondary imperfectives, as in (1d). Finally, both simple and secondary imperfectives can be perfectivized by a superlexical prefix: a prefix which does not express the result (in the narrow conventional sense as in Ramchand 2004, Svenonius 2004; but see Arsenijević 2007a,b, Žaucer 2009 for a resultative analysis of superlexical prefixes) and expresses a meaning related to the quantity of the event, as in (1e)–(1f), respectively.

(1) a. Pio   je čaj.
        drink.PTCP.IPFV AUX tea
        'He was drinking tea.'

---

[1]In the paper I qualify verbs as telic or atelic (i.e. unspecified), while it is actually the entire VP that can be telic or atelic and not the verb alone. In Slavic, however, a set of verbs is restricted to fitting in telic VPs only, and therefore describing them as telic is not false. Other verbs are unspecified for telicity, as argued in §1.2. Note also that the nature of and criteria for attesting telicity are highly debated issues, both in the general linguistic theory and in its application to Slavic languages. In the current paper, I do not go deeper into this discussion, but rather stick to the tests which display consistency when implemented on the Slavic linguistic material.

[2]In this paper, where relevant, verbs are specified for belonging to the traditional classes of perfective or imperfective verbs by the last item in their glosses. This item is added after a period, and does not correspond to any morpheme in the original example, which is meant to specify that this specification applies to the entire verb, and not to the last glossed morpheme.

Throughout the paper, I also use the standard marking of the grammaticality status of the example: ? for slightly degraded, ?? for strongly degraded, * for ungrammatical and % for examples grammatical in some varieties, i.e. for some speakers. The sign # is used for examples which are pragmatically or semantically ill-formed.

The relevant verbs in the examples are glossed following an assumed morphological analysis, i.e. decomposed into morphemes represented by their default morphs in order to keep the identity of morphemes coded across examples (one exception is allomorphy triggered by imperfectivization, where the exact allomorphs are given to keep represented the illustrated morphological operation). Due to phonological alternations, some of the morphemes in some of the given examples surface with different morphs.

All the examples in the paper are constructed by the author, who is a native speaker of the Ekavian standard Serbo-Croatian and the Torlakian dialect. For each constructed example, it has been verified in the corpus that the structural pattern used is attested in the syntactic and semantic environment equivalent to that which is constructed.

b. Od-pio   je   čaj.
from-drink.PTCP.PFV AUX tea
'He took a sip from the tea.'

c. Pi-nu-o   je   čaj.
drink-SUFF-PTCP.PFV AUX tea
'He took a sip from the tea.'

d. Od-pi-ja-o   je   čaj.
from-drink-SUFF-PTCP.IPFV AUX tea
'He was taking a sip / sips from the tea.'

e. Po-pi-o   je   čaj.
over-drink.-PTCP.PFV AUX tea
'He drank all the tea.'

f. Iz-od-pi-ja-o   je   čaj.
out-from-drink-SUFF-PTCP.PFV AUX tea
'He took sips of the tea to its exhaustion.'

A large number of observations, generalizations and problems have been reported and discussed in the rich literature in this field. This paper tackles one narrow, but core question in this domain: Is telicity universally marked by affixes in Slavic? In order to answer it, I discuss several related issues, most importantly the relevant opposition behind the traditional division of Slavic verbs into perfectives and imperfectives, including the structural representation and semantic content of the relevant asymmetry and the relation between the members of the so-called aspectual pairs. A remark is due regarding aspectual pairs, as their reality represents another unresolved issue in Slavic linguistics. I take two verbs to be an aspectual pair if one of them is perfective, the other imperfective, there is an independently attested morphological operation that derives one from the other, and, abstracting away from aspect, they mean the same. For polysemous verbs, it suffices that there is at least one meaning of the perfective and one of the imperfective verb such that the condition of semantic equivalence abstracting away from aspect applies to their combination. The availability of non-shared interpretations poses no problem for this relation. The criterion used to establish that two verbs form a pair is that a sentence can be constructed following the general pattern illustrated for two verbs in (2), such that the imperfective verb fits the first verbal slot and the imperfective the second.

(2)   a. Marija satima jede   sendviče   iznova i   iznova, i
         M   hours eat.PRES.3.SG.IPFV sandwiches again and again   and

Boban Arsenijević

upravo je   pojela   poslednji.
just   AUX.3.SG.PFV eat.PTCP.F.SG.PFV last
'Marija has been eating sandwiches again and again for hours, and
she just ate the last one.'

b. Jovan satima dotrčava   kući iznova i   iznova, i
J   hours run.to.PRES.3.SG.IPFV home again and again   and
upravo je   dotrčao   poslednji put.
just   AUX.3.SG.PFV run.to.PTCP.F.SG.PFV last   time.
'Jovan has been coming home running again and again for hours, and
he just came home running for the last time.'

I provide arguments from Serbo-Croatian (SC; all the examples in the paper are
from SC unless otherwise indicated) supporting the statements in (3).

(3) a. The strong generalization made by Łazorczyk (2010), that telicity is uni-
        versally reflected in affixal material, taking affixes as the feature con-
        tent of some relevant syntactic heads rather than the morphs surfacing
        on the verb, holds in SC without the need to postulate null prefixes.
     b. Morphologically simple verbs passing all or some tests as telic are
        either idiomatically stored and thus irrelevant for the generalization
        above, or are rather unrestricted for telicity (i.e. in the traditional view
        – atelic) with telic interpretations emerging from pragmatics.

The paper is organized as follows. §1.1 introduces the database that I use to in-
spect the relevant quantitative properties of the relevant verb classes, and §1.2
presents the relevant existing views of Slavic verbal aspect. §2 presents the struc-
tural model at the syntax-semantics interface assumed to underlie the aspectual
morphology and semantics in Slavic languages. In §3, I discuss the affixless per-
fectives and argue that they are all idiomatic, i.e. non-compositional, and hence ir-
relevant for the generalization about affixal marking of aspect. §4 gives a general
overview of the four classes of traditional imperfectives regarding telicity, with
a special attention to secondary imperfectives and simple imperfectives passing
some tests as telic. The latter class is then scrutinized in §5 with respect to the
issue of null prefixes, and it is argued that these verbs do not support the intro-
duction of null prefixes either. §6 concludes.

1.1 The empirical base

Besides the common sources of empirical data, including previous literature,
corpora and grammaticality judgments, the research reported includes quanti-

tative insights from the Database of the Western Slavic verbal system (Arsenijević et al. in preparation). The database consists of 5300 SC and 3000 Slovenian verbal lemmata retrieved from the srWac, hrWac and bsWac corpora for SC (Ljubešić & Erjavec 2011) and from the Slovenian National Corpus for Slovenian (FidaPLUS 2000). The verbs are selected based on frequency: the top 3000 highest frequency lemmata from each of the corpora are included and annotated. As srWaC, hrWaC and bsWaC are corpora of different SC varieties, the SC database contains the union of all three sets of 3000 verbs from the three corpora. Different morphophonological shapes that the same verbs had in two or all three varieties (e.g., Ekavian, Ijekavian, Ikavian versions or those emerging from using different integration suffixes to adopt borrowed verbs or to imperfectivize native ones) were introduced as separate entries, and annotated as variants of one verb. Each verb is annotated for a fixed set of over 40 different properties, including frequency, lexical and grammatical aspect as verified by the chosen tests, argument structure (taking accusative, genitive, dative, PP, clausal arguments; reflexivity), the characteristic morphemes (the root, prefixes, suffixes), their special properties (e.g. root-allomorphy), prosodic characteristics (position of the high tone, long syllables), theme vowels and others.

In the present investigation, the database was used to determine the quantitative properties of significance for the research such as the relative sizes of various relevant classes of verbs or their frequencies.

1.2 The background: The asymmetry underlying the opposition between the traditional Slavic perfective and imperfective verbs

As the central question of the paper concerns verbal aspect and affixation, the aim in this section is to highlight some of the relevant notions and introduce the views that are particularly important for the discussion to come, as a bridge to a more precise formulation of the research goals. The relation between the lexical and the grammatical aspect in Slavic languages and the role of prefixation have received numerous accounts, and still continue to evade an overarching analysis (Borer 2005, Borik 2006, Ramchand 2004, Arsenijević 2006, a.o.).[3] Regarding the nature of the morphologically marked opposition between the two classes of verbs in Slavic languages traditionally referred to as perfectives and imperfectives, Łazorczyk (2010) argues that Slavic verbs are only marked for the lexical aspect, and that the grammatical aspect is not marked up until the structural

---

[3] For a definition of notions like lexical and grammatical aspect, i.e. (a)telicity and (im)perfectivity, as well as quantization and homogeneity, incrementality etc. – see Milosavljević (2023).

Boban Arsenijević

level of inflection, i.e. it may only be marked by specific verb forms. Arsenijević (2018) divides Slavic verbs into those that are marked as perfective and those that are unmarked, hence ambiguous, but with an imperfective bias emerging via antipresupposition: that the speaker has not used a verb specified as perfective implies that the speaker did not want to convey a perfective meaning, but its disjunctive alternative: the imperfective one. Milosavljević (in preparation) argues for a hybrid between these views: as in Łazorczyk (2010), verbs in Slavic are only marked for lexical aspect, and as in Arsenijević (2018), they can be either strictly telic (the traditional perfective verbs) or unspecified for telicity (traditional imperfectives). Like in Arsenijević (2018), the atelic bias of the traditional imperfectives stems from antipresupposition, but is often additionally supported by the aktionsart (it is more difficult to impose a telic interpretation on verbs denoting states than on verbs denoting processes, which are in turn more difficult than verbs denoting culminating events, such as secondary imperfectives). Like in Łazorczyk (2010), grammatical aspect is specified at a higher structural level, in strong dependence on the value of lexical aspect (see e.g. Borik 2006 for discussion). The analysis I develop here builds on Milosavljević's view. In what follows, I spell out the exact telic and atelic interpretations between which traditional imperfective verbs are ambiguous (a more fine-grained discussion is offered in §4.2).

Based on the presented view, in the rest of the paper, I use the term ASPECTUALLY UNSPECIFIED (AU) VERBS for the traditional imperfective verbs, and ASPECTUALLY SINGULAR (AS) VERBS for the traditional perfective verbs in Slavic. AU verbs normally head verbal expressions that pass tests as atelic, i.e. homogeneous predicates (following Bennett & Partee 1972, Verkuyl 1972, Bach 1986, Krifka 1989 a.o. in assuming that properties of quantity mereologically modelled as quantization and homogeneity underlie the notions of telicity and atelicity, respectively). The predicate describing a state in (4a), or one describing a process, as in (4b), indeed by default show atelic behavior. As verbs normally head verbal predicates that display telic behavior and describe events involving a phase transition (which makes them quantized), as in (4c). Finally, there are also AU verbs which describe eventualities that involve a phase transition, as in (4d). I refer to this as the secondary imperfective pattern since it most frequently occurs with traditional secondary imperfectives (verbs derived by imperfectivizing a perfective verb, in the adopted terminology: secondary AU verb), but crucially for the present discussion – there are other classes (apparently) displaying this pattern too. Verbal expressions headed by these verbs normally pass tests both as telic and as atelic, and can be assigned four different readings.

6

1 Specification of telicity in Serbo-Croatian, without null prefixes

(4) a. Marija je spava-la (??za) dva sata.
   M AUX sleep-PTCP.IPFV for two hours
   (Intended:) 'Marija slept for/in two hours.'

   b. Marija je ras-la (??za) 15 godina.
   M AUX grow-PTCP.IPFV for 15 years
   (Intended:) 'Marija grew for/in 15 years.'

   c. Marija se u-spava-la *(za) dva sata.
   M REFL in-sleep-PTCP.PFV for two hours
   (Intended:) 'Marija fell asleep in / for two hours.'

   d. Marija se u-spavlj-iva-la (za) dva sata.
   M REFL in-sleep-SUFF-PTCP.IPFV for two hours
   'Marija was falling asleep in / for two hours.'

   i. process / preparatory stage: (Intended:) 'Marija was working on getting herself to sleep for two hours / in two hours.'

   ii. phase transition (slow motion): (Intended:) 'Marija was falling asleep for two hours / in two hours.'

   iii. an unbounded series of iterations: 'A series of iterations of events of Maria falling asleep (in two hours) was going on (for two hours).'

   iv. the general-factual reading: (Intended:) 'At least once in the past, Maria fell asleep for two hours / in two hours.'

While expressions headed by AS verbs are strictly telic, those headed by imperfective verbs display atelic behavior, but are not restricted to it. As soon as a possible source of quantization is introduced into the predicate describing the event – in terms of any kind of overtly, or contextually specified bounds – the predicate begins to display the secondary imperfective pattern, including passing the temporal duration modification test as telic (for a detailed discussion see Milosavljević 2023, in preparation). This is illustrated in (5), where the latent source of quantization is a measure phrase as in (5a) and (5b), i.e. the goal phrase as in (5c). I argue in this paper that these predicates have aspectually unspecified interpretations. The set of eventualities matching their extension includes pragmatically salient subsets which satisfy telic predicates (i.e. subsets consisting solely of bounded events). The latent sources of telicity in (5) merely support the pragmatic strengthening of the interpretation in the sense of Horn (1989) to one of these subsumed telic meanings.

Boban Arsenijević

(5) a. Marija je spavala   svoju dozu (za) dva sata.
       M   AUX slept.PTCP.IPFV her   dose   for two hours
       'Marija had her dose of sleep for/in two hours.'

   b. Marija je rasla   dva centimetra (za) godinu dana.
       M   AUX grew.PTCP.IPFV 2   cm   for year days
       'Marija grew two centimeters for/in a year.'

   c. Marija je putovala   do Lajkovca (za) dva sata.
       M   AUX travelled.PTCP.IPFV to Lajkovac for 2   hours
       'Marija (has) travelled to Lajkovac in/for two hours.'

As indicated by the examples in (1), (4) and (5), the simplest verbal predicates are unspecified for aspect, and there are various ways to assign them a telic interpretation. I argue in this paper that there are two degrees of strength of this assignment. Consider the verbal expression headed by a simple verb in (6a), which I analyze as unspecified for telicity with a strong bias for an atelic interpretation due to antipresupposition (the availability of a direct telic counterpart indicates that telicity was not intended). On the one hand, this predicate can be imposed telicity by prefixation, as in (6b) where a lexical prefix contributes a result, or in (6c), where a superlexical prefix specifies a bounded quantity. Alternatively, the suffix *-nu* may strongly impose telicity by specifying a quantity smaller than some contextually provided standard as in (6d). Both strong ways of imposing telicity make the verb perfective in the traditional sense.

On the other hand, a quantized incremental theme as in (6e) or a result (i.e. goal) specification, as in (6f), when the verb licenses one, may impose an interpretation which makes prominent a subset of events from the extension of the predicate, which itself matches a telic characteristic predicate. The example in (6e) makes prominent the set of eventualities measured out and thus telicized by the bounds of the daily dose of planking, and that in (6f) the set of eventualities telicized by a pair of a presupposed initial point and the explicated final point (muscle cramps). Finally, quantization may come from a measure phrase, as in (6g) (see also Pereltsvaig 2000, Szucsich 2001, Milosavljević 2023, in preparation for a discussion of temporal adverbials imposing telicity). In this latter set of cases, the verb remains AU, and the overall interpretation preserves its default atelic status. In §2, I argue that these bounds only provide a specification of atoms for the lexical component of the predicate, but do not necessarily include the contribution of the syntactic head responsible for telicity.

(6) a. Marija je radila   plenking (??za) dva sata.
       M   AUX do.PTCP.IPFV planking   for two hours
       'Marija did planking for two hours.'

b. Marija je   do-radila   plenking ??(za) dva minuta.
   M   AUX to-do.PTCP.PFV planking   for two minutes
   'Marija finished her planking in two minutes.'

c. Marija je   od-radila   plenking ??(za) dva minuta.
   M   AUX from-do.PTCP.PFV planking   for two minutes
   'Marija did her planking in two minutes.'

d. Marija je   rad-nu-la   plenking (za) dve sekunde.
   M   AUX do-SUFF-PTCP.PFV planking for two seconds

   i.   Without *za*: 'Marija did two seconds of planking.'

   ii. With *za*: 'Marija did a little bit of planking in two seconds.'

e. Marija je   radila   svoj dnevni plenking (za) dva sata.
   M   AUX do.PTCP.IPFV her daily   planking for two hours
   'Marija did her daily portion of planking for/in two hours.'

f. Marija je   radila   plenking do grča   mišića (za) dva sata.
   M   AUX do.PTCP.IPFV planking to spasm muscles for two hours
   'Marija did planking until her muscles cramped for/in two hours.'

g. Marija je   radila   plenking pet minuta za sat vremena.
   M   AUX do.PTCP.IPFV planking 5   minutes for hour time
   'Marija has done (at least once) an aggregate of five minutes of
   planking in one hour.'

Based on observations of this type, where prefixes and the semelfactive suffix correspond to obligatory telic interpretations, and other sources of quantization to rather latent telicity, the literature in the area of Slavic verbal aspect establishes a strong link between telicity and verbal prefixes. Fleischhauer & Gabrovska (2019) argue that the only way to derive telic verbal predicates in Slavic is prefixation, and Łazorczyk (2010) goes as far as claiming that the mapping is bijective: there is no telic verb without the suffix *-nu* or a prefix, nor is there *-nu* or a prefix that does not introduce telicity. For telic expressions showing no visible telicizing affixes, she postulates a null prefix. Expressions involving a morphologically simple verb with a latent quantization, and more generally all the expressions with an iterative interpretation, which can only be defined on the background of a telic predicate, raise the question whether the simple verbs heading them too involve a null prefix, whose contribution gets overwritten by a structural layer which re-imposes unspecification, or the attested interpretations are pragmatically promoted for truly simple verbs.

A related question concerns AU verbs which are prefixed. Łazorczyk (2010) considers that these prefixes introduce telicity, which is then neutralized by an

Boban Arsenijević

atelicizing operation (typically, secondary imperfectivization by a suffix). Considering that these verbs too pass tests both as telic and as atelic, a prominent analytic option is that the embedded telic structure is available for the tests of telicity. This would mean that the full predicate is atelic, but tests may also target its compositional components, and gives a reductionist advantage to one of the two analyses invoked above – the one with a null prefix also for the latently quantized simple AU verbs. The reductionist advantage lies in the fact that all (latently) quantized predicates can be generalized to involve a prefix, rather than having to define particular subclasses, some of which do and some do not involve a prefix. In light of the main goal of this paper, to scrutinize the arguments for null verbal prefixes in Slavic languages, this expands the empirical focus of the paper also to the simple AU verbs that may have progressive and iterative interpretations.

2 The assumed theoretical view

I present my view of the composition of verbal predicates using, for convenience, the framework of Distributed Morphology (DM, Halle & Marantz 1993), but it may equally well be formulated in terms of Nanosyntax (Starke 2010) or another realizational framework, as nothing crucially depends on the specific properties of DM. I take roots to denote predicates and to be able to take arguments (e.g. Travis 2012). The structure consisting of the root and its arguments is uncategorized, but I label it as the √-phrase (√P) for the purpose of reference in the text, without implying a syntactic projection. Once a root structure is categorized, its arguments may move up to positions introduced by functional projections.

√Ps can only merge with a category feature. The one relevant for the discussion is the verbal category. This is illustrated in (7) and Figures 1-2, where two *v*Ps are schematically presented, one without and one with a specified goal. The subject leaves the √P in both cases, but in (7b), i.e. Figure 2, there is additionally a predicate of the small clause, which remains inside the *v*P. The verbal category is assumed to be realized by the theme vowel (TH) in SC, and since the examples represent just the *v*P – the inflection is completely missing. I remain agnostic regarding the way the root ends up forming a word with the categorizer (via head-movement, PF dislocation or in some other way) as it is not relevant for the topic of discussion.

(7) a. ptica let-i-
         bird   fly-TH

b. ptica let-i-   na jug
   bird fly-TH on south
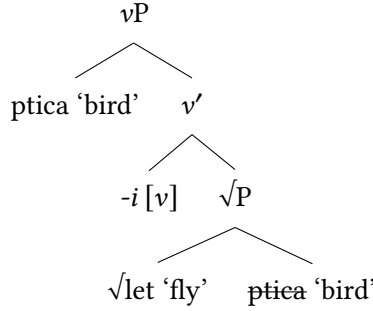
Figure 1: Syntactic representation of (7a)
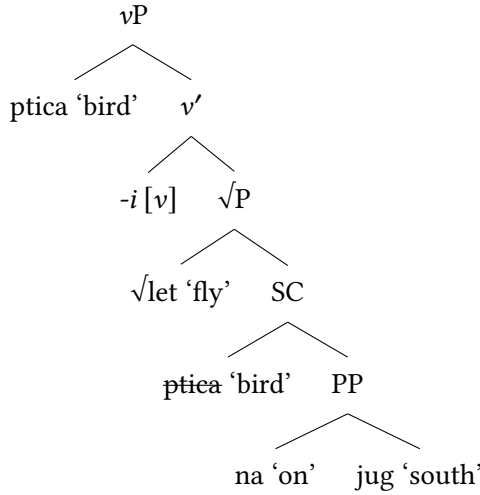
Figure 2: Syntactic representation of (7a)

I assume the verbal category to have a double contribution. It restricts the ontological class of the predicate to eventualities, and to kinds, by introducing a variable restricted to event kinds as the referential argument of the expression, and imposes division on the complement, thus acting as a grinder. In result, the *v*P denotes a non-atomic join lattice (as opposed to Chierchia's 1998 atomic join lattice for nominal kinds) satisfying the predicate in its complement (the base of

the lattice consists of parts of events). The vP in (7a), i.e. Figure 1, thus denotes all the possible sums over the maximal set of events of birds flying and all their parts, and the vP in (7b), i.e. Figure 2, all the possible sums over the maximal set of events of birds flying south and all their parts.

Recall that as explicated in §1.2, I consider that the aspectual division characteristic of Slavic verbs traditionally described as one of perfectivity is rather an opposition between telic and unspecified verbs. Following Borer (2005) and Łazorczyk (2010), I assume that it structurally corresponds to the presence or absence of a functional projection immediately above the category projection vP, which I label QP. Verbs with a QP above their vP are telic, i.e. they fall in the traditional class of perfectives, and those without it are atelic, i.e. AU verbs.

Unlike Borer (2005), Łazorczyk (2010), who take Q^0 to effect quantization, I take it (with Milosavljević in preparation – see his work for further arguments for this view and for references to relevant previous discussions) to impose a singular interpretation, i.e. to restrict the non-atomic join lattice to its base and impose atomicity on it. It is basically Filip & Rothstein's (2005) maximality operator, except that in the current approach it applies to the base of the lattice rather than to the entire predicate (here it would mean the entire lattice, in which case the derived denotation would be the sum of the entire base).[4] The meaning derived is the set of individual maximal event kinds satisfying the predicate denoted by the $\sqrt{}$P. On this view, a near equivalence can be established between telic verbal predicates (the denotations of AS verbs) and nominal singulars, as well as between atelic verbal predicates (denotations of AU verbs) and mass nouns. Consequently, the semantic effect of the respective head can be considered the same: whatever the way that singularity is imposed by singular number on nominals, is also the way it is imposed by the verbal counterpart (e.g. by blocking, or failing to provide, the sum operation needed to form the lattice).

The question emerges why telic predicates, i.e. predicates headed by AS verbs, tend to be used with the perfective viewpoint aspect, and atelic predicates, i.e. those whose expression involves AU verbs, with the imperfective viewpoint aspect. I assume that this mapping is pragmatically induced. A perfective viewpoint aspect presupposes boundedness – else it is logically impossible to take a perspective on the eventuality from a time outside of its temporal trace, or to have the trace be contained in the reference time, which are the standard ways of modelling the perfective viewpoint. In light of the view that all events

---

[4]One difference to Filip & Rothstein (2005) is that in their approach the verb includes in its denotation the atomicity crucial for the application of the maximality operator, while in the present approach atomicity is provided in a latent way by subconstituents of the VP, or simply by the context.

are presupposed to have initial bounds (Arsenijević 2006), a final bound suffices for quantization. If whenever the viewpoint aspect is perfective, the event predicate satisfies telicity, then perfective viewpoint aspect will present a pragmatically stronger interpretation of telic verbal expressions, and will thus undergo strengthening in Horn's (1989) sense whenever the context supports it. In result, quantized predicates, typically headed by AS verbs, will be the default way of describing eventualities viewed from the perfective perspective. On the other hand, if the reference time is properly included in the temporal interval of the eventuality, then within the reference time, it is impossible to epistemically verify the boundedness of the predicate (the ground for the imperfective paradox). Therefore, AU verbs are the default way of describing eventualities viewed in the imperfective perspective. That both these present pragmatic rather than semantic effects is evidenced by the fact that they can be cancelled: the general-factual use of AU verbs involves a perfective viewpoint, and instances of the imperfective paradox involve the use of AS verbs in interrupted progressive (hence imperfective viewpoint) contexts.

The feature representing the singular operator in QP needs to operate on a unit of counting, but is in itself underspecified for it. In the typical case, it receives this specification from the structurally closest compositional component of the $\sqrt{P}$ contributing the characteristic predicate of the atom to the aggregate predicate. I hence model this specification as a feature that is copied from the respective sub-predicate as the value of the singular atomizing feature in the head of QP. When such a predicate is absent from the structure, the singular feature receives the default value and the corresponding interpretation, where the unit of counting is the smallest eventuality satisfying the predicate for some contextually specified level of granulation. The singular feature with the default value is realized as the semelfactive suffix *-nu*, as illustrated in (8a), i.e. Figure 3.

When the singular feature takes a specific value and thus imposes atoms defined by the respective characteristic predicate as the unit of counting, valuation obtains via agreement: the singular feature probes into the c-commanded structure and agrees with the most local predicate specifying a possible unit of counting. Typically, this is a source, as in (8b), i.e. Figure 4, a goal, as in (8c), i.e. Figure 5, or a result predicate (of another kind). The contrast between, on the one hand (8a)–(8b), where the inclusion of the goal in the event is not entailed, and on the other (8c), where it is, is exactly predicted by the analysis: agreement with the goal, realized by a goal-oriented prefix on the verb, results in a restriction of the counting units to event-atoms specified by reaching the goal, and hence it cannot be negated. Effectively, in this example, the agreement of Q^0 with the predicate of the small clause, i.e. its promotion from a regular sub-argument into

Boban Arsenijević

the value of the feature singular, changes the interpretation of the small clause from the direction into the goal of the motion event. The absence of agreement or agreement with the source, as in the first two examples, allows for the negation of the reaching of the goal, since it leaves the small clause with the source interpretation.

(8) a. Ptica je let-nu-la   na jug, ali nije   stigla.
        bird AUX fly-SEM-PTCP on south, but textscneg.aux arrived
        'The bird flew south a little bit, but hasn't arrived.'

     b. Ptica je od-let-e-la   na jug, ali nije   stigla.
        bird AUX from-fly-TH-PTCP on south, but NEG.AUX arrived
        'The bird flew away towards south, but hasn't arrived.'

     c. Ptica je do-let-e-la   na jug, #ali nije   stigla.
        bird AUX to-fly-TH-PTCP on south, but textscneg.aux arrived
        'The bird came to the south flying, #but textscneg.aux arrived.'
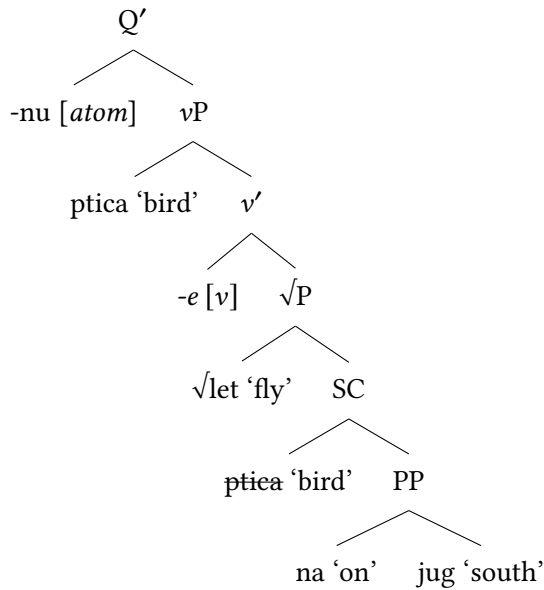
Figure 3: Syntactic representation of (8a)

These examples show that telicity, i.e. in the present account singularity, depends on the syntactic marking and not on the lexical description. The same lexical description (i.e. the same √P taking a path PP) composes into an atelic

14
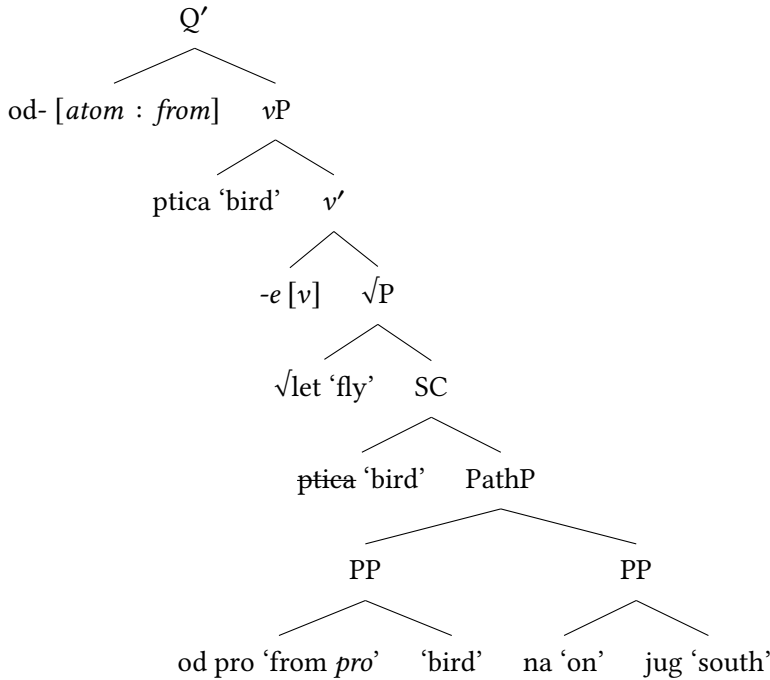
Figure 4: Syntactic representation of (8b)

predicate if no QP projects, as in (7b), i.e. Figure 2. Its path component (*na jug* 'to the south' in the examples above) is only a latent telicizer – it realizes this capacity only if a QP agrees with its predicate head. This is the reason why Quaglia et al. (under review) describe what I label QP as the result-Voice phrase – the projection that introduces the result as an argument of the verb by agreeing with the predicate of a respective phrase, copying its content and realizing it as a clitic. QP fits better as a label as it also includes the option with an unvalued singular feature realized by the suffix -*nu* as well as valuation by various adverbials (see Milosavljević 2023) or source prefixes. The fact that an atelic event predicate often stands in a superset relation to a (discourse-prominent) telic event predicate becomes relevant in §4.2, where simple AU verbs are discussed whose interpretation involves a prominent role of a salient telic predicate, as I argue – without involving a structural level specifying telicity.

Finally, I follow Arsenijević (2018) and Simonović et al. (2021) in analyzing secondary imperfectivization as reverbalization. Simonović et al. (2021) start from the observation that certain secondary imperfectives are derived by stacking an

Boban Arsenijević
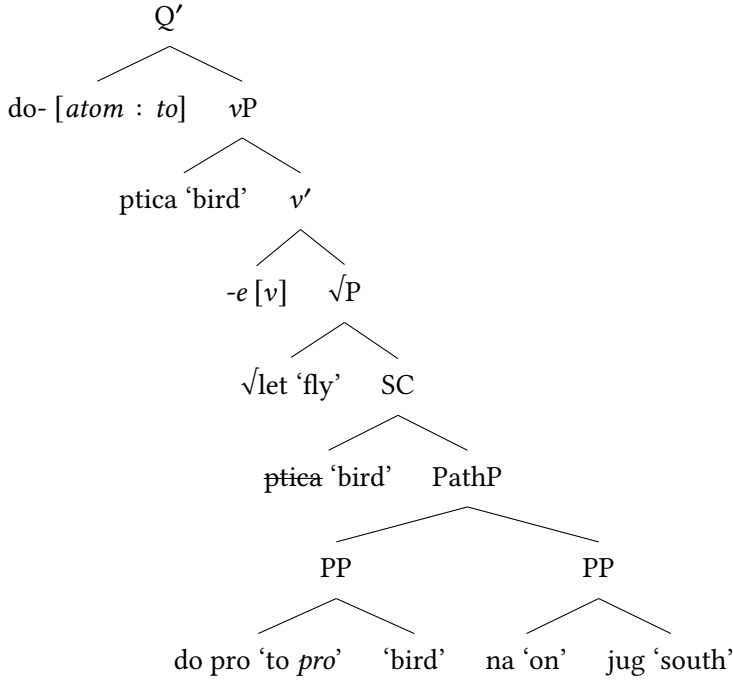
Figure 5: Syntactic representation of (8c)

additional theme vowel on top of the existing one, and that all imperfectivizing suffixes can be analyzed into two of the independently attested theme vowels with a consonant in between which plausibly realizes a glide. An analysis is developed where indeed secondary imperfectivization is always effected by either a single theme vowel or a sequence of two theme vowels. Considering that secondary imperfectivization targets AS verbs and assuming that theme vowels realize the category head, this implies that secondary imperfectivization amounts to deriving an unrestricted verb from a verb which is restricted to singularity. A new unrestricted verb is derived by merging the verbal structure with a new verbal category head, i.e. deriving a new verb from it. As in the present view, the category head *v* grinds the predicate in the complement, the contribution of the QP is neutralized and the verb denotes an AU predicate again. This is represented in (9).

(9) Ptica je do-let-e-a-la na jug.
bird AUX to-fly-TH-TH-PTCP on south
'The bird was coming to the south flying.'

16

(*do-let-e-a-la* is realized /doletala/ for reasons that I do not discuss)
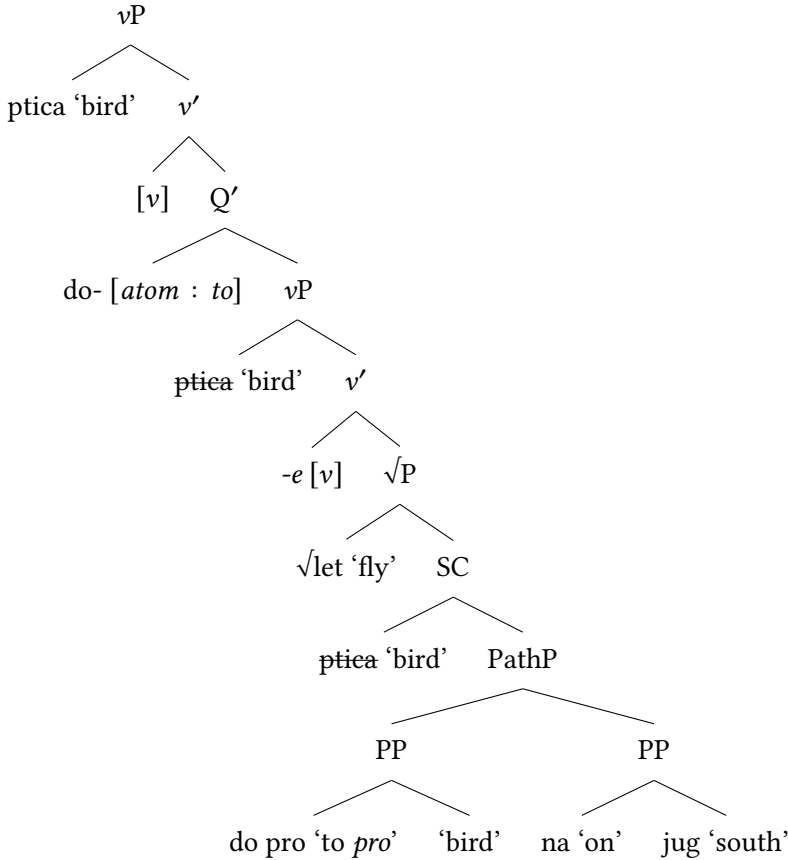
Figure 6: Syntactic representation of (9)

The theme vowel of the lower *v*P cannot be fully realized – it obligatorily contracts with the final segment of the root. While the theme ⟨j(e), i⟩, as in the example in (9), i.e. Figure 6, contracts without a trace, other themes, including ⟨i, i⟩ as illustrated in (10), palatalize the final segment of the base, or display other phonological effects.[5]

---

[5]All theme vowels in SC come in two allomorphs, surfacing in different subsets of verb forms. In the present paper, therefore, each theme vowel is represented as an ordered pair of the allomorph surfacing in the present tense and that surfacing in the infinitive, in that order.

Boban Arsenijević

(10) a. rod-i-ti   / rađ-a-ti
          bear-TH-INF.PFV bear-TH-INF.IPFV
          'give birth'

      b. set-i-ti   / seć-a-ti
          remember-TH-INF.PFV remember-TH-INF.IPFV
          'remember'

      c. u-prav-i-ti   / u-pravlj-a-ti.
          in-straight-TH-INF.PFV in-straight-TH-INF.IPFV
          'steer'

In the view presented, Łazorczyk's (2010) generalization about obligatory affixes in Slavic languages then translates as a requirement that the features copied to Q^0 by agreement be realized, whether or not the QP is embedded in a reverbalizing vP. Note that the view that Q^0 specifies singularity rather than quantization does not bear on the particular issue of affixation. The model outlined crucially departs from Borer (2005) and Łazorczyk (2010) in the reverbalization view of secondary imperfectivization. This issue is in particular relevant for the imperfective verbs compatible with the *za*-phrase, because it raises the question whether these verbs are vPs without a QP, in which case their behavior in tests of telicity needs to be explained, but their affixless realization is expected, or they are vPs projected on top of a QP, in which case their lack of prefixes needs to be explained (e.g., in terms of null prefixes), but their telic behavior on certain tests is expected. In §5, I argue for the former option.

3 Simple telic perfectives

Every Slavic language has a class of simple telic perfective verbs – i.e., verbs without prefixes or suffixes (other than inflection endings) that pass tests as telic. All these verbs describe achievements (or semelfactives), which makes them less compatible with durative adverbials. For this reason, I use the conjunction test (Verkuyl 1972) to illustrate their telicity in (11), where neither of the verbs allows for a single event interpretation characteristic of atelic verbal predicates.

(11) a. Jovan je stavio   mleko u frižider sinoć   i jutros.
          J   AUX put.PTCP.PFV milk in fridge last.night and this.morning
          'Jovan put the milk in the fridge last night and this morning.'
                                                              (two events only)

      b. Marija je spasila   psa iz reke sinoć   i jutros.
          J   AUX save.PTCP.PFV dog from river last.night and this.morning
          'Marija saved the dog from the river last night and this morning.'

(two events only)

Łazorczyk (2010) discusses verbs of this type in Russian and Polish, and postulates phonologically null prefixes to maintain the strong generalization that telicity is universally marked by an affix. This theoretical move both complicates the system and opens some additional questions such as the conditions on null realization of prefixes (when does the same feature get an overt and when a null realization), the grammatical status of null prefixes (what kind of empty category are they), their competition with overt prefixes. This calls for a thorough consideration of alternative analyses.

Simple telic perfectives have been downplayed in the literature as an enumerable closed class, plausibly listed in the lexicon (e.g. Toporišič 2000). If all these verbs are stored in the lexicon and idiomatic, then they do not pose a problem for the generalization that Slavic languages obligatorily mark singularity (i.e. telicity) by affixes, as the generalization only concerns compositionally derived telicity. Łazorczyk (2010) gives an ambiguous view of the issue. In one place (pg. 80), she compares simple perfectives with English irregular plurals, pointing out that both are small closed classes (hence likely listed). In another (pg. 28-29), however, she stipulates that null prefixes are productive, pointing out that in Russian simple loan verbs are easily used as perfective, and that in Bulgarian there are also a larger number of simple perfectives.

In Slavic languages, verbs are borrowed as biaspectuals. On the present approach, biaspectual verbs are irrelevant for the necessity of null prefixes, due to the fact that the meaning of AU verbs, identified with the homogeneous kind denotation of the vP, includes the base of the lattice, i.e. the denotation of the singular predicate and in the absence of competition (i.e. of a restricted atomic minimal pair) can be used for singular denotations. However, the claim that simple AS verbs too are productive, as indicated for Bulgarian, indeed supports the introduction of null prefixes.

As SC is similar to Bulgarian in having, at least at first glance, a larger number of simple AS verbs, I focus on establishing whether indeed this class can be considered productive, or it rather shows the quantitative properties of classes idiomatically listed in the lexicon. This question is best answered by a quantitative investigation into the size and frequency of the class of simple telic perfectives. A closed unproductive class fits a relatively small size (several dozens at most) and a high frequency. An open productive class makes the inverse prediction.

Among the 5300 SC verbs in the database by Arsenijević et al. (in preparation), 46 are annotated as telic simple verbs (throughout the quantitative report, by verbs, I refer to verbal lemmata in the corpus). This makes 5.5% of all the

simple verbs in the database, i.e. 0.85% of all the verbs included. On a more thorough analysis, it turns out that the class is even smaller, since the original set of verbs includes geographic variants of the same verb as well as verbs which on a closer look display the semelfactive suffix in certain forms and/or varieties. After cleaning up these verbs, the number of simple telic perfectives is reduced to 29 (this number cannot be used to calculate the percentage as the rest of the base is not cleaned from geographic variants, but indicates that they are significantly lower).

Besides being small, the class also includes at least three verbs with a somewhat archaic feel (*bataliti* 'quit', *turiti* 'put', *latiti se* 'tackle'), and not a single borrowed verb or neologism. Its average frequency (105.15 tokens per million) is more than three times higher than the average for the database (32.05) – another marker of low productivity (e.g., Plag 2012: 22-35). All in all, the quantitative data are compatible with treating these verbs as idiomatic and thus orthogonal to Łazorczyk's (2010) generalization. The stem of these verbs (the component consisting of the root and the theme vowel) is likely lexically stored with the semantics matching a QP, without a QP being projected, compositionally interpreted and realized as a prefix.

4 Imperfectives and telicity

Since secondary imperfectives are assumed in the present paper to be reverbalized telic event kinds, and hence each of these verbs embeds a structure which represents a telic event kind, the generalization investigated in the paper accents the question whether there are simple verbs with a semantics equivalent to secondary imperfectives (i.e. having progressive and iterative meanings). If there are such verbs, they too become relevant for the generalization that telicity is universally marked by an affix. The reason is that secondary imperfectives are taken to include a QP, and therefore morphologically simple verbs expressing the semantics of secondary imperfectives might also be a class in which the QP is present but not realized – contra Łazorczyk's (2010) generalization. This section identifies a class of verbs that at the first sight match the described pattern, and discusses them in light of the generalization.

Before focusing on simple imperfectives, a discussion is due of imperfectives more generally, and their behavior regarding aspectual pairs. This discussion is intended to show two things. The first is to identify the class of simple imperfectives indicated above and the second is to argue that the relevant, iterative interpretation of such simple imperfectives always also has a perfective realization by a verb involving an overt prefix. As it is well known that there is a strong

correlation between the combination of meaning and argument structure on the one hand and the prefix on the other, this supports the analysis in terms of null prefixes, counterparts of those visible on the perfective pair. In the rest of this section I pursue a more detailed analysis of these issues, leading to the conclusion that simple imperfectives are never full equivalents of secondary imperfectives, and that they consequently do not involve null prefixes either.

4.1 Four classes of imperfectives regarding aspectual pairs

The notion of aspectual pairs holds a prominent place in the theory of Slavic verbal aspect. An aspectual pair consists of two verbs with exactly the same meaning and argument structure, distinguished minimally in their aspect: one of them belongs to AS verbs and the other to imperfectives. The prototypical aspectual pair involves a perfective verb and its secondary (i.e. derived) imperfective, but as discussed below, pairs may also be argued to exist where the perfective seems to morphologically include the imperfective (Janda & Lyashevskaya 2011), as well as where both members appear to display the same degree of morphological complexity (these are the pairs whose perfective members are the simple perfectives from §3). The case where the perfective seems to derive from its imperfective pair by prefixation has been subject to debate with respect to the role of the prefix. As the verbal prefix in Slavic languages contributes conceptual content beyond its grammatical effect, the question is how can the prefixed verb still mean the same as its prefixless imperfective pair. In the Russian grammatical tradition, two different answers to this question have been proposed. On one, the prefix in such cases is void of any conceptual content (Vinogradov 1938, Šaxmatov 1941, Švedova 1980). On the other, referred to as the implication, or the overlap approach, the meaning of the prefix is included in the meaning of the verbal base – hence, it does not add any new content (Isačenko 1960, Timberlake 2004, Janda & Lyashevskaya 2011).

In a somewhat modified version of Maslov's (1948) classification of imperfective verbs regarding their aspectual pairs, I divide them into four classes: (i) secondary imperfectives, illustrated in Table 1,[6] (ii) simple imperfectives that have prefixed perfective pairs, while when the semelfactive suffix *-nu* is added their meaning is changed beyond the aspectual contrast, illustrated in Table 2 (the suffixed perfectives of these verbs are typically rare in use, need to be productively

---

[6]The morphological analysis assumed includes the theme vowel of the base verb in its secondary imperfective, even though in some examples, including those used in these examples, it is not visible on the surface (e.g. in the lengthening of the vowel). For arguments in favor of this analysis and reason for the lack of surface effects, see Simonović et al. (2021).

derived, and bear the flavor of a neologism), (iii) simple imperfectives that have perfective pairs with the semelfactive suffix -*nu*, while all their prefixed counterparts display semantic shifts, as in Table 3 and (iv) simple imperfectives that have no proper aspectual partners – as both the prefixed and the suffixed variant bear additional or shifted semantics as in Table 4. The difference between the last two classes is that the simple AU verbs forming an aspectual pair via suffixation denote cumulative atomized predicates, i.e. predicates describing iterations of a more or less clearly individuated atom (waving consists of atomic waves, banging of atomic bangs, nodding of atomic nods), while simple imperfectives without perfective partners have prototypical mass properties. The former then present another type of verbal expressions which build on atomic lexical descriptions but are not singular due to the lack of a QP (recall the discussion around example (8)). Of particular importance for the discussion are simple imperfectives with prefixed perfective partners.

Table 1: Aspectual pairs including a secondary imperfective

IMPERFECTIVE (class (a))		
iz-bac-i-iva-ti	u-trlj-a-ava-ti	do-trč-a-ava-ti
out-throw-TH-SUF-INF	in-rub-TH-SUF-INF	to-run-TH-SUF-INF
'throw out'	'rub in'	'run to'
PERFECTIVE (a minimal pair)		
u-baciti	pro-trljati	do-trčati
in-throw	through-rub	to-run
'throw in'	'rub a little'	'run to'

4.2 Simple imperfectives with perfective pairs: Their aspectual properties

Whether secondary imperfective verbs, derived from telic perfectives, are telic, atelic or both has been a matter of debate. One group of authors argue that all imperfectives are atelic (Borer 2005, MacDonald 2008, Łazorczyk 2010), another treats them as possibly telic (Arsenijević 2006, Borik 2006, Braginsky & Rothstein 2008, Stanojević 2012, Fleischhauer & Gabrovska 2019). I assume here, as discussed in §1.2, that secondary imperfectives are unspecified for telicity, but they are derived from telic predicates over event kinds.

1 Specification of telicity in Serbo-Croatian, without null prefixes

Table 2: Simple imperfectives with prefixed perfective pairs

IMPERFECTIVE (class (b))		
ređati	pržiti	kriviti
arrange	fry	blame
'arrange'	'fry'	'blame'
PREFIXED PERFECTIVE (a minimal pair)		
po-ređati	iz-pržiti	o-kriviti
over-arrange	out-fry	around-blame
'arrange'	'fry'	'blame'
SUFFIXED PERFECTIVE (not a minimal pair)		
ređ-nu-ti	prž-nu-ti	kriv-nu-ti
arrange.SUFF-INF	fry.SUFF-INF	blame.SUFF-INF
'arrange a bit'	'fry a bit'	'blame a bit'

Table 3: Simple imperfectives with suffixed perfective pair

IMPERFECTIVE (class (c))		
mahati	lupati	klimati
wave	bang	nod
'wave'	'bang'	'nod'
PREFIXED PERFECTIVE (not a minimal pair)		
od-mahati	u-lupati	raz-klimati
from-wave	in-bang	away-nod
'wave back'	'whisk'	'loosen'
SUFFIXED PERFECTIVE (a minimal pair)		
mah-nu-ti	lup-nu-ti	klim-nu-ti
wave-SUFF-INF	bang-SUFF-INF	nod-SUFF-INF
'wave'	'bang'	'nod'

Boban Arsenijević

Table 4: Simple imperfectives without perfective pair

IMPERFECTIVE (class (d))		
sedeti	mrzeti	smrdeti
sit	hate	stink
'sit'	'hate'	'stink'
PREF. PERFECTIVE (not a minimal pair)		
pre-sedeti	za-mrzeti	u-smrdeti
across-lead	for-hate	in-stink
'sit through'	'start hating'	'make stinky'
SUFFIXD PERFECTIVE (not a minimal pair)		
?sed-nu-ti	?mrz-nu-ti	?smrd-nu-ti
sit-SUFF-INF	hate-SUFF-INF	stink-SUFF-INF
'sit a bit'	'hate a bit'	'stink a bit'

Consider the tests of telicity in (12). On the temporal adverbial test, secondary imperfectives pass both the test for telicity and for atelicity. On the temporal conjunction test, they turn out to be atelic: they can combine with a conjunction of *at-x-time* expressions with a single event interpretation. So why do different tests give different results (see also Mittwoch 2010, 2013 and Milosavljević 2023 for a critical assessment of the tests of telicity)?

(12) a. Marija je rasklapala   pušku dva minuta.
       M   AUX disassemble.PTCP.IPFV rifle   two minutes

    i. Process / preparatory stage: 'Marija was removing parts of the rifle for two minutes (without necessarily reaching the completion).'

    ii. Phase transition (slow motion): 'Marija was completing her disassembling of the rifle for two minutes (the completion is being reached).'

    iii. An unbounded series of iterations: 'A series of iterations of events of Maria disassembling the rifle was going on for two minutes.'

  b. Marija je rasklapala   pušku za dva minuta.
     M   AUX disassemble.PTCP.IPFV rifle   for two minutes
     An unbounded series of iterations: 'A series of iterations of events of Maria disassembling the rifle in two minutes was going on.'

1 Specification of telicity in Serbo-Croatian, without null prefixes

    c. Marija je   rasklapala   pušku u pola pet i   u pet.
        M   AUX disassemble.PTCP.IPFV rifle   in half five and in five
        'Marija was disassembling a rifle at half past five and at five o'clock.'

I argue, based on the discussion in §2, that the reason for ambiguity is that the tests target different structural levels, corresponding to different predicates. One level is the QP, and the other the reverbalizing *v*P. The former is accessible to the temporal duration adverbial (and it has to be the one with *za* 'for', since only that one matches the QP), but not to the conjunction of temporal adverbials locating the epistemic evaluation time (i.e. reference time), because the epistemic evaluation time is only specified for the reverbalized structure. The reverbalizing *v*P is hence accessible to both kinds of temporal adverbials, but without another QP on top of the reverbalizing *v*P, the temporal duration adverbial has to be the bare one (on the adopted view of aspect, *za*-adverbials must be taken to require restriction to singularity). On a closer scrutiny, hence, secondary imperfectives are AU, but they embed a telic event kind, which yields an illusion of telicity with adverbials for duration.

    In the present paper, I do not dwell on this aspect of the proposal, but turn to its consequences for the main argument of the paper.[7] Secondary imperfectives all include an affix realizing the QP, in line with Łazorczyk's (2010) generalization. However, if there are simple imperfectives which are semantically equivalent to secondary imperfectives, then they embed a QP but do not realize it morphologically. This directs our attention to the simple imperfectives with prefixed perfective counterparts. These verbs have available the same readings as secondary imperfectives: the two progressive interpretations (zooming in onto the process subevent or onto the phase transition) and the iterative one, as illustrated in (13). Moreover, they combine with the *in*-phrase, which is interpreted as a measure of

---

[7]Prompted by a suggestion by the editors, I provide one quick argument in favor of this view. The analysis predicts that when occurring together as modifiers of secondary imperfectives, being embedded deeper than the bare ones, *za*-adverbials are harder to move higher in the structure than bare temporal duration adverbials. Indeed, for instance, fronting for focalization (with the focal stress indicated in (i) below by the capital letters) is more readily accessible to the bare adverbials, than to the *za*-adverbials (the latter only works as correction).

    (i)   a. DVA SAta je Marija rasklapala   pušku za dva minuta.
           two hours AUX M   disassemble.PTCP.IPFV rifle for two minutes
           'It was for two hours that Marija was disassembling the rifle in two minutes.'

        b. Za DVA miNUta je Marija rasklapala   pušku dva sata.
           for two minutes AUX M   disassemble.PTCP.IPFV rifle two hours
           'It was in two minutes that Marija was disassembling the rifle for two hours.'

Boban Arsenijević

the temporal interval of a related telic event predicate (describing the repeating unit in the iterative interpretation), as well as with the *for*-phrase, which is interpreted as a measure of the temporal interval of the event denoted by the derived imperfective as in the reading in (13a-iii). The conjunction test verifies atelicity, as shown in (13c).

(13) a. Marija je punila   pušku dva minuta.
         M   AUX charge.PTCP.IPFV rifle   two minutes

    i. Process / preparatory stage: 'Marija was putting bullets in the rifle for two minutes.'

    ii. Phase transition (slow motion): 'Marija was on the verge of finishing charging the rifle for two minutes.'

    iii. An unbounded series of iterations: 'A series of iterations of events of Maria charging the rifle was going on for two minutes.'

  b. Marija je punila   pušku za dva minuta.
      M   AUX charge.PTCP.IPFV rifle   for two minutes
      An unbounded series of iterations: 'A series of iterations of events of Maria charging the rifle in two minutes was going on.'

  c. Marija je punila   pušku u pola pet i   u pet.
      M   AUX charge.PTCP.IPFV rifle   in half five and in five
      'Marija was charging a rifle at half past five and at five o'clock.'

The remaining two bigger types of simple imperfectives, those with suffixed perfective counterparts and those without any, have only one interpretation, as illustrated in (14) and (15), which makes them uninteresting for the current investigation – except as evidence that other patterns exist.

(14) a. Marija je mahala   (*za) dva minuta.
         M   AUX wave.PTCP.IPFV for two minutes
         (Intended:) 'Marija was waving for/in two minutes.'

  b. Marija je mahala   u pola pet i   u pet.
      M   AUX wave.PTCP.IPFV in half five and in five
      'Marija was waving at half past five and at five o'clock.'

(15) a. Marija je spavala   (*za) dva minuta.
         M   AUX sleep.PTCP.IPFV for two minutes
         (Intended:) 'Marija was sleeping for/in two minutes.'

  b. Marija je spavala   u pola pet i   u pet.
      M   AUX sleep.PTCP.IPFV in half five and in five
      'Marija was sleeping at half past five and at five o'clock.'

26

To sum up, two classes of simple imperfectives, those illustrated in (14) and (15), are plain atelic predicates, even though one of them involves atomic lexical descriptions. This indicates that they lack the QP, which is in line with their lack of aspectual affixation. The third class, simple imperfectives with prefixed perfective pairs, are more similar to secondary imperfectives, in involving an atomic lexical description, having both iterative and progressive interpretations and being compatible with the *za*-phrase (taken to attest telicity). The relevant question is whether these verbs involve a QP as the above grouping suggests, and therefore require the postulation of null prefixes, or they describe eventualities that lend themselves well to the singular interpretation, but do not have it structurally realized, and hence structurally correspond to a primary *v*P rather than to a reverbalizing one. In the latter case, the consequence is that no verbs in SC involve the compositional contribution of a QP without overtly realizing it through affixation.

Before providing a deeper analysis, it is important to establish whether simple imperfectives patterning with secondary imperfectives are productive or whether they too can be considered a listed idiomatic class and therefore irrelevant for the discussion.

4.3 Quantitative insights

The database of SC verbs (Arsenijević et al. in preparation) includes 1886 derived AU verbal lemmata, more than double the number of simple ones, which count 720 items. Exactly 800 of the derived AU verbs are secondary imperfectives, i.e. verbs derived from perfectives by an imperfectivizing suffix. All of them form aspectual pairs with their perfective bases, following the pattern in Table 1. The remaining derived AU verbs fall in three classes: those that are not part of a minimal aspectual pair (758 verbs), those that have an aspectual partner formed by an additional prefix (40 verbs) and those derived from on nouns, adjectives and borrowings, typically with a biaspectual interpretation (288 verbs, which due to their non-verbal base, are not listed in the tables below).

Among the 720 simple imperfectives in the database, 344 have prefixed and 39 suffixed aspectual partners, 17 are derived from simple perfectives by adding a theme vowel and 320 do not form aspectual pairs. The quantitative data are summarized in Table 5.

I argued in §3 that simple perfectives are lexically listed and idiomatic. The 17 simple imperfectives deriving from them by reverbalization are then also not problematic for the generalization that Q^0 must be realized by an affix. However,

Boban Arsenijević

Table 5: Classes of imperfectives and their sizes, summarized

	Derived	Simple
Prefixed partner	40	344
Suffixed partner	/	39
Partner has different theme vowels	/	17
No pair	758	320
Secondary imperfective	800	na

if it turns out that they indeed embed structures deriving singular (i.e. telic) predicates, the 344 simple imperfectives with prefixed perfective pairs are less likely to be listed. This makes below 8% of all the verbs in the database and 47.78% of all the simple imperfectives, a fraction unlikely to be listed as idiomatic. The average frequency of this class is 22.33 tokens per million, which is lower than the average for the database, at 32.05. This too is compatible with treating the class as productive.

5 Simple imperfectives with prefixed aspectual pairs are truly simple

Łazorczyk (2010) generalizes that telicity (in my approach singularity) must be marked in QP by a prefix or by the semelfactive suffix -*nu*. Such verbs may then be reverbalized, thus becoming AU again. Reverbalization too must be morphologically marked in Slavic, and as argued by Simonović et al. (2021) – this marking consists in (sequences of) theme vowels. This is illustrated for prefixed perfectives in (16a) for a single theme and in (16b) for a sequence, as well as in (16c) for simple perfectives.

(16) a. u-vid-e-ti > u-vid-e-a-ti, /uviđati/
        in-see-TH-INF.PFV   in-see-TH-TH-INF.IPFV
        'realize/see'

    b. po-plav-i-ti > po-plav-i-i-a-ti, /poplavljivati/
        over-flood-TH-INF.PFV   over-flood-TH-TH-TH-INF.IPFV
        'flood'

    c. stav-i-ti > stav-i-a-ti, /stavljati/
        put-TH-INF.PFV   put-TH-TH-INF.IPFV
        'put'

The prediction for imperfective verbs lacking prefixes or the semelfactive suffix is thus that if they are underlying secondary imperfectives – they will have at least two theme vowels each, and if they are truly simple – they will have exactly one.

(17)   a.   cvat-∅-ti, /cvasti/
          bloom-TH-INF.IPFV
          'bloom'

      b.   vežb-a-ti
          exercise-TH-INF.IPFV
          'exercise'

      c.   kvar-i-ti
          spoil-TH-INF
          'spoil'

Empirical data indicate the latter. The verbs in question give no ground for identifying more than one theme vowel. This is most obvious on verbs with the theme ⟨∅, e⟩, illustrated in (17a) (there are 34 such verbs among the simple imperfectives compatible with the *za*-phrase, which include 720 verbs). Since the theme ⟨∅, e⟩ never occurs in reverbalizing sequences, in verbs of this class, an additional theme ⟨a, a⟩ (the only theme able to reverbalize alone) or a sequence of themes occurring as a reverbalizer would be clearly visible. Examples with other themes are given in (17b) and (17c).

This view is further supported by the 17 imperfective aspectual partners reported as simple above in §4.3. Under the here adopted analysis from Simonović et al. (2021), these verbs actually need to be treated as derived by secondary imperfectivization. The reason is that they can be convincingly argued to involve a thematic vowel on top of that realized on the perfective pair, as in those contexts in which the lower theme is expected to surface – it indeed does. This is illustrated in (18), where in (18a), the lower theme is null, hence invisible, in (18b) the final consonant of the root fully absorbs the front theme vowel, but in the contexts like (18c), where the contraction results in a phonological change, the change is attested on the surface.[8] This strengthens the assumption that in SC, secondary imperfectivization is never null, and that the imperfectives that do not show any traces of it are indeed simple.

---

[8]Simonović et al. (2021) analyze certain morphological realizations of the verbal category to involve a floating high vowel, which is realized only when it resolves the hiatus and else silent. These are represented in the examples in (18) and in the following in the superscript.

Boban Arsenijević

(18) a. pad-∅-ti   > pad-$^\varnothing$-a-ti, /padati/
          fall-TH-INF.PFV   fall-TH-TH-INF.IPFV
          'fall'

      b. bac-i-ti   > bac-^{i}-a-ti, /bacati/
          throw-TH-INF.PFV throw-TH-TH-INF.IPFV
          'throw'

      c. stav-i-ti   > stav-^{i}-a-ti, /stavljati/
          put-TH-INF.PFV   put-TH-TH-INF.IPFV
          'put'

Semantic evidence also goes in this direction. I report two relevant observations. The first is that both secondary imperfectives and simple imperfectives compatible with the *za*-phrase fail to license the progressive interpretation in combination with the *za*-phrase. The narrow iterative interpretation with a series of events in one reference time is also unavailable – the only iterative reading available is the general-factual use distributed over a plural reference time (an instantiation of the event kind satisfying both the temporal adverbial and the verbal predicate has taken place in each from the set of relevant reference times). This is illustrated in (19). The sentence with a perfective verb in (19a) has an interpretation which involves reference to an event in the past, while the two sentences with imperfective verbs, the one with a simple imperfective in (19b) and the one with a secondary imperfective in (19c), only can be used if the question under discussion is whether events of Jovan running (in)to the school in ten minutes, i.e. Jovan interrogating Marija in ten seconds, have taken place in each of a set of discourse-given or accommodated reference times – but not to actually refer to a series of such events. The failure to refer to an individual event is also reflected in the fact that the latter two sentences cannot have the progressive interpretation (Jovan was in the process of running (in)to the school in ten minutes and Jovan was in the process of interrogating Marija in ten seconds, respectively). This asymmetry is triggered by the *za*-phrase, as without it, all three sentences can have the progressive interpretation, in addition to other options (see (21a)).

(19) a. Jovan je   po-je-∅-o   kolač za deset sekundi.
          J   AUX over-eat-TH-PTCP.PFV cake for 10   seconds
          'Jovan completed an event of eating a cake and it took 10 seconds.'

      b. Jovan je   trč-a-o   u školu za deset minuta.
          J   AUX run-TH-PTCP.IPFV in school for 10   minutes
          'Jovan used to get to school running in ten minutes.'

   c. Jovan je   iz-pit-i-^{u}a-o   Mariju za deset sekundi.
     J   AUX out-ask-TH-TH-PTCP.IPFV M   for 10   seconds
     'Jovan used to complete interrogations of Marija in ten seconds.'

Secondary imperfectives uncontroversially embed telic structures, i.e. QPs. At first glance, the observed parallel seems to support the view that simple imperfectives compatible with the *za*-phrase embed a QP too, i.e. that they are secondary imperfectives which fail to show the morphological signature of reverbalization, and should be modelled in terms of null prefixes. The failure of such verbs modified by the *za*-phrase to refer to a single event is then due to the *za*-phrase occurring at the level of the QP, below the reverbalizing *v*P. The latter derives a kind, and therefore the *za*-phrase can only be interpreted at the kind level.

However, also the alternative, that simple imperfectives compatible with the *za*-phrase are primary *v*Ps (i.e. verbalized √Ps), has the potential to account for this interpretation. Assume that the predicate denoted by the √P is modified by the *za*-phrase. It thus contributes to the predicate that is verbalized in the same fashion as the goal PP – before the meaning is homogenized by the category head. After categorization, an AU predicate is derived denoting a sum of events of running to school in 10 minutes, all their parts, and all the sums thereof. Since AU predicates are weaker than AS predicates, under multiple reference times, the interpretation gets pragmatically strengthened (Horn 1989) to the AS interpretation, i.e. to including one maximal event per reference time.

This view raises two questions. One is, if the *za*-phrase can modify the √P, then how is it excluded from other atelic verbal predicates, i.e. how does it derive the behavior that has qualified it as a test for telicity? The *za*-phrase requires that the modified predicate specifies a possible atom, not necessarily that it is singular. This is exactly what characterizes the simple imperfectives that resemble the secondary ones. One of the other two classes are verbs denoting states (the pattern in Table 4), and their roots clearly specify no atoms. The other includes event predicates which are inherently atomic, but do not specify or likely lead to a result (the pattern in Table 3). These verbs do not combine with the *za*-phrase because their atoms are conceptualized to take a point in time, and hence resist this type of modification just like semelfactives do. This is confirmed by the fact that when an appropriate context is set, which implies a prolonged duration, modification is actually possible. This is illustrated in (20).[9]

---

[9](20a) is acceptable if the *za*-phrase measures the epistemic evaluation time: for a second, Jovan was waving, but this interpretation is orthogonal to the issue.

Boban Arsenijević
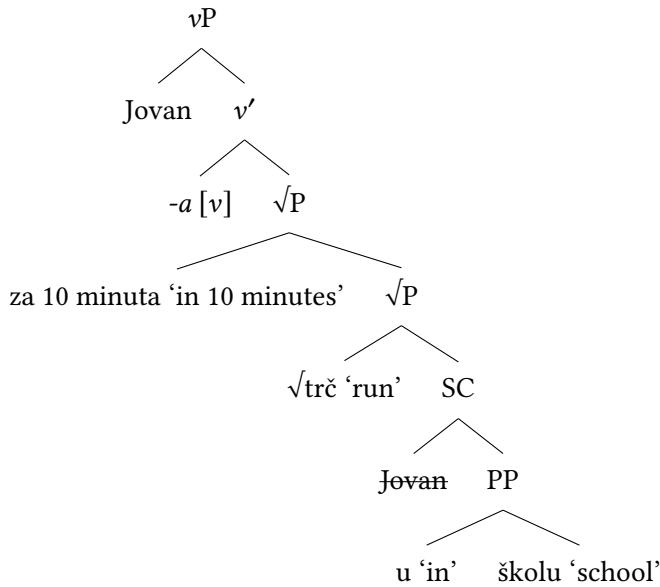
Figure 7: Syntactic representation of (19b)

(20) a. ?? Jovan je   mah-a-o   za sekund.
         J   AUX wav-TH-PTCP.IPFV for second
         'Jovan used to wave in a second.'

      b. ?? Jovan je   mah-nu-o   za sekund.
         J   AUX wav-TH-PTCP.PFV for second
         'Jovan waved once in a second.'

      c. *Context:* The task was to wave a big flag as fast as possible, while
         always making full waves from one horizontal direction of the flag
         to the opposite. Fastest full waves were recorded and the wavers
         were ranked. Jovan was the fastest.

         Jovan je   mah-a/nu-o   za sekund.
         J   AUX wav-TH-PTCP.IPFV/PFV for second
         'Jovan managed to wave in a second (on at least / exactly one
         occasion).'

Classes c) and d) above are hence excluded on different grounds, either due to not
licensing atomic conceptualization, or due to specifying atoms whose temporal
trace cannot be non-trivially measured.

32

The other question is how these verbs when combined with the *za*-phrase receive the interpretation of a general-factual imperfective distributed over a plural reference time. The issue is even more striking in light of the observation that this combination cannot have a progressive interpretation (denoting the process stage of an ongoing event of, e.g., running to school in 10 minutes). I argue that the same explanation holds for simple imperfectives that applies to the secondary ones, which show the same pattern of behavior. Namely, on the progressive interpretation, the sentences in (19b) and (19c) exemplify the imperfective paradox, as at the epistemic evaluation time it can only be verified that the event of Jovan running to school is taking place, but not how long it will take to completion, or even that it will be completed. The progressive readings are degraded exactly because the speaker cannot know the duration of an event before its completion (i.e. the speaker cannot describe an incomplete event in terms of an event kind resorting to the temporal duration of completed events). They are hence not grammatically unavailable, but rather pragmatically blocked.

The blocking above looks like the imperfective paradox, where too the issue is that an event is described including a result, yet at the time of epistemic evaluation it is impossible to evaluate whether the result obtains. The difference is likely in the fact that the result in the relevant cases is pragmatically established as a plausible defining property of a natural class, while the duration expressed by the *za*-phrase, with an infinite range of possible measures – each standing for a different natural class, is not. As pointed by Olav Mueller Reichau (p.c.), for predicates including the *za*-phrase that do match an established natural kind, such as e.g. God's creation in seven days, the progressive interpretation becomes available.

The second observation that supports the universal simple analysis of simple imperfectives concerns the status of the result, i.e. goal predicate. Recall that in §2, I have shown that the semantic specification of the result at the level of the $\sqrt{P}$ does not suffice to derive singularity (i.e. telicity), and moreover that without the agreement of the Q^0 with the result predicate, the result predicate is not bound by the speech act predicate (i.e. it is not asserted in assertions). Furthermore, it was shown that semantic effects of result agreement are preserved after reverbalization (i.e. secondary imperfectivization), in spite of the grinding effect of reverbalization, arguably due to the pragmatic competition with the respective simple imperfective.

Similar asymmetry can be observed between simple imperfectives compatible with the *za*-phrase and secondary imperfectives. Consider the examples in (21).

Boban Arsenijević

(21) a. Jovan je trč-*u*a-o   u školu (dužim putem).
         J   AUX run-TH-PTCP.IPFV in school longer way
         'Jovan was running to school (the longer way).'

      b. Jovan je u-trč-*u*a-*u*a-o   u školu (#dužim putem).
         J   AUX in-run-TH-TH-PTCP.IPFV in school   longer way
         'Jovan was entering the school running (#the longer way).'

Without the path modifier, example (21a) with a simple imperfective can mean
the same as (21b), which includes a secondary imperfective. This again at first
glance supports the null prefix analysis, under the assumption of full composi-
tionality. However, with the adverbial modifying the path, the sentence with a
secondary imperfective is pragmatically ill-formed, while the one with a simple
imperfective is fine. This is the case because the secondary imperfective on the
progressive interpretation tends to refer to the narrow phase transition to the re-
sult state (i.e. from Jovan being outside the school to him being inside the school),
and the path of this transition is conceptualized as a point in space, which cannot
be modified for length (even on the slow motion / temporal zooming in interpre-
tation licensing the progressive). The simple imperfective rather refers to the
preparatory stage, i.e. to the motion event leading to the phase transition.[10]

Irrespective of the analysis of aspectual morphology, the asymmetry in (21)
argues for different syntactic structures and types of meanings for simple and
secondary imperfectives. Simple imperfectives have no QP, and secondary im-
perfectives embed one. The latter fact restricts their denotation to sums of parts
of events involving the specified result (in the given case, to parts of the event of
switching from being outside to being inside the school). In light of the analysis
proposed in §1.2, where the aspectual semantic restrictions of AS and AU verbs
are largely pragmatically determined, with an important role played by the con-
trasts between aspectual pairs – the fact that verbs of both classes have prefixed
perfective aspectual partners even more clearly implies that their compositional
semantics, and hence also their structures are different. A plausible difference
suggested by their morphology is that secondary imperfectives do involve a QP
and a reverbalizing secondary *v*P, while simple imperfectives never do.

---

[10]Secondary imperfectives involving goal/result- and source-oriented prefixes show the effect
of shrinking to the point of phase transition. Those with path-oriented prefixes do not, as
illustrated in (i).

 (i) Marija je uz-trč-*u*a-*u*a-la   uz   Rtanj (dužim putem).
      M   AUX up_along-run-TH-TH-PTCP.IPFV up_along Rtanj longer way
      'Marija was running up the mountain Rtanj (the longer way).'

Evidence provided in this section thus supports the view in which simple imperfectives compatible with the *za*-phrase are not telic and do not embed a telic structure. Consequently, they do not require the positing of null prefixes.

6 Conclusion

The starting point of the investigation was the strong generalization from Ła-zorczyk (2010) that in Slavic languages, telicity is necessarily marked by an affix, and that affixless verbs which show telic behavior involve null prefixes. The main question tackled by the paper was whether the strong generalization can be maintained without the introduction of null prefixes, i.e. whether the empirical data renders null prefixes necessary to maintain the hypothesis. On the material from SC, I argued that neither of the affixless verb classes showing (aspects of) telic behavior involve null prefixes. More precisely: proper simple perfectives are all idiomatized and stored in the lexicon with a non-compositional telic meaning. Affixless imperfectives compatible with the *za*-phrase do not show true telic behavior, and do not embed the structure corresponding to a telic eventuality. This simplifies the model by eliminating null prefixes, while still preserving the strong generalization about affixes and telicity. I presented morphological and semantic asymmetries, as well as quantitative corpus-based evidence in support of this view.

Abbreviations

AUX	auxiliary		PTCP	participle
DAT	dative		PFV	perfective
GEN	genitive		TH	theme vowel
IPFV	imperfective		√	root
INF	infinitive		SG	singular

Acknowledgements

Boban Arsenijević

References

Arsenijević, Boban. 2006. *Inner aspect and telicity*. Utrecht: LOT Publication.

Arsenijević, Boban. 2007a. A unified analysis of two classes of Slavic verb prefixes. In Sylvia Blaho (ed.), *Console XIV proceedings*, 21–36. Leiden: Student Organisation of Linguistics in Europe.

Arsenijević, Boban. 2007b. Slavic verb prefixes are resultative. *Cahiers Chronos* 17. 197–213. DOI: 10.1163/9789401204446_013.

Arsenijević, Boban. 2018. Serbo-Croatian suffix −*va* and the illusion of the imperfective specification. presented at the conference AMBIGO, held at the University of Göttingen. https://rb.gy/iqblpq.

Arsenijević, Boban, Katarina Gomboc Čeh, Franc Marušič, Stefan Milosavljević, Petra Mišmaš, Jelena Simić, Marko Simonović & Rok Žaucer. in preparation. *WeSoSlaV: Database of the Western South Slavic verbal system*.

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9(1). 5–16.

Bennett, Michael & Barbara Partee. 1972. *Toward the logic of tense and aspect in English*. Santa Monica, CA: System Development Corporation. DOI: 10.1002/9780470751305.ch4.

Borer, Hagit. 2005. *Structuring sense volume 2: The normal course of events*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199263929.001.0001.

Borik, Olga. 2006. *Aspect and reference time*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199291298.001.0001.

Braginsky, Pavel & Susan Rothstein. 2008. Vendlerian classes and Russian aspectual system. *Journal of Slavic Linguistics* 16(1). 1–55. DOI: 10.1353/jsl.0.0009.

Chierchia, Genaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6(4). 339–405. DOI: 10.1023/A:1008324218506.

FidaPLUS. 2000. *Slovenian national corpus fidaplus*. https://www.sketchengine.eu/fida-plus-corpus/.

Filip, Hana & Susan Rothstein. 2005. Telicity as a semantic parameter. In James Lavine, Steven Franks, Hana Filip & Mila Tasseva-Kurktchieva (eds.), *Formal approaches to Slavic linguistics (FASL 14): The Princeton University meeting*, 139–156. Ann Arbor, MI: Michigan Slavic Publications.

Fleischhauer, Jens & Ekaterina Gabrovska. 2019. Chapter 5 Perfectivity and atelicity: The role of perfective aspect in aspectual composition. In *Cross-linguistic perspectives on the semantics of grammatical aspect*. Leiden: Brill. 97–126. DOI: 10.1163/9789004401006_006.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20*. Ken Hale & Samuel J. Keyser (eds.). Cambridge, MA: The MIT Press. 111–176.

Horn, Laurence R. 1989. *A natural history of negation.* Chicago, IL: University of Chicago Press.

Isačenko, Alexander V. 1960. *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim – čast' vtoraja: morfologija.* Bratislava: Izdatel'stvo akademii nauk.

Janda, Laura A. & Olga Lyashevskaya. 2011. Aspectual pairs in the Russian National Corpus. *Scando-Slavica.* 57(2). 201–215. DOI: 10.1080/00806765.2011.631780.

Krifka, Manfred. 1989. *Nominalreferenz und Zeitkonstitution: Zur Semantik von Massentermen, Pluraltermen und Aspektklassen.* München, Wilhelm Fink.

Łazorczyk, Agnieszka Agata. 2010. *Decomposing Slavic aspect: The role of aspectual morphology in polish and other Slavic languages.* Los Angeles, CA: University of Southern California. (Doctoral dissertation).

Ljubešić, Nikola & Tomaž Erjavec. 2011. HrWaC and slWaC: Compiling web corpora for Croatian and Slovene. In Ivan Habernal & Václav Matoušek (eds.), *Text, speech and dialogue*, 395–402. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-642-23538-2_50.

MacDonald, Jonathan E. 2008. *Syntactic nature of inner aspect: A minimalist perspective.* Amsterdam: John Benjamins Publishing Company. DOI: 10.1075/la.133.

Maslov, Jurij S. 1948. Vid i leksičeskoe značenie glagola v sovremennom russkom literaturnom jazyke. *Izvestija Akademii nauk SSSR—Otdelenie literatury i jazyka* 7(4). 303–316.

Milosavljević, Stefan. 2023. The sequence of similar events interpretation, simple imperfective forms, and Slavic aspectual composition. In *this volume.* Language Science Press.

Milosavljević, Stefan. in preparation. *Specification of event duration and aspectual composition in Slavic.* Graz: University of Graz. (Doctoral dissertation).

Mittwoch, Anita. 2010. Event measurement and containment, syntax, and event structure. In *Lexical semantics, syntax, and event structure.* Malka Rappaport Hovav, Edit Doron & Ivy Sichel (eds.). Oxford: Oxford University Press. 252–266. DOI: 10.1093/acprof:oso/9780199544325.003.0012.

Mittwoch, Anita. 2013. On the Criteria for distinguishing accomplishments from activities, and two types of aspectual misfits. In *Studies in the composition and decomposition of event predicates.* Boban Arsenijević, Berit Gehrke & Rafael Marín (eds.). Dordrecht: Springer. 27–48. DOI: 10.1007/978-94-007-5983-1_2.

Pereltsvaig, Asya. 2000. On accusative adverbials in Russian and Finnish. In *Adverbs and adjunction.* Artemis Alexiadou & Peter Svenonius (eds.). (Linguistics in Potsdam 6). Potsdam: Universitätsverlag Potsdam. 155–176.

Boban Arsenijević

Plag, Ingo. 2012. *Morphological productivity: Structural constraints in English derivation*. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110802863.

Quaglia, Stefano, Boban Arsenijević, Marko Simonović & Svitlana Antonyuk. under review. Allomorphy, morphological operations and the order of Slavic verb-prefixes. *Journal of Slavic Linguistics*.

Ramchand, Gillian. 2004. Time and the event: The semantics of Russian prefixes. *Nordlyd* 32(2). 323–366. DOI: 10.7557/12.72.

Šaxmatov, Aleksey A. 1941. *Sintaksis russkogo jayzka*. Moscow: Učpedgiz.

Simonović, Marko, Stefan Milosavljević & Boban Arsenijević. 2021. Serbocroatian secondary imperfectivisers consist of theme vowels. forthcoming.

Stanojević, Veran. 2012. Perfektivnost i teličnost u srpskom jeziku sa osvrtom na situaciju u francuskom jeziku. *Naučni sastanak slavista u Vukove dane* 41(1). 141–155.

Starke, Michal. 2010. Nanosyntax: a short primer to a new approach to language. *Wiener Linguistische Gazette* 1(36). DOI: 10.7557/12.213.

Švedova, Nataľja Ju. (ed.). 1980. *Russkaja grammatika*. Moscow: Izd. Nauka.

Svenonius, Peter. 2004. Slavic prefixes inside and outside VP. In Peter Svenonius (ed.), *Nordlyd 32(2): Slavic prefixes*, 205–253. University of Tromsø. DOI: 10.7557/12.68.

Szucsich, Luka. 2001. Adjunct positions of nominal adverbials in Russian. In *Current issues in formal Slavic linguistics*. Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn & Luka Szucsich (eds.). (Linguistik International 5). Frankfurt: Peter Lang. 106–116.

Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press.

Toporišič, J. 2000. *Slovenska slovnica, 4. izdaja*. Maribor: Založba Obzorja. DOI: 10.1093/acprof:oso/9780199263929.001.0001.

Travis, Lisa. 2012. Arguments From The Root Vs. Arguments From The Syntax. In *The End of Argument Structure*. María Cristina Cuervo & Yves Roberge (eds.). Leiden, The Netherlands: Brill. 261–291. DOI: 10.1163/9781780523774_012.

Verkuyl, H.J. 1972. *On the compositional nature of the aspects* (Foundations of Language Supplementary Series). Springer Netherlands.

Vinogradov, Viktor V. 1938. *Sovremennyj russkij jazyk. grammatičeskoe učenie o slove*. Moscow: Učpedgiz.

Žaucer, Rok. 2009. *A VP-internal/resultative analysis of 4 "VP-external" uses of Slavic verbal prefixes*. Ottawa, ON: University of Ottawa. (Doctoral dissertation).

Chapter 2

Doubling in South Slavic relative clauses and the predictability of morphosyntactic features

Julia Bacskai-Atkari

University of Amsterdam / University of Potsdam

The paper investigates the morphosyntactic properties of relative markers in South Slavic. In Slavic languages, just like in many other European languages, relative clauses can be introduced by two kinds of relative markers: (i) relative complementisers, which are invariant in their form, and (ii) relative pronouns, which are inflected (for case, number, and gender, depending on the language). Slavic languages regularly use wh-based complementisers and/or pronouns. Crucially, the two cannot co-occur: this ban is not grounded in the syntactic structure per se, but it derives from the feature incompatibility of two wh-based relative markers, as these are regularly equipped with an uninterpretable relative feature. The only exception is Macedonian: in this case, however, there is independent evidence for the complementiser to have a different feature makeup, suggesting that while morphological properties are good predictors for the relevant syntactic constraints, they are not deterministic.

Keywords: demonstrative pronoun, feature checking, finiteness, inflection, interrogative clause, relative clause

1 Introduction

There are various elements that can overtly mark and introduce relative clauses; two examples from English are given in (1) below:

(1)   a.   This is the problem *which* we should solve first.

Julia Bacskai-Atkari

    b. This it the problem *that* we should solve first.

On the one hand, there are differences in the etymology (cf. Hopper & Traugott 1993, Heine & Kuteva 2002): relative markers can be interrogative-based, as *which* in (1a) above (also: *who(m)*, *whose* etc.), or demonstrative-based, as *that* in (1b) above.

    On the other hand, there are differences in the position of these elements: relative markers can be relative pronouns as the interrogative-based English pronouns *which*, *who(m)* etc. and the demonstrative-based German pronouns *der*/*die*/*das* etc., or they can be relative complementisers, as the demonstrative-based English *that* and the interrogative-based South German *wo* (cf. Bayer 1984, Salzmann 2006, 2017, Brandner & Bräuning 2013, Weiß 2013). Given the positional differences, it is not surprising that doubling patterns consisting of an overt relative operator and an overt relative complementiser are attested, as illustrated in (2):

(2)   % This is the problem *which that* we should solve first.

As indicated (%), this pattern is not accepted in all varieties of English (it is, for instance, excluded from the standard variety).

    Regarding Germanic, Bacskai-Atkari (2020) made the observation that while overt relative pronouns and overt relative complementisers can be combined, these combinations appear to be restricted by the etymology, in that only asymmetric combinations are attested as genuine REL+REL combinations; that is, as combinations where both elements are attested as relative markers on their own as well.[1] This observation raises several questions. First, it should be clarified how strong the generalisation is cross-linguistically: in this article, I am going to examine Slavic data in this respect, as Slavic languages are known to have the various kinds of relative markers mentioned above. Consider the following examples from Bosnian-Croatian-Serbian (henceforth: BCS):

(3)   a. čovjek *što* puši
             man   that smokes
             'a/the man that smokes/is smoking'   (Gračanin-Yuksek 2013: 27)

    b. čovjek *koji*   puši
             man   which.NOM smokes
             'a/the man who smokes/is smoking'   (Gračanin-Yuksek 2013: 26)

---

[1]As will be discussed in §2, this is not merely the result of what items are available. Both in English and in German, wh-based pronouns are available; in addition, both of these languages have varieties where wh-based complementisers are attested. Nevertheless, WH+WH combinations are not attested in these varieties either.

40

The relative clause is introduced by the complementiser *što* in (3a) and by the relative pronoun *koji* (inflected for case) in (3b). Both of these elements are wh-based: as will be discussed in §3, this is the regular Slavic pattern (see Auderset 2020 for typological insights). The relevance of this pattern for testing the validity of the above-mentioned hypothesis is clear: While Germanic languages tend to have asymmetric patterns due to the availability of demonstrative-based relative markers, the wh-based Slavic patterns may provide us insights into whether the lack of wh+wh patterns is systematic or rather coincidental in nature.

Second, the question arises how apparently excluded combinations can be analysed synchronically: while pointing to the etymology may be satisfactory for descriptive purposes, it is highly unlikely that it can be taken as a grammatical constraint per se. In this article, I will argue that the etymological differences correspond to differences formulated in terms of morphosyntactic features.

Third, related to this, the question arises what independent evidence we have for the featural properties of individual elements. Without such independent evidence, simply translating etymological differences into features would again amount to mere descriptive adequacy. The present paper argues that the combinations are restricted by the distribution of [rel] features that are ultimately determined by the etymology but they can show subsequent deviations.

The paper is structured as follows. In §2, I am going to briefly discuss the observations for Germanic. In §3, I will present the data from (South) Slavic and I will provide an analysis for the doubling patterns in §4.

2 Germanic

In Germanic languages, we can observe doubly filled COMP effects involving an overt pronoun and an overt complementiser; these can be assigned the schematic structure shown in Figure 1.[2]

The combination of a wh-pronoun and a d-complementiser can be observed in non-standard varieties of English (see van Gelderen 2009) and marginally also in Swedish, as shown by the data in (4):

---

[2]I adopt a single CP analysis for doubling in relative clauses, following Bacskai-Atkari (2020); under this view, there are no designated projections for left-peripheral elements, unlike in cartographic approaches (going back to Rizzi 1997). Note also that while doubling is attested in these languages, it is altogether not very frequent (unlike in embedded interrogatives, where doubly filled COMP effects are widely attested). Bacskai-Atkari (2022) attributes this to discourse factors: the relative pronoun is essentially redundant (at least when the relative complementiser is overt).

Julia Bacskai-Atkari

```
                    CP
                   /  \
              which    C′
                      /  \
                     C    TP
                     |
                    that
```
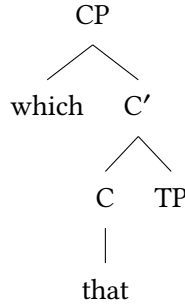
Figure 1: The structure of doubly filled COMP

(4) a. It's down to the community *in which that* the people live.

(van Gelderen 2013: 59)

b. Detta är studenten   *vilken som* bjöd   in Mary.
   this   is the.student which that invited in Mary
   'This is the student who invited Mary.'

(Bacskai-Atkari & Baudisch 2018: 247)

The combination of a d-pronoun and a wh-complementiser can be observed in South German dialects (Brandner & Bräuning 2013, Weiß 2013, Fleischer 2017), illustrated for Hessian and for (North) Bavarian in (5a) and in (5b), respectively:[3]

(5) a. Des Geld, *des   wo* ich verdiene, des   geheert mir.
      the.n money that.n REL I   earn.1SG that.n belongs I.DAT
      'The money that I earn belongs to me.'   (Fleischer 2017)

b. Mei Häusl   (…), *dös   wos* dorten unten (…) steht
   my house.DIM   that.n REL there below   stands
   'My little house, which stands down there'   (Weiß 2013: 780)

Given the differences between elements related to position and etymology, there are four logically possible configurations; out of these, only two are attested as genuine REL+REL combinations (that is, where both members are independently and productively attested as relative markers). This is shown in Table 1.

While the asymmetric combinations are straightforward, the D+D combination is at least questionable. On the surface, this kind of combination is attested in Waasland Dutch (Boef 2013), as shown in (6).

---

[3]In these varieties, the wh-based complementisers also regularly introduce relative clauses on their own. The complementiser *wo* has a wider distribution geographically; note that it is not used as a declarative complementiser or as a mere finiteness marker.

Table 1: Combinations of genuine relative markers

	d-complementiser	wh-complementiser
d-pronoun	−/??	+
wh-pronoun	+	−

(6) Dat is de man *die dat* het verhaal verteld heeft.
    that is the man who that the story told has
    'That is the man who has done it.'   (Boef 2008: 93)

In this case, however, it is very probable that the combination cannot be considered as genuine REL+REL. In Dutch, relative clauses introduced by a single *dat* (as a complementiser) are found in Vlaams-Brabant Dutch (Boef 2013) and thus not in the same area where the doubling pattern is attested: in the doubling pattern in (6), then, the complementiser marks finiteness, not [rel].[4]

In other words, there is no strong evidence for the existence of genuine D+D doubling. More importantly, no combinations of the form "wh-pronoun + wh-complementiser" are attested (even though they would be logically possible in certain varieties, such as in English with the complementiser *what* and in South German with the complementisers *wo* and *was*).

---

[4]The availability of *dat* as a finiteness marker is also independently motivated: it is also attested in embedded constituent questions across Dutch dialects, that is, in environments where it cannot be a declarative complementiser (see Schallert et al. 2018 for a recent discussion). Another potential counterexample to the generalisation in Table 1 comes from Old English (see van Gelderen 2009), as illustrated below:

(i) ac gif we asmeagaþ þa   eadmodlican dæda þa þe he worhte, þonne ne þincþ
    but if we consider those humble   deeds that that he wrought then not seems
    us þæt nan wundor
    us that no wonder
    'But if we consider the humble deeds which he wrought, that will seem no wonder to
    us.'   (*Blickling Homilies* 33; Watanabe 2009: 364, citing Allen 1980)

In Old English, we find the above doubling pattern as an intermediate stage in the process of reanalysis of one of the d-pronouns (*that*) into a complementiser, removing the original complementiser þe (van Gelderen 2009): this suggests that þe was possibly only a finiteness marker, or that the pronoun was initially still a demonstrative but not [rel]. This (and the Waasland Dutch pattern) crucially differs from the present-day English pattern, where *that*-relatives are common and productive: in other words, there is no reason to assume that patterns like (4a) would involve a mere finiteness marker.

Julia Bacskai-Atkari

3 The data

3.1 Relative markers in South Slavic

South Slavic languages are particularly interesting regarding the above generalisation, since these languages regularly use wh-based elements (cf. Kljajevic 2012: 36, Auderset 2020) as relative markers. In addition, both major strategies (that is, pronouns versus complementisers) are attested in (South) Slavic languages.

Consider again the examples from BCS in (3), repeated here for the sake of convenience as (7):

(7)   a. čovjek *što* puši
        man   that smokes
        'a/the man that smokes/is smoking'   (Gračanin-Yuksek 2013: 27)

      b. čovjek *koji*   puši
        man   which.M.NOM smokes
        'a/the man who smokes/is smoking'   (Gračanin-Yuksek 2013: 26)

In (7a), the relative clause is introduced by the complementiser *što*; in (7b), it is introduced by the relative pronoun *koji*, which is, unlike the complementiser, inflected for case. This becomes evident if we compare the elements above, which occur in subject relative clauses, to their counterparts in direct object relative clauses, as shown in (8a) and (8b), and in indirect object relative clauses, as shown in (8c) and (8d):

(8)   a. čovjek *što* ga   Jan vidi
        man   that 3SG.ACC.CL Jan sees
        'a/the man who Jan sees'   (Gračanin-Yuksek 2013: 27)

      b. čovjek *kojeg*   Jan vidi
        man   which.M.ACC Jan sees
        'a/the man who Jan sees'   (Gračanin-Yuksek 2013: 27)

      c. čovjek *što* mu   Jan   pokazuje put
        man.NOM that 3SG.DAT.CL Jan.NOM shows   way.ACC
        'a/the man to whom Jan shows/is showing the way'
                      (Gračanin-Yuksek 2013: 27)

      d. čovjek *kojem*   Jan   pokazuje put
        man.NOM which.M.DAT Jan.NOM shows   way.ACC
        'a/the man to whom Jan shows/is showing the way'
                      (Gračanin-Yuksek 2013: 27)

As can be seen, while *što* does not change its form, the relative pronoun is in-
flected for accusative and dative case.[5] Importantly, relative operators are phono-
logically identical to their interrogative counterparts (also inflected for case, num-
ber and gender); *što* is phonologically identical to the most unmarked interroga-
tive form (nominative/accusative; the dative would be *čèmu*). The interrogative
patterns are illustrated in (9) below:

(9)   a. *Što*   je Marija videla?
          what.ACC AUX Mary   seen
          'What did Mary see?'   (Halpern 1995: 77)

      b. *Koji*   čovek je   voleo Mariju?
          which.M.NOM man   AUX seen Mary.ACC
          'Which man saw Mary?'   (Halpern 1995: 78)

      c. *Koju*   žabu   je lane liznulo?
          which.F.ACC frog.ACC AUX fawn lick.PTCP
          'Which frog did the fawn lick?'   (Kljajevic 2012: 34)

The syntactic positions of the relevant elements are illustrated in Figure 2 and
Figure 3. We can observe the same variation between complementisers and pro-
nouns in Macedonian, as shown in (10).

(10) a. Covekot   *koj*   vleze   e moj   sosed.
          man.the.M.SG who.M.SG come.AOR.3SG is my.M.SG neighbour
          'The man who came in is my neighbour.'   (Bužarovska 2009: 232)

      b. Covekot   *što* go   sretnavme e moj   sosed.
          man.the.M.SG that 3SG.ACC.CL meet.AOR.1PL is my.M.SG neighbour
          'The man whom we met is my neighbour.'   (Bužarovska 2009: 232)

---

[5]Note also another difference between the two strategies in (8), which cannot be seen in (7):
the direct object and the indirect object relative clauses with *što* contain a resumptive pronoun
(*ga* and *mu*, respectively), while this is not the case in the counterparts containing the relative
pronoun. Resumptive pronouns are used to lexicalise the gap in certain languages: since in
this respect they are similar to relative pronouns, it is actually expected that they should not
co-occur with the relative pronoun while they can (and in the given cases, must, see Gračanin-
Yuksek 2013: 27) surface when the relative clause is introduced by a complementiser. In this
respect, the presence/absence of resumptive pronouns in (8) is yet another indicator for the
structural difference between the relative markers under scrutiny. Note that the absence of
resumptive pronouns in subject relative clauses is also expected: resumptive pronouns are
more likely to occur in functions that are lower in the so-called Noun Phrase Accessibility
Hierarchy , and subjects constitute the highest function, so that the use of resumptive pronouns
in this function is extremely rare cross-linguistically (Keenan & Comrie 1977).

Julia Bacskai-Atkari
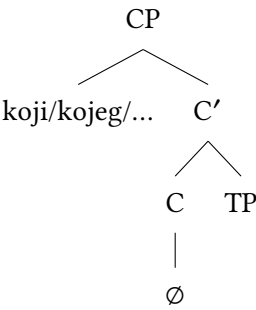
Figure 2: The position of relative pronouns in Slavic
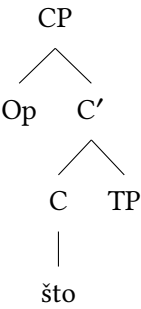
Figure 3: The position of relative complementisers in Slavic

Again, both elements are interrogative-based. This is illustrated in (11) below:

(11) a. *Što*   jade deteto?
          what eats child.the
          'What does the child eat?'   (Lazarova-Nikovska 2013: 134)

      b. *Koj*   te   potseti?
          who.cl 2sg.acc.cl reminded.3sg.perf.prs
          'Who reminded you?'   (Tomić 2006)

Slovene also makes use of both strategies, as illustrated in (12):

(12) a. Poznam človeka, *katerega* so   iskali.
          know.1sg man.acc which.acc aux.3pl looked.for
          'I know the man who they were looking for.'   (Hladnik 2010: 10)

      b. Poznam človeka, *ki* so   ga   iskali.
          know.1sg man.acc that aux.3pl m.acc.cl looked.for
          'I know the man that they were looking for.'   (Hladnik 2010: 10)

46

The relative pronoun is inflected and it is obviously a wh-based element (Mitrović 2016: 225); the complementiser *ki* lacks an interrogative counterpart in the modern language (Mitrović 2016: 225) but it derives from Proto-Indo-European **kʷís* 'who, what' and Slovene *ki* developed into an interrogative complementiser after the 14th century (Mitrović 2016: 225). As Hladnik (2010: 38) notes, citing Cazinkić (2001), *ki* is often perceived to be a reduced form of the relative pronoun, which is etymologically wrong. Further, prescriptive rules favour the pronoun strategy over the complementiser strategy (Hladnik 2010: 38): this is in fact reminiscent of the situation in West Germanic.

3.2 A note on Bulgarian

Bulgarian represents a special case within South Slavic regarding relative markers. Both strategies (the pronoun strategy and the complementiser strategy) can be observed in Bulgarian, with the colloquial complementiser *deto* (Rudin 2014) and with regular relative pronouns, as shown by the corpus examples taken from Bužarovska (2009) in (13):

(13) a. Imaše   xora, *koito* ne viždaxa   ništo   pred
          have.IMPERF.3SG people who.PL not see.IMPERF.3SG nothing before
          sebe si.
          own CL
          'There were people who saw nothing in front of them.'

                                                        (Bužarovska 2009: 249)

  b. Da bjaxa   mi   kazali,   če ima xora, *deto* bjagat
      SM be.PL.IMPERF 1SG.DAT.CL told.PL.PART that has people that run.3PL
      ot   dobroto   kato zajci ot   kopoj...
      from good.the.N.SG like rabbits from hound.M.SG
      'If I had been told that there are people who run away from good like
      rabbits from a hound...'   (Bužarovska 2009: 249)

The relative operator is evidently wh-based; as for *deto*, it also goes back to an interrogative operator (Bužarovska 2009: 234; see Krapova 2010: 1241 for a more detailed analysis) and, as mentioned above, it counts as colloquial, reminiscent of the prescriptive preferences for relative pronouns in Slovene and in West Germanic.

Note that the situation in Bulgarian is in fact somewhat more complex, as wh-pronouns in relative pronouns are apparently complex: *kojto* consists of the wh-base *koj* and the element *-to* (this pattern is productive, e.g. *kakvo-to* 'what'

or *kolko-to* 'how much'), whereby the status of *-to* is subject to much debate, as discussed by Rudin (2014) in detail. The most important question in this respect is whether the combination is primarily syntactic (involving distinct syntactic positions) or morphological (involving a single syntactic node). Unlike *što*, *-to* is not available as a complementiser in other constructions and it does not resemble a wh-element either (Rudin 2014: 322). Rudin (2009) analyses this element as a specifically relative complementiser: in this case, Bulgarian would in fact show doubling, but note that as *-to* is not a wh-based element, this does not go against the generalisation under scrutiny here, i.e. that wh+wh combinations are regularly not attested; further, *-to* is not available as a relative marker on its own, so that a genuine rel+rel doubling pattern would not arise either. Rudin (2014: 324) remarks that the complementiser approach faces problems with complex wh-phrases such as *kolkoto goljam* 'how big', where *-to* appears to be incorporated into the wh-phrase. According to Rudin (2014), a further problem lies in the fact that the complementiser account would predict more parallelism with *što*, which is problematic as e.g. *što* in Macedonian is banned from comparatives but Bulgarian *-to* is not. This is, however, not a strong counterargument: as argued by the present paper, relative complementisers may show different behaviour (and distribution) due to their different featural properties; further, relative complementisers appearing in comparatives show considerable variation, and *što* is in fact available in comparatives in BCS (see Bacskai-Atkari 2016 for discussion). Other analyses include treating *-to* as a definiteness marker (e.g. Izvorski 2000; see Rudin 2009 and Rudin 2014: 322–323 for counterarguments) or as a morphological marker of relative pronouns (Hauge 1999, see Rudin 2014: 325 for some concerns): in these cases, however, no complex left periphery is involved and these account would again not be problematic for the issues discussed in the present paper. For this reason, Bulgarian *-to* will not be discussed in §4.

3.3 Interim summary and outlook

In sum, it is evident that South Slavic languages by default show variation between the relative complementiser strategy and the relative operator strategy. It is worth mentioning that this kind of variation is not restricted to South Slavic but can be more generally observed across Slavic languages, though the exact distribution and acceptability patterns differ.

In West Slavic, the standard option seems to be the use of relative pronouns, but once non-standard varieties are also taken into account, we can also find relative complementisers in these languages, i.e. Czech and Polish *co*, Slovak *čo* (Šimík 2008, Guz 2017, Minlos 2012).

In East Slavic, both relative pronouns and relative complementisers are attested: while Russian *čto* is a markedly colloquial option (Meyer 2017), Ukrainian and Belarusian *ščo* seems to be more widespread (Danylenko 2018).

In other words, the variation between the relative complementiser strategy and the relative operator strategy is not restricted to South Slavic languages but can be found more generally in Slavic languages. The complementiser strategy is overall more restricted; South Slavic seems to offer the best testing ground for potential wh+wh combinations. For this reason, I am going to restrict myself to the discussion of South Slavic data in the discussion to follow.

4 Doubling

4.1 A note on features

I adopt standard minimalist assumptions regarding formal features, going back to Chomsky (1995); see also Zeijlstra (2014). According to this, the kind of features that can participate in morphosyntactic operations are called formal features: this set of features intersects with semantic features. Interpretable formal features are in the intersection; uninterpretable features are pure formal features (they cannot be interpreted at LF) and need to be checked off (or, in more recent terms, valued); this can be done via a matching interpretable feature. Note that the presence of any uninterpretable feature, [u-F], on a certain element implies only that the particular feature is not interpretable on that given element in LF and it does not imply in any way that the given element would lack other semantic features (or meaning).

4.2 The analysis of doubling patterns

As mentioned in §2, doubling patterns appear to be asymmetric; this observation led Bacskai-Atkari (2020) to the hypothesis that the observed differences may be due to differences in the interpretability of [rel] features. According to this, we should have the following distribution: d-pronouns and d-complementisers are [i-rel] and wh-pronouns and wh-complementisers are [u-rel].[6]

---

[6]One might wonder why this should be so: so far, this hypothesis gives the right empirical predictions, yet it would be desirable to detect more general properties behind the particular feature distribution. As far as Germanic is concerned, it is evident that demonstrative-based elements constitute the older strategy (see Ringe & Taylor 2014: 467 for Old English *þe* and Axel-Tober 2017: 46 for Old High German *the*): wh-based elements were introduced later into headed relative clauses, via analogy (from free relatives and interrogatives). Apart from this,

Julia Bacskai-Atkari

At any rate, the asymmetric patterns ensure proper feature checking, as shown in Figure 4 and Figure 5. In both configurations, the uninterpretable feature is properly checked off by its interpretable counterpart. By contrast, symmetric patterns are essentially problematic for feature checking. In the case of two [i-rel] features, the movement of the operator is not motivated; in the case of two [u-rel] features, the uninterpretable feature cannot be checked off. Relative complementisers regularly encode finiteness, [fin].

CP
which[u-rel]   C′
            C[i-rel],[fin]   TP
            that[i-rel],[fin]
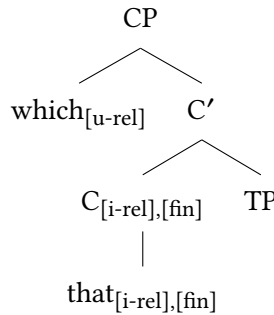
Figure 4: Features in WH+D

Regarding the former, we observed in §3 that some d-pronoun + d-complementiser combinations seem to exist, even though they were classed as not genuine. In the case of Waasland Dutch, the complementiser *dat* marks finiteness, and is thus underspecified for [rel]. In the case of Old English, *þe* was in the process of losing its [i-rel] specification, ultimately changing into being underspecified for [rel] and marking finiteness only, similarly to the Waasland Dutch combination.[7] This suggests that D+D patterns can be accounted for in this model: an

---

note that the source elements differ in terms of definiteness features: demonstratives are definite, while the wh-base itself is indefinite (see Watanabe 2009, who also shows that the indefinite wh-base in English was also quantificational, making the clause into a complete proposition, which was incompatible with headed relatives). Relative pronouns are co-referential with the head noun under a matching analysis (cf. Salzmann 2017: 55–179) and definite pronouns are thus natural candidates as anaphors: indeed, the reanalysis of demonstrative markers into C-elements is traditionally considered to have evolved from paratactic structures involving a genuine demonstrative pronoun, since such examples are indeed possible and attested (but see Axel-Tober 2017 for a critical evaluation of this as the sole trigger of the relevant changes) – unlike with interrogative pronouns. In this sense, it is possible that the features [i-rel] and [u-rel] are ultimately related to the definite versus indefinite distinction, respectively. Future research will have to determine whether this idea is on the right track and, if so, how the diachronic feature inheritance can be modelled.

[7]Note that this does not make two projections necessary (i.e., one for clause type and one for

CP
der[i-rel]   C′
C[u-rel],[fin]   TP
wo[u-rel],[fin]
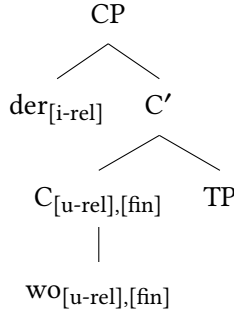
Figure 5: Features in D+WH

underspecified complementiser is used to lexicalise the complementiser and the abstract [u-rel] feature can be checked off regularly by the pronoun, as illustrated for Waasland Dutch in Figure 6.[8]

CP
die[i-rel]   C′
C[u-rel],[fin]   TP
dat[fin]
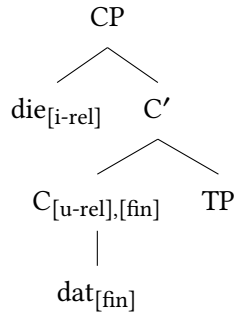
Figure 6: Doubling in Waasland Dutch

Note that there is independent evidence for the d-complementiser as underspecified for [rel]: the same complementiser appears in declaratives, where there

finiteness, as in cartographic approaches like that of Rizzi 1997 or Baltin 2010), as also shown by Bacskai-Atkari (2020) for embedded interrogatives. Intervening elements (which are often used as arguments for designated projections in cartographic approaches) are not attested in Germanic between clause-type markers (including finiteness markers).

[8]The mismatch between the underlying syntactic feature bundle and the inserted vocabulary item is in line with the core property of Distributed Morphology called Underspecification, according to which the inserted Vocabulary Items (the phonological expressions of abstract words) are not necessarily fully specified for the particular syntactic positions where they are inserted (see McGinnis-Archibald 2016: 401–405 for a summary; see Halle & Marantz 1994, Harley & Noyer 1999). This is a basic property of Late Insertion and it does not go against inclusiveness (Chomsky 1995: 225).

are no head nouns. One might wonder why lexicalising the (finite) C position is necessary: this seems to be a general tendency in Germanic (Bacskai-Atkari 2018, 2020) and it is not of further interest in this paper.

Crucially, the more problematic wh+wh patterns are not attested in Germanic. However, South Slavic shows variation here: while such combinations are not attested in BCS (Goodluck & Stojanović 1996: 292) and Slovenian (Hladnik 2010: 12–13), this pattern appears to be possible in Macedonian (Rudin 2014: 320). This is illustrated by the following example:

(14) čovekot *koj-što* zboruva
      the.man who-that talks
      'the man who is talking'   (Rudin 2014: 316)

The pattern in (14) seems to be productive: it is attested with all relative pronouns. The only exception is when the pronoun also has the form *što*, so that the sequence **što što* is ungrammatical (Rudin 2014: 320, citing Kramer 1999). This may well be a phonological constraint (and as such it is not direct evidence against the pronominal status of the second *što* element): as shown by Bošković (2002), similar constraints can be observed in multiple wh-fronting in Slavic languages.

Importantly, both relative markers in (14) are clearly interrogative-based, as their surface-identical counterparts are available as interrogative operators, as shown in (11) above and in (15) below:

(15) a. *Koj* zboruva?
          who talks
          'Who is talking?'   (Rudin 2014: 315)
      b. *Što* sakaš?
          what want.2sg
          'What do you want?'   (Rudin 2014: 320)

The data thus suggest that (14) apparently has a wh+wh pattern, which seems to contradict the hypothesis mentioned above. In order to determine to what extent (14) actually poses a problem for the theory, the distribution of the complementiser should be examined further. In Macedonian, *što* is also available as a declarative complementiser (Rudin 2014), as demonstrated in (16):

(16) Se   raduvam, *što* ve   gledam.
      REFL rejoice.1sg that you.PL.ACC see.1sg
      'I am happy that I see you.'   (Tomić 2006: 419)

This differs from the wh-based complementisers in Germanic, which may also be the reason for the differences regarding the doubling patterns in relative clauses. Regarding the status of *što* in relative clauses, Rudin (2014: 320) provides strong arguments that it should definitely taken to be a complementiser (contrary to Tomić 2012). First, the doubly filled COMP patterns such as (14) indicate that it cannot be a pronoun, as it appears in addition to the relative pronoun:[9] note that the word order constraint follows from the internal structure of the CP (Bacskai-Atkari 2018, 2020). Second, there is independent evidence for *što* being a complementiser otherwise, see (16) above. Third, prepositions cannot take relative *što* as a complement (the same applies to English *that*).[10]

Based on these observations, the structure in itself is not problematic, as it appears to demonstrate the same underlying syntax as the doubling patterns mentioned above and it can be derived from the structures in Figure 2 and Figure 3 in a straightforward way, as shown in Figure 7.
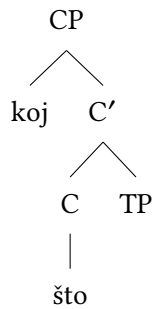
Figure 7: Doubling in Macedonian

The question is rather what the feature specification of *što* is. In essence, there are two possibilities: (i) underspecification for [rel], just like *dat* in Waasland Dutch, or (ii) specification as [i-rel].

---

[9]Unlike interrogative pronouns, which can co-occur in a single clause, there can only be a single relative pronoun in a relative clause: the head noun is co-referential with the relative pronoun, which can be base-generated only in a single position. See also Rudin (2014: 320).

[10]This is shown by the following example:

(i) * studentkata, *za   što* zboruvame
     student   about that speak.1PL
     Intended: 'the student about whom we speak'   (Rudin 2014: 320)

Rudin (2014: 320), citing Tomić (2012) and Kramer (1999), confirms that such patterns are impossible in relative clauses. Note that this of course does not imply anything about the interrogative pronoun *što* in questions.

Julia Bacskai-Atkari

Regarding the first hypothesis, we can establish the following. Underspecification in itself is plausible under a late insertion approach (Halle & Marantz 1993; see also the discussion in this section above), inasmuch as the abstract underlying head is lexicalised by a partial match (see Figure 6 for Waasland Dutch). This assumption is less problematic if the abstract head is [u-rel] than when it is [i-rel], since uninterpretable features are deleted anyway after check-off, so that Vocabulary Insertion taking place in the morphological component (after Spell-Out) does not actually see [u-rel]. The same argumentation does not follow automatically for [i-rel], though: leaving the C position in Macedonian as underspecified or as [u-rel] would require the relative pronoun to be specified as [i-rel], but there is no independent evidence for Macedonian wh-operators to be different from the general properties of wh-based relative markers, that is, creating an exception for wh-based relative pronouns in Macedonian as [i-rel] would be ad hoc.[11] In principle, this possibility cannot be excluded but making such an assumption without independent evidence would be merely descriptive at this stage.

On the other hand, however, we have independent evidence for *što* having different properties from the Germanic pattern. In the hypothesis formulated in (ii) above, *što* is [i-rel], which actually implies a difference from the Germanic pattern. There are two points of interest here. First, doubling patterns in Germanic are primarily attested in embedded interrogatives and much less in relative clauses (Bacskai-Atkari 2022), due to the lexicalisation preference on C: the same does not apply to Slavic. In other words, while both language groups may show doubling patterns in relative clauses, the underlying reasons are likely to be different and thus it cannot be expected that the two groups show parallel behaviour in all respects. Second, regarding the status of *što*, it should be noted that such relative declarative complementisers in South Slavic introduce factives and not all kinds of declarative clauses, unlike what we can observe in Germanic.

Consider the following examples from BCS:

(17) a. Jesam ti   rekao *da* je   Marija orišla na odmor?
         AUX.1SG you.DAT told   that AUX.3SG Marija gone on vacation
         'Did I tell you that Marija went on vacation?'   (Arsenijević 2020: 341)

      b. Jesam ti   rekao *što* je   Marija orišla na odmor?
         AUX.1SG you.DAT told   that AUX.3SG Marija gone on vacation
         'Did I tell you that Marija went on vacation?' (it is a fact that she did)
                                                        (Arsenijević 2020: 341)

---

[11]This crucially differs from the Dutch scenario, where the d-pronoun can regularly be assumed to have an [i-rel] specification, in line with the general hypothesis.

In (17a), the embedded clause is non-factive: it may or may not be true that Marija went on vacation. In (17b), however, the embedded clause is factive: this is the context where *što* can appear. As Arsenijević (2020) argues, *što*-declaratives have referential properties and are thus similar to relative clauses (see Krapova 2010: 1266 for Bulgarian and Macedonian and Bužarovska 2009 and Browne 1986: 69 for Macedonian; see also Aboh 2005 for factives being a special kind of relative clauses).[12] However, notice that there is no head noun and no relative operator movement in such configurations: this indicates that *što* cannot be [u-rel] in these constructions, as there would be no element to check off this feature. In other words, while the interrogative element can be assumed to have a regular [u-rel] feature, this feature is lost in factive declaratives.[13] This leads to the configuration shown in Figure 8.

Note that the loss of [u-rel] does not make [i-rel] automatically available on the inserted lexical items as an inherent property: in particular, there is no wʜ+wʜ doubling in ordinary relative clauses in BCS, so that there is no reason to assume that BCS in ordinary relative clauses would be [i-rel]. By contrast, we can observe wʜ+wʜ doubling in ordinary relative clauses in Macedonian, indicating that Macedonian *što* is available as [i-rel]. In this way, we can set up an implicational hierarchy: wh-based declaratives are a prerequisite for wʜ+wʜ doubling in ordinary relative clauses but not vice versa, that is, the existence of wh-based

---

[12]This may be related to the fact that *što*-relatives in BCS are used in relative clauses where the head noun is familiar (see Arsenijević 2020: 341–342). Note that the familiarity of the referent (as expressed by the head noun) does not equal definiteness on the relative pronoun, as familiarity and definiteness are distinct (though not unrelated) properties. Consider the following example:

    (i)   I saw a/the shopkeeper who was wearing a kilt.

In (i), the head noun is either indefinite or definite: this does not affect the relative marker (the pronoun *who*).

[13]Note that the similarities between (factive) declaratives and relative clauses does not make the two constructions equal. In particular, they differ in terms of operator movement, as shown by Arsenijević (2009). In (headed) relative clauses, the matrix correlate (the head noun) is co-referential with the relative pronoun, which is interpreted in the relativisation site (the base position) and in the CP-domain (the landing site): such elements undergo movement. By contrast, while Arsenijević (2009) assumes that there is also a matrix correlate in (factive) declaratives, the co-referential nominal element in the subordinate clause has its relativisation site at the top of their structure, that is, in the projection that specifies the illocutionary force of the clause. In other words, this configuration involves a higher projection site and no relative operator movement; consequently, the feature checking relation discussed in the present article does not apply.

Julia Bacskai-Atkari
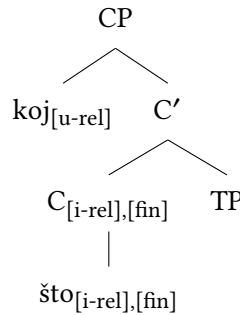
Figure 8: Features involved in doubling in Macedonian

declaratives does not imply the existence of wh+wh doubling in ordinary relative clauses.

5 Conclusion

In this paper, I examined doubling in South Slavic relative clauses, concentrating on the effects of the morphological inventory: crucially, both wh-pronouns and wh-complementisers are available in these languages. The typological predictions based on Germanic and Slavic are the following: (i) genuine wh+wh combinations are not attested, and (ii) the only exception is Macedonian, where the wh-complementiser *što* has different properties (as supported by independent evidence), indicating that further (featural) reanalysis is possible. This indicates that while morphological properties are decisive for most patterns, they do not prohibit further grammaticalisation even in languages where the original wh-element is still available. In this sense, morphological properties are not deterministic as morphosyntactic features may deviate from the original, predictable patterns.

Abbreviations

1	first person	N	neutral
2	second person	NOM	nominative
3	third person	PART	particle
ACC	accusative	PERF	perfective
AOR	aorist	PL	plural
AUX	auxiliary	PRS	present tense
CL	clitic	PTCP	participle
DAT	dative	REFL	reflexive
DIM	diminutive	REL	relative
F	feminine	SG	singular
IMPERF	imperfective	SM	subject marker
M	masculine		

Acknowledgements

References

Aboh, Enoch Oladé. 2005. Deriving relative and factive clauses. In Laura Brugè et al. (eds.), *Contributions to the Thirtieth "Incontro di Grammatica Generativa"*, 265–285. Venice: Cafoscarina.

Allen, Cynthia. 1980. *Whether* in Old English. *Linguistic Inquiry* 11(4). 789–793.

Arsenijević, Boban. 2009. Clausal complementation as relativizationl. *Lingua* 119(1). 39–50. DOI: 10.1016/j.lingua.2008.08.003.

Arsenijević, Boban. 2020. Referential properties of subordinate clauses in Serbo-Croatian. In Teodora Radeva-Bork & Peter Kosta (eds.), *Current developments in Slavic linguistics: Twenty years after (based on selected papers from FDSL 11)*, 341–353. Frankfurt am Main: Peter Lang.

Auderset, Sandra. 2020. Interrogatives as relativization markers in Indo-European. *Diachronica* 37(4). 474–513. DOI: 10.1075/dia.19030.aud.

Julia Bacskai-Atkari

Axel-Tober, Katrin. 2017. The development of the declarative complementizer in German. *Language* 93(2). 29–65. DOI: 10.18148/hs/2017.v0i0.19.

Bacskai-Atkari, Julia. 2016. Towards a cross-linguistic typology of marking polarity in embedded degree clauses. *Acta Linguistica Hungarica* 63(4). 389–409. DOI: 10.1556/064.2016.63.4.1.

Bacskai-Atkari, Julia. 2018. Doubly filled COMP in Czech and Slovenian interrogatives. In Roland Lenertová Denisa Meyer, Radek Šimík & Luka Szucsich (eds.), *Advances in formal Slavic linguistics 2016*, 1–23. Berlin: Language Science Press. DOI: 10.5281/zenodo.2545509.

Bacskai-Atkari, Julia. 2020. German V2 and Doubly Filled COMP in West Germanic. *The Journal of Comparative Germanic Linguistics* 23(2). 125–160. DOI: 10.1007/s10828-020-09117-x.

Bacskai-Atkari, Julia. 2022. Discourse-driven asymmetries between embedded interrogatives and relative clauses in West Germanic. In Nicholas Catasso, Marco Coniglio & Chiara De Bastiani (eds.), *Language change at the interfaces: Intrasentential and intersentential phenomena*, 189–214. Amsterdam: John Benjamins. DOI: 10.1075/la.275.06bac.

Bacskai-Atkari, Julia & Lisa Baudisch. 2018. *Clause typing in Germanic: A questionnaire and its results*. Potsdam: University of Potsdam. URL: https://publishup.uni-potsdam.de/frontdoor/index/index/docId/40681, last accessed: 07.04.2018.

Baltin, Mark. 2010. The nonreality of doubly filled Comps. *Linguistic Inquiry* 41(2). 331–335. DOI: 10.1162/ling.2010.41.2.331.

Bayer, Josef. 1984. COMP in Bavarian syntax. *The Linguistic Review* 3(3). 1–38. DOI: 10.1515/tlir.1984.3.3.209.

Boef, Eefje. 2008. Variatie in Nederlandse lange-afstandsrelatieven. *Taal en Tongval* 60. 90–120.

Boef, Eefje. 2013. *Doubling in relative clauses: Aspects of morphosyntactic microvariation in Dutch*. Utrecht: LOT.

Bošković, željko. 2002. On multiple *wh*-fronting. *Linguistic Inquiry* 33(3). 351–383. DOI: 10.1162/002438902760168536.

Brandner, Ellen & Iris Bräuning. 2013. The particle *wo* in Alemannic: Only a complementizer? *Linguistische Berichte* 234. 131–169.

Browne, Wayles. 1986. *Relative clauses in Serbo-Croatian*. Zagreb: Institute of Linguistics.

Bužarovska, Eleni. 2009. Restrictive relative clauses in standard Macedonian and standard Bulgarian. In Mila Dimitrova-Vulchanova & Olga Mišeska Tomić (eds.), *Investigations in the Bulgarian and Macedonian nominal expression*, 230–260. Trondheim: Tapir Academic Press.

Cazinkić, Robert. 2001. Kategorizacija in razvrstitev oziralnikov *ki* in *kateri*. *Slavistična revija* 49. 55–73.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: MIT Press.

Danylenko, Andrii. 2018. A tale of two pathways: On the development of relative clause chaining in East Slavonic. In Björn Hansen, Jasmina Grković-Major & Barbara Sonnenhauser (eds.), *Diachronic Slavonic syntax: The interplay between internal development, language contact and metalinguistic factors*, 361–386. Berlin: De Gruyter. DOI: 10.1515/9783110531435-013.

Fleischer, Jürg. 2017. *Relativsatz-Einleitung*. In *SyHD-atlas*. Available at http://www.syhd.info/apps/atlas/#relativsatz-einleitung (last accessed: 14 July 2017).

Goodluck, Helen & Danijela Stojanović. 1996. The structure and acquisition of relative clauses in Serbo-Croatian. *Language Acquisition* 5(4). 285–315.

Gračanin-Yuksek, Martina. 2013. The syntax of relative clauses in Croatian. *The Linguistic Review* 30(1). 25–49. DOI: 10.1515/tlr-2013-0002.

Guz, Wojciech. 2017. Definiteness and specificity in two types of Polish relative clauses. In Joseph Emonds & Markéta Janebová (eds.), *Language use and linguistic structure: Proceedings of the Olomouc Linguistics Colloquium 2016*, 323–342. Olomouc: Palacký University.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Halle, Morris & Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics* 21. 275–288.

Halpern, Aaron. 1995. *On the placement and morphology of clitics.* Stanford, CA: CSLI Publications.

Harley, Heidi & Rolf Noyer. 1999. Distributed morphology. *GLOT International* 4(4). 3–9.

Hauge, Kjetil Rå. 1999. *A short grammar of contemporary Bulgarian.* Bloomington: Slavica.

Heine, Bernd & Tania Kuteva. 2002. *World lexicon of grammaticalization.* Cambridge: Cambridge University Press. DOI: 10.1017/9781316479704.

Hladnik, Marko. 2010. *Restrictive relative clauses in Slovene and its spoken varieties.* Utrecht: Utrecht Institute of Linguistics OTS. (MA thesis).

Hopper, Mike & Elizabeth Traugott. 1993. *Grammaticalization.* Cambridge: Cambridge University Press.

Julia Bacskai-Atkari

Izvorski, Roumyana. 2000. A solution to the subcomparative paradox. In Roger Billerey-Mosier & Brook Danielle Lillehaugen (eds.), *WCCFL 19: The proceedings of the 19th West Coast Conference on Formal Linguistics*, 232–245. Somerville: Cascadilla Press.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.

Kljajevic, Vanja. 2012. *Comprehension of Wh-dependencies in Broca's aphasia.* Newcastle upon Tyne: Cambridge Scholars Publishing.

Kramer, Christina E. 1999. *Macedonian.* Madison: University of Wisconsin Press.

Krapova, Iliyana. 2010. Bulgarian relative and factive clauses with an invariant complementizer. *Lingua* 120. 1240–1272. DOI: 10.1016/j.lingua.2009.08.002.

Lazarova-Nikovska, Ana. 2013. On interrogative sentences in Macedonian: A generative perspective. *RCEAL Working Paper in English and Applied Linguistics* 9. 129–159.

McGinnis-Archibald, Martha. 2016. Distributed morphology. In Andrew Hippisley & Gregory Stump (eds.), *The Cambridge handbook of morphology*, 390–423. Cambridge: Cambridge University Press. DOI: 10.1017/9781139814720.015.

Meyer, Roland. 2017. The C system of relatives and complement clauses in the history of Slavic languages. *Language* 93(2). 97–113. DOI: 10.18148/hs/2017.v0i0.21.

Minlos, Philip. 2012. Slavic relative *čto / co*: Between pronouns and conjunctions. *Slověne* 1. 74–91. DOI: 10.31168/2305-6754.2012.1.1.5.

Mitrović, Moreno. 2016. A relative syntax and semantics for Slovenian. In Franc Marušič & Rok Žaucer (eds.), *Formal studies in Slovenian syntax*, 221–252. Amsterdam: John Benjamins. DOI: 10.1075/la.236.10mit.

Ringe, Donald & Ann Taylor. 2014. *The development of Old English.* Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199207848.001.0001.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer. DOI: 10.1007/978-94-011-5420-8_7.

Rudin, Catherine. 2009. The Bulgarian relative marker *-to*. In Steven Franks, Vrinda Chidambaram & Brian Joseph (eds.), *A linguist's linguist: Studies in South Slavic linguistics in honor of E. Wayles Browne*, 403–422. Bloomington: Slavica.

Rudin, Catherine. 2014. Sorting out *-to* and *što*: Bulgarian and Macedonian relative markers. In Jacek Witkoś & Sylwester Jaworski (eds.), *New insights into Slavic linguistics*, 315–326. Frankfurt am Main: Peter Lang.

Salzmann, Martin. 2006. Resumptive pronouns and matching effects in Zürich German relative clauses as distributed deletion. *Leiden papers in Linguistics* 3(1). 17–50.

Salzmann, Martin. 2017. *Reconstruction and resumption in indirect A'-dependencies: On the syntax of prolepsis and relativization in (Swiss) German and beyond.* Berlin: Mouton de Gruyter. DOI: 10.1515/9781614512202.

Schallert, Oliver, Alexander Dröge & Jeffrey Pheiff. 2018. Doubly-filled COMPs in Dutch and German: A bottom-up approach. Manuscript. Universität München / Universität Marburg. Available at https://ling.auf.net/lingbuzz/003979 (last accessed: 17 June 2022).

Šimík, Radek. 2008. Specificity in (Czech) relative clauses. In Jacek Witkoś & Gisbert Fanselow (eds.), *Elements of Slavic and Germanic grammars: A comparative view*, 177–198. Frankfurt am Main: Peter Lang.

Tomić, Olga Mišeska. 2006. *Balkan Sprachbund morpho-syntactic features.* Dordrecht: Springer. DOI: 10.1007/1-4020-4488-7.

Tomić, Olga Mišeska. 2012. *A grammar of Macedonian.* Bloomington: Slavica.

van Gelderen, Elly. 2009. Renewal in the left periphery: Economy and the complementiser layer. *Transactions of the Philological Society* 107(2). 131–195. DOI: 10.1111/j.1467-968X.2009.01216.x.

van Gelderen, Elly. 2013. *Clause structure.* Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139084628.

Watanabe, Akira. 2009. A parametric shift in the D-system in Early Middle English: Relativization, articles, adjectival inflection, and indeterminates. In Paola Crisma & Giuseppe Longobardi (eds.), *Historical syntax and linguistic theory*, 358–374. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199560547.003.0021.

Weiß, Helmut. 2013. Satztyp und Dialekt. In Jörg Meibauer et al. (eds.), *Satztypen des Deutschen*, 764–785. Berlin: Walter de Gruyter. DOI: 10.1515/9783110224832.764.

Zeijlstra, Hedde. 2014. On the uninterpretability of interpretable features. In Peter Kosta et al. (eds.), *Minimalism and beyond*, 109–129. Amsterdam: John Benjamins. DOI: 10.1075/lfab.11.05zei.

Chapter 3

Aspect separated from aspectual markers in Russian and Czech

Petr Biskup

Universität Leipzig

This article is concerned with the derivation of morphological aspect in Russian and Czech. It investigates four aspectual markers: prefixes, the secondary imperfective suffix, the semelfactive marker and the habitual suffix. It argues that not only in Russian (see Tatevosov 2011, 2015) but also in Czech aspect interpretation is separated from prefixes and the secondary imperfective suffix. Moreover, it extends the separation to the semelfactive suffix and the habitual marker. Specific morphological aspect properties of Russian and Czech predicates are derived by an Agree analysis with minimality based on dominance relations in the complex verbal head.

Keywords: Agree, aspect, prefixes, habitual suffix, secondary imperfective, semelfactive suffix

1 Introduction: Aspectual markers

This section introduces four aspectual markers, prefixes, the secondary imperfective marker, the semelfactive suffix and the habitual suffix. I call these morphemes aspectual markers since they are relevant to morphological aspect (they can change the perfective/imperfective value of the base predicate) and/or since they are relevant to aspect more generally, e.g. because of bringing about (a)telicity, habituality or new aktionsart properties.

1.1 Prefixes

Lexical prefixes (also called internal, qualifying, resultative) as well as superlexical (external, modifying, aktionsart) prefixes almost always perfectivize the

imperfective simplex verb (for discussion of the two types of prefixes, see e.g. Isačenko 1962, Petr 1986, Lehmann 1993, Schoorlemmer 1995, Babko-Malaya 1999, Svenonius 2004, Arsenijević 2006, Romanova 2006, Gehrke 2008, Tatevosov 2013, Szucsich 2014, Biskup & Zybatow 2015, Caha & Ziková 2016, Biskup 2019, Klimek-Jankowska & Błaszczak 2021, 2022). For the perfectivizing effect of lexical prefixes, see examples (1) and (2). [1]

(1) a. kleiť[IPF]
         stick
         'to stick on'
     b. na-kleiť[PF]
         on-stick
         'to stick on'   (Russian)

(2) a. chovat[IPF]
         raise
         'to raise'
     b. vy-chovat[PF]
         out-raise
         'to raise'   (Czech)

With respect to the perfectivizing effect of superlexical prefixes, consider examples (3) and (4).

(3) a. delať[IPF]
         do
         'to do'
     b. na-delať[PF]
         CUM-do
         'to do a lot'   (Russian)

(4) a. plést[IPF]
         knit
         'to knit'
     b. do-plést[PF]
         COMP-knit
         'to complete knitting'   (Czech)

---

[1]Lexical prefixes are glossed with a meaning of the corresponding preposition and superlexical prefixes are glossed with the appropriate aktionsart abbreviation.

Both Russian and Czech also have simplex verbs that are perfective. If they combine with a lexical or a superlexical prefix, they remain perfective, as demonstrated by the Russian examples in (5) and the Czech examples in (6).

(5) a. [vy-[kupiť]^{PF}]^{PF}
         out-buy
         'to buy sb.'s freedom'
     b. [na-[kupiť]^{PF}]^{PF}
         CUM-buy
         'to buy a lot'   (Russian)

(6) a. [do-[dat]^{PF}]^{PF}
         to-give
         'to deliver'
     b. [do-[říci]^{PF}]^{PF}
         COMP-say
         'to say to the end'   (Czech)

Lexical and superlexical prefixes can co-occur, as shown by the following examples. Also in this case, the predicate remains perfective. In addition, it holds that the superlexical prefix must occur outside the lexical prefix, as demonstrated by the contrast between examples (7a), (8a) and examples (7b) and (8b).

(7) a. [pere-[vy-polniť]^{PF}]^{PF}
         EXC-out-fulfill
         'to overfulfill'
     b. * [vy-[pere-polniť]^{PF}]^{PF} (Russian)
         out-EXC-fulfill

(8) a. [pře-[vy-chovat]^{PF}]^{PF}
         REP-out-raise
         're-educate'
     b. * [vy-[pře-chovat]^{PF}]^{PF} (Czech)
         out-REP-raise

1.2 The secondary imperfective marker

In this section, I consider the effect of the secondary imperfective suffix on the morphological aspect of the base predicate. Let us begin with Russian.

The secondary imperfective suffix derives an imperfective predicate from a perfective predicate, which can contain a lexical prefix, as in examples (9) and (10).

Petr Biskup

(9)   a. [za-[rabot-a]$^{\text{IPF}}$]$^{\text{PF}}$-ť
        behind-work-TH-INF
        'to earn'

      b. [[za-[rabat]$^{\text{IPF}}$]$^{\text{PF}}$-yva]$^{\text{IPF}}$-ť
        behind-work-SI-INF
        'to earn'   (Russian)

(10)   a. [po-[moč']$^{\text{IPF}}$]$^{\text{PF}}$
        along-can
        'to help'

      b. [[po-[mag]$^{\text{IPF}}$]$^{\text{PF}}$-a]$^{\text{IPF}}$-ť
        along-can-SI-INF
        'to help'   (Russian)

The imperfectivizing suffix can also derive an imperfective predicate from a perfective stem with a superlexical prefix, as in (11), or from a perfective stem without a prefix, as shown in (12).

(11)   a. [za-[rabot-a]$^{\text{IPF}}$]$^{\text{PF}}$-ť
        INC-work-TH-INF
        'to start working'

      b. [[za-[rabat]$^{\text{IPF}}$]$^{\text{PF}}$-yva]$^{\text{IPF}}$-ť
        INC-work-SI-INF
        'to start working'   (Russian)

(12)   a. [d-a]$^{\text{PF}}$-ť
        give-TH-INF
        'to give'

      b. [d-a]$^{\text{PF}}$-va]$^{\text{IPF}}$-ť
        give-TH-SI-INF
        'to give'   (Russian)

Certain superlexical prefixes can also attach outside the imperfectivizing suffix (see e.g. Ramchand 2004, Gehrke 2008, Tatevosov 2013, Szucsich 2014, Klimek-Jankowska & Błaszczak 2021, 2022) and they perfectivize the predicate again, as illustrated in example (13).

(13)   a. [[vy-[talk]$^{\text{IPF}}$]$^{\text{PF}}$-iva]$^{\text{IPF}}$-ť
        out-push-SI-INF
        'to push out'

3 *Aspect separated from aspectual markers in Russian and Czech*

    b. [po-[[vy-[talk]^{IPF}]^{PF}-iva]^{IPF}]^{PF}-t́
        DIST-out-push-SI-INF
        'to push out one after another'   (Russian)

Some superlexical prefixes can occur both inside the imperfectivizing suffix, as the inceptive *za-* in (11), and outside the secondary imperfective marker, as the inceptive *za-* in the following example.

(14) a. [[ot-[kry]^{IPF}]^{PF}-va]^{IPF}-t́
        away-cover-SI-INF
        'to open'
      b. [za-[[ot-[kry]^{IPF}]^{PF}-va]^{IPF}]^{PF}-t́
        INC-away-cover-SI-INF
        'to start opening'   (Russian)

Standardly, the secondary imperfective suffix is taken to have three forms: *-yva-*/*-iva-*, as in (9b), (11b) and (13), *-va-*, as in (12b) and (14), and *-a-*/*-ja-*, as in (10b); see e.g. Vinogradov et al. (1952), but there are also alternative analyses like Isačenko (1962) and Matushansky (2009). A closer look at the data under discussion reveals that *v* is present in *-va-* because of blocking hiatus; compare examples (12) and (14) with example (10b).

    In Czech, an analogous pattern is observed: the secondary imperfective suffix derives an imperfective verb from a perfective stem and the base predicate can contain either a lexical prefix or a superlexical prefix. Examples (15b) and (16b) show an imperfective predicate derived from a lexically prefixed verb.

(15) a. [za-[bí]^{IPF}]^{PF}-t
        behind-beat-INF
        'to kill'
      b. [[za-[bí]^{IPF}]^{PF}-je]^{IPF}-t
        behind-beat-SI-INF
        'to kill'   (Czech)
(16) a. [vy-[pros-i]^{IPF}]^{PF}-t
        out-beg-TH-INF
        'to beg'
      b. [[vy-[proš]^{IPF}]^{PF}-ova]^{IPF}-t
        out-beg-SI-INF
        'to beg'   (Czech)

Petr Biskup

In contrast, example (17b) demonstrates an imperfective predicate derived from a superlexically prefixed predicate.

(17) a. [do-[plés]$^{\text{IPF}}$]$^{\text{PF}}$-t
 comp-knit-inf
 'to complete knitting'

 b. [[do-[plét]$^{\text{IPF}}$]$^{\text{PF}}$-a]$^{\text{IPF}}$-t
 comp-knit-si-inf
 'to complete knitting' (Czech)

The imperfectivizing suffix can also derive an imperfective predicate from an unprefixed perfective verb, as illustrated in examples (18) and (19).

(18) a. [d-á]$^{\text{PF}}$-t
 give-th-inf
 'to give'

 b. [d-á]$^{\text{PF}}$-va]$^{\text{IPF}}$-t
 give-th-si-inf
 'to give' (Czech)

(19) a. [vrát-i]$^{\text{PF}}$-t
 return-th-inf
 'to return'

 b. [vrac]$^{\text{PF}}$-e]$^{\text{IPF}}$-t
 return-si-inf
 'to return' (Czech)

In Czech, too, certain superlexical prefixes attach to the stem after the imperfectivizing suffix. Hence, they perfectivize the secondary imperfective predicate, as illustrated in the following example, based on example (15).

(20) a. [[za-[bí]$^{\text{IPF}}$]$^{\text{PF}}$-je]$^{\text{IPF}}$-t
 behind-beat-si-inf
 'to kill'

 b. [po-[[za-[bí]$^{\text{IPF}}$]$^{\text{PF}}$-je]$^{\text{IPF}}$]$^{\text{PF}}$-t
 dist-behind-beat-si-inf
 'to kill one after another' (Czech)

68

Some superlexical prefixes can attach to the verb both before the imperfectivizing suffix, as in (17), and after the imperfectivizing marker, as in (21c). Both examples contain an occurrence of the completive prefix *do-*. [2]

(21)   a.   [vy-[plés]$^{\text{IPF}}$]$^{\text{PF}}$-t
        out-string-INF
        'to string'
      b.   [[vy-[plét]$^{\text{IPF}}$]$^{\text{PF}}$-a]$^{\text{IPF}}$-t
        out-string-SI-INF
        'to string'
      c.   [do-[[vy-[plét]$^{\text{IPF}}$]$^{\text{PF}}$-a]$^{\text{IPF}}$]$^{\text{PF}}$-t
        COMP-out-string-SI-INF
        'to complete stringing'   (Czech)

It is obvious from the examples that there are three secondary imperfective markers in Czech: -(*v*)*a*-, present in (17), (18) and (21), -*ova*-, occurring in (16), and the suffix -(*j*)*e*-, which is present in (15) and (19) and which is not productive (see Petr 1986). The examples also suggest that *v* in -*va*- and *j* in -*je*- block hiatus; compare (18) with (17b) and (15b) with (19b). In fact, the pattern could be simplified if we decomposed -*ova*- and the Russian -*yva*-/-*iva*-. They follow the general Slavic -V*va*- pattern, with a vowel, -*v*- blocking hiatus and (the iterative) -*a*- (see e.g. Kuznecov 1953 and Lunt 2001). For ease of exposition, I will treat the imperfectivizing markers as a whole in what follows.

Thus, the relevant part of the linearized structure with aspectual markers and their aspectual effects looks like (22). *LP* stands for lexical prefixes, *SP* for superlexical prefixes and *SI* for the secondary imperfective suffix.

(22)   [SP$_{\text{higher}}$[[SP$_{\text{lower}}$[LP[√root]$^{\text{PF/IPF}}$]$^{\text{PF}}$]$^{\text{PF}}$SI]$^{\text{IPF}}$]$^{\text{PF}}$

Recall that some superlexical prefixes merge lower and others higher than the imperfectivizing suffix (and some of them can merge in a lower as well as in a higher position).

1.3 The semelfactive marker

The semelfactive suffix consists of -*n*- and some vowel in Slavic (the original form was *-*nVn*-; see Wiemer & Seržant 2017). It selects a root with a punctual

---

[2]In this respect, Czech differs from Russian, which only allows completive *do-* in the lower position (see Tatevosov's 2008 discussion of intermediate prefixes).

Petr Biskup

or instantaneous property and derives a perfective stem, as illustrated in the Russian example (23) and the Czech example (24).[3]

(23)   a. krik
           shout
           'shout'
       b. krik-nu-t$^{\text{PF}}$
           shout-SEML-INF
           'to shout out'   (Russian)

(24)   a. bod
           point
           'point'
       b. bod-nou-t$^{\text{PF}}$
           point-SEML-INF
           'to stab'   (Czech)

The semelfactive marker differs from the suffix *-nV-* present in other verbs like degree achievements. The degree achievement *-nV-* selects a root denoting a property and does not have a perfectivizing effect on the verb (see Taraldsen Medová & Wiland 2019 for the relation and differences between the two *-nV-* suffixes).

Since the semelfactive suffix attaches directly to the root and verbalizes it, as shown by the contrasts in (23) and (24), I assume that it spells out the verbalizing head *v*. If correct, then we expect the semelfactive suffix to be in complementary distribution with other themes representing the verbalizing *v*. This prediction is borne out, as demonstrated below. The examples in (25a) and (26a) show a grammatical combination of the root and a theme vowel, whereas the examples in (25b)–(25c) and (26b)–(26c) – based on grammatical forms (23b) and (24b) – demonstrate that the co-occurrence of the theme vowel and the semelfactive suffix leads to ungrammaticality in both orders.[4,5]

---

[3]Some Russian verbs take the expressive, extended marker *-anu-* (and some both *-nu-* and *-anu-*); see e.g. Isačenko (1962) and Švedova (1980).

[4]A reviewer suggests analyzing the marker *-nu-* as a sequence of the semelfactive marker (with the perfective feature) and the theme vowel, which would have the advantage that all theme vowels would be analyzed identically: as verbalizers without aspectual features. The disadvantage, however, is that then the verbalizer (the theme vowel) would not be adjacent to the root, contrary to the standard assumption. In addition, the elements behave like a unit, e.g. with respect to elision; cf. the following Czech alternatives in the past tense: *tiskl/tisknul* 'printed'.

[5]To avoid hiatus, I insert /v/ between the semelfactive suffix and the theme vowel in (25b) and (26c), a strategy known from secondary imperfectives.

(25) a. krič-a-ť
         shout-TH-INF
         'to shout'

     b. * krik-nu-va-ť
         shout-SEML-TH-INF
         Intended: 'to shout out'

     c. * krič-a-nu-ť
         shout-TH-SEML-INF
         Intended: 'to shout out'   (Russian)

(26) a. bod-a-t
         point-TH-INF
         'to stab'

     b. * bod-a-nou-t
         point-TH-SEML-INF
         Intended: 'to stab'

     c. * bod-nou-va-t
         point-SEML-TH-INF
         Intended: 'to stab'   (Czech)

Given that the semelfactive marker represents the verbalizing head *v*, the complementary distribution of this suffix and the secondary imperfective marker – shown in (27) and (28) – cannot be based on structural blocking, as proposed e.g. by Markman (2008) for Russian.

(27) * krik-nu-va-ť
       shout-SEML-SI-INF
       Intended: 'to shout out'   (Russian)

(28) * bod-nou-va-t
       point-SEML-SI-INF
       Intended: 'to stab'   (Czech)

The reason for ungrammaticality of cases like (27) and (28) can be rather semantic. For instance, Jabłońska (2007) argues that semelfactives – being instantaneous – do not have a process part in their event structure, on which the progressive operator of secondary imperfectives could operate. Another possibility is to assume that the secondary imperfective suffix spells out an atelicizer/eventizer, which combines with complex events, i.e. accomplishments ($\lambda R.\lambda e.\exists s[R(e)(s)]$, see Łazorczyk 2010 and Tatevosov 2015). It is obvious that semelfactives are not of the

appropriate eventive type; they do not introduce a change of state (e.g. Smith 1991) and they are taken to be achievements by Vendler (1957).[6]

There is also a possibility to exclude cases like (27) and (28) by morphological blocking, where the existence of the simpler imperfective forms *kričať* in (25a) and *bodat* in (26a) prevents the use of the more complex forms (27) and (28). The advantage of the second and the third possibility is that in contrast to the argument by Jabłońska (2007) they can also answer the question of why (27) and (28) are not possible with the iterative (non-progressive) reading of the imperfectivizing suffix.[7]

As to structural properties of the semelfactive -$n(V)$-, it needs to be placed outside lexical prefixes, as demonstrated in (29), with SEML representing the verbalizing head *v*.

(29) $[SP_{higher}[[SP_{lower}[_v \text{ SEML } [LP[\sqrt{root}]^{PF/IPF}]^{PF}]^{PF}]^{PF}SI]^{IPF}]^{PF}$

The rationale behind is that root nominalizations can contain lexical prefixes but cannot include the semelfactive -$n(V)$-. As shown in (30) for Russian and in (31) for Czech, root nominalizations can contain lexical prefixes but can include neither lower superlexicals nor higher superlexical prefixes (see also Caha & Ziková 2016 for Czech data). The Russian *podkop* can only have the meaning 'tunnel'; the attenuative superlexical interpretation of *pod-* is not available in this case. Similarly in the Czech (31), *příkop* can only mean 'ditch' and the prefix *pří-* cannot have the attenuative interpretation.

(30)   a.   pod-kop
           under-dig
           'tunnel'

---

[6]The second reasoning could also explain the incompatibility of the degree achievement -$n(V)$- with the imperfectivizing suffix in cases like (i.b). Alternatively, one may suggest that the ungrammatical status of (i.b) has an economy reason because degree achievement verbs like *sochnuť* in (i.a) are imperfective (without the imperfectivizing suffix).

(i)   a.   soch-nu-ť
           dry-DA-INF
           'to dry'

     b.   * soch-nu-va-ť
           dry-DA-SI-INF
           Intended: 'to dry'   (Russian)

[7]As pointed out by a reviewer, the claim that the complementary distribution of the semelfactive suffix and the secondary imperfective marker is not based on structural blocking is also supported by the fact that in languages like South-East Serbo-Croatian, the two markers are combined quite productively, as in *tak-n-uje-m* 'I touch repeatedly'.

b. * pod-kop
ATT-dig
Intended: 'little kick' (Russian)

(31) a. pří-kop
at-dig
'ditch'

b. * pří-kop
ATT-dig
Intended: 'little kick' (Czech)

This means that the boundary of root nominalizations must be placed between the projection containing lexical prefixes and the projections with lower super-lexicals (and the projection with the semelfactive suffix) in (29).

There is, however, an interesting distinction between Russian and Czech with respect to nominalizations and the semelfactive suffix. While in Czech the suffix can be a part of stem nominalizations, in Russian it is not possible; consider the contrast between (32) and (33).

(32) * kop-nu-t-i-e
dig-SEML-N/T-NMLZ-NOM.SG
Intended: 'a dig/kick' (Russian)

(33) kop-nu-t-í
dig-SEML-N/T-NMLZ.NOM.SG
'a dig/kick' (Czech)

This can be related to the fact that in contrast to Czech nominalizations, Russian stem nominalizations are structurally less complex and do not contain the aspectual projection, as discussed in the next section.

As illustrated in (23) and (24), the semelfactive suffix perfectivizes the stem, as do prefixes. If both elements co-occur, then unsurprisingly the predicate remains perfective, irrespective of whether the prefix is lexical or superlexical. For a lexical prefix, consider the Russian example in (34) and for a superlexical prefix consider the Czech example (35), with an attenuative reading.

(34) [vs-[krik-nu-ť]^{PF}]^{PF}
up-shout-SEML-INF
'to give a scream' (Russian)

Petr Biskup

(35)   [na-[prask-nou-t]^{PF}]^{PF}
       ATT-crack-SEML-INF
       'to crack partially'   (Czech)

Generally, it is difficult to find examples of semelfactive predicates with a superlexical prefix. This results from the fact that semelfactive predicates refer to bounded singleton events that are punctual, which clashes with the fact that superlexical prefixes typically modify the spatiotemporal path of the event expressed by the base predicate. Moreover, the perfective aspect of semelfactive verbs pose a problem for the imperfective selection properties of some superlexical prefixes.

As the comparison of (36a) and (36b) shows, the semelfactive *-n(V)-* is responsible for the ungrammatical status of the verb prefixed by the delimitative prefix *po-*.

(36)   a.   po-bod-a-t
            DEL-point-TH-INF
            'to stab to a certain extent several times'

       b.   *po-bod-nou-t
            DEL-point-SEML-INF
            Intended: 'to stab in a short time frame'   (Czech)

Building on the data, I propose the following meaning for the semelfactive *-n(V)-*.

(37)   ⟦SEML⟧ = $\lambda P \lambda e[P(e) \wedge \text{ATOM}(e) \wedge |e| = 1]$

It derives predicates with a single occurrence of the event described by the stem and the event is atomic. That is, there is no proper part of the event (it is punctual), which means that the predicate is not divisive, which in turn means that it is quantized (see Borer 2005). Because of the minimal (atomic) property of the semelfactive *-n(V)-*, there is no path in the event that could be accessible to the delimitative *po-* in cases like (36b).[8] The ungrammaticality of (36b) cannot be based on unsatisfied selection properties of the prefix *po-* if delimitative *po-* and attenuative *po-* form a natural class. Specifically, the attenuative prefix can also adjoin to perfective predicates in Czech, as in [*po-*[*otevřít*]^{PF}]^{PF} 'to open a little'.

The single occurrence property of the semelfactive *-n(V)-* in (37) is responsible for the fact that the iterative reading is not available in cases like *kriknut'* 'to shout out' and *bodnout* 'to stab' in (23b) and (24b), respectively. In contrast, predicates

---

[8]The minimal property is a (language) idealization; in the real world, there can be some trajectory involved e.g. in the stab movement (cf. Rothstein 2004).

with the identical root but without the semelfactive *-n(V)-* like *kričať* 'to shout' and *bodat* 'to stab' in (25a) and (26a) allow the iterative interpretation.[9]

1.4 The habitual marker

Russian habitual forms like (38b) – derived from (38a) – are classified as colloquial or archaic and it is often claimed that they only occur in the past tense (see Isačenko 1962, Zaliznjak & Šmelev 1997, Paducheva 2015, but see also Tatevosov 2013).[10]

(38) a. pis-a-ť[IPF]
          write-TH-INF
          'to write'
      b. pis-yva-ť[IPF]
          write-HAB-INF
          'to write repeatedly'   (Russian)

In contrast, Czech derives analogous imperfective forms quite productively (Filip 1993, Filip & Carlson 1997, Esvan 2007, Nübler 2017, but see also Berger 2009); consider example (39). Certain authors even consider forms like (39b) to be an instantiation of a 'third aspect' (see e.g. Kopečný 1962).[11]

(39) a. ps-á-t[IPF]
          write-TH-INF
          'to write'
      b. ps-á-va-t[IPF]
          write-TH-HAB-INF
          'to write repeatedly'   (Czech)

The examples above show that in both languages, the habitual suffix derives an imperfective verb from an imperfective base.

In Czech, there are also reduplicative forms like (40), which are usually described as expressive predicates denoting a longer (or temporally distant, see Filip 1993) habitual event. They are imperfective, too.

---

[9]The single occurrence property can be defined in terms of a maximality operator; see Egg (2018).

[10]I use the term HABITUAL but various terms can be found in the literature: "iterative", "frequentative" and "generic".

[11]Against expectations, Polish is even more restricted than Russian with respect to habitual forms like *pis-ywa-ć* 'to write repeatedly'. There are only a few verbs (see Grzegorczykowa et al. 1984 and Łaziński 2020).

Petr Biskup

(40) ps-á-vá-va-t[IPF]
      write-TH-HAB-HAB-INF
      'to write repeatedly for a long time/long ago'   (Czech)

In contrast to Russian, it is also possible to derive a habitual predicate from a secondary imperfective verb in Czech, as shown by the pair in (41). The derived verb is again imperfective.

(41) a. vy-pis-ova-t[IPF]
          out-write-SI-INF
          'to excerpt'
      b. vy-pis-ová-va-t[IPF]
          out-write-SI-HAB-INF
          'to excerpt repeatedly'   (Czech)

Examples (39b) and (41b) show that the habitual marker is outside the theme and the imperfectivizing suffix, respectively. Building on the structural proposal in (29), that means that the habitual suffix must also be higher than lexical prefixes and lower superlexical prefixes.

   In fact, the habitual marker is even higher than higher superlexical prefixes and the aspectual projection. The argument goes as follows. It has been argued that Russian *nie*-nominals are aspectless (see Švedova 1980, Schoorlemmer 1995, Gehrke 2008, Tatevosov 2011, 2020); hence phasal verbs can combine with prefixed nominals derived from a perfective stem like in (42).

(42) načal   na-pis-a-n-i-e
      started on-write-TH-N/T-NMLZ-ACC.SG
      'started writing'   (Russian; based on Tatevosov 2011: ex. (18))

On the contrary, Czech stem nominalizations have the morphological aspect (e.g. Procházková 2006). For this reason, the phasal verb is compatible with the imperfective nominals in (43a) and (44a) but is not compatible with the perfective nominals in (43b) and (44b).

(43) a.   začal   vy-pis-ová-n-í
           started out-write-SI-N/T-NMLZ.ACC.SG
           'he started writing out'
      b. * začal   vy-ps-á-n-í
           started out-write-TH-N/T-NMLZ.ACC.SG
           Intended: 'he started writing out'   (Czech)

(44)   a.   začalo   na-kup-ová-n-í
               started on-buy-SI-N/T-NMLZ.NOM.SG
               'buying started'

       b. * začalo   na-koup-e-n-í
               started on-buy-TH-N/T-NMLZ.NOM.SG
               Intended: 'buying started'   (Czech)

Czech stem nominalizations can be prefixed with higher superlexical prefixes like the cumulative *na-* in example (45a), in contrast to Russian *-nie* nominals, which only allow superlexicals in the lower position (see Tatevosov 2011). Note that the prefix *na-* is indeed cumulative because the prefixed predicate can take a plural object like in *naházení židlí na něco* 'throwing chairs on sth.' but cannot combine with a quantized singular object like in *naházení židle na něco* 'throwing a chair on sth.'. Crucially, stem nominalizations cannot contain the habitual suffix, as demonstrated in (45b).

(45)   a.   na-ház-e-n-í
               CUM-throw-TH-N/T-NMLZ.NOM.SG
               'throwing a lot of sth.'

       b. * ps-á-vá-n-í
               write-TH-HAB-N/T-NMLZ.NOM.SG
               Intended: 'repeated writing'   (Czech)

This means that stem nominalizations include the structure in (29). Their structure includes higher superlexical prefixes but also the aspectual projection in Czech, which hosts the perfective or the imperfective operator responsible for the morphological aspect interpretation.[12] At the same time, the data suggest

---

[12]In the case of the perfective operator, the event time is included in the reference time, as in (i.a), and with the imperfective operator, the reference time is included in the event time, as shown in (i.b) (both taken from Paslawska & von Stechow 2003: 322).

   (i)   a.   PERFECTIVE = $\lambda P \lambda t \exists e. \tau(e) \subseteq t \wedge P(e)$

         b.   IMPERFECTIVE = $\lambda P \lambda t \exists e. t \subseteq \tau(e) \wedge P(e)$

For predicates with a result state introduced by a prefix, one can add the state variable and the trace function mapping the state to its time, as in (ii) (taken from Biskup 2019: 43).

   (ii)   PERFECTIVE = $\lambda R \lambda t \exists s \exists e [R(s)(e) \wedge \tau(e) \subseteq t \wedge \tau(e) \supset\subset \tau(s)]$

The presence of the appropriate operator is tested with the standard diagnostics for perfectivity and imperfectivity, i.e. (in)compatibility with the auxiliary 'to be', (im)possibility of the future interpretation of the present form, (in)compatibility with phase verbs and the formation of participles. Note that I follow the two-component approach to aspect and distinguish the morphological (grammatical, outer) aspect from the lexical (situation, inner) aspect.

Petr Biskup

that the habitual suffix is higher than superlexical prefixes and the aspectual projection.

The high position of the habitual affix finds support in the fact that the marker can scope over quantificational adverbs, which are very high in the clausal structure; consider the following example.

(46) Z   dovolené ps-á-va-l   velmi zřídka.
     from vacation write-TH-HAB-PART.M.SG very rarely
     'It was almost always the case that when he was on vacation, he sent a
     letter very rarely.'   (Czech)

I assume for the time being that the meaning of the habitual marker is 'almost always', as shown in the translation in (46). The rationale behind is that the meaning of always is too strong. Given that sentence (47) is anomalous, the meaning of the habitual marker cannot be 'always'. That would derive a fully acceptable sentence.

(47) * Člověk bý-vá-Ø   smrtelný.
       man   be-HAB-3.SG mortal
       'Man is almost always mortal.'   (Czech)

Given the high structural position of the habitual marker, the question arises why it is not compatible with the semelfactive *-n(V)-*, as illustrated in (48) and (49). The answer is not complicated. The habitual suffix selects an imperfective predicate but the semelfactive affix derives perfective verbs.

(48) * krik-nu-va-ť
       shout-SEML-HAB-INF
       Intended: 'to shout out repeatedly'   (Russian)

(49) * bod-nou-va-t
       point-SEML-HAB-INF
       Intended: 'to stab repeatedly'   (Czech)

In both languages, the habitual suffixes are identical to the secondary imperfective suffixes. Russian mostly uses the marker *-yva-/-iva-*, as in (38b), but the markers *-va-* and *-a-/-ja-* can also be found; consider verbs in (50) and (51). These examples again suggest that *-va-* and *-a-* are phonologically conditioned allomorphs.

(50) a. pe-t'$^{\text{IPF}}$
          sing-INF
          'to sing'
      b. pe-va-t'$^{\text{IPF}}$
          sing-HAB-INF
          'to sing repeatedly'   (Russian)

(51) a. vid-e-t'$^{\text{IPF}}$
          see-TH-INF
          'to see'
      b. vid-a-t'$^{\text{IPF}}$
          see-HAB-INF
          'to see repeatedly'   (Russian)

In Czech, habitual suffixes form a subset of the secondary imperfective markers. Beside *-va-*, there is also its allomorph *-a-*, as in (52), and the marker *-e-*, which is not productive (see Petr 1986).

(52) a. jís-t$^{\text{IPF}}$
          eat-INF
          'to eat'
      b. jíd-a-t$^{\text{IPF}}$
          eat-HAB-INF
          'to eat repeatedly'   (Czech)

In what follows, I argue that – albeit homophonous – the habitual markers are not secondary imperfective suffixes. First, there are morphological aspect differences. While the imperfectivizing suffix derives an imperfective predicate from a *perfective* verb, the habitual suffix derives an imperfective predicate from an *imperfective* base.

There are also interpretational differences. Secondary imperfective verbs can have the progressive interpretation, the iterative interpretation, the factual and the habitual/generic interpretation. In contrast, predicates with the habitual suffix can only have the habitual/generic interpretation, as demonstrated by the (*repeatedly*) translations in this section. An analogous distinction is observed in cases with iterative adverbs, as in (53). In sentence (53a), two interpretations are available: The first, cardinality interpretation has three iterated events of writing during one vacation. The second one is the habitual quantificational interpretation, which is probably stronger than the habitual interpretation of predicates

Petr Biskup

with the overt habitual marker. In contrast, with the habitual suffix, as in (53b), only the habitual interpretation is available, with *z dovolené* going to the restrictor and *psával třikrát* to the nucleus of the habitual quantifier ALMOST ALWAYS (or of the standard generic operator).

(53) a. Z   dovolené ps-a-l   třikrát.
          from vacation write-TH-PART.M.SG three.times
          'From vacation, he sent a letter three times.'
          'From vacation, he used to send a letter three times.'

      b. Z   dovolené ps-á-va-l   třikrát.
          from vacation write-TH-HAB-PART.M.SG three.times
          'It was almost always the case that when he was on vacation, he sent
          a letter three times.'   (Czech)

The next argument is based on differences in nominalizations. As already shown by the ungrammatical form *psávání in (45b), the habitual marker cannot be included in stem nominalizations. However, the secondary imperfective suffix can be a part of such nominalizations, as illustrated in (54b) (and simplex verbs can also be nominalized, as shown in (54a)).

(54) a. ps-a-n-í
          write-TH-N/T-NOM.SG
          'writing'

      b. vy-pis-ová-n-í
          out-write-SI-N/T-NOM.SG
          'excerpting''   (Czech)

As to phonological properties of the secondary imperfective suffix and the habitual marker, there are many similarities. Both affixes can induce a vowel change, most typically the change from the phoneme /o/ to /a/, which is a relic of the Proto-Indo-European vowel gradation (lengthening, see e.g. Nandris & Auty 1969). For the Russian imperfectivizing suffix, consider (55) and for the habitual marker, see (56).[13]

(55) a. s-pr*o*s-í-t^{PF}
          with-ask-TH-INF
          'to ask'

---

[13]In the perfective form in (55a), the phoneme /o/ is reduced and surfaces as the phone [ɐ] given its positioning in the first pretonic syllable.

    b. s-pr*áš*-iva-t̕^{IPF}
       with-ask-SI-INF
       'to ask'   (Russian)

(56) a. ch*o*d-í-t̕^{IPF}
       walk-TH-INF
       'to walk'

    b. ch*áž*-iva-t̕^{IPF}
       walk-HAB-INF
       'to walk repeatedly'   (Russian)

The examples also show that both aspectual morphemes can shift the accent to the root and that the underlying front theme vowel can palatalize the root consonant in the derived forms in (55b) and (56b).

    Lengthening processes are observed in Czech, too. In (57) the imperfectivizing marker -(*v*)*a*- lengthens the preceding theme vowel. Similarly, in (58) the habitual marker -(*v*)*a*- lengthens the preceding theme -*i*-. This lengthening also applies in reduplicated form, as already shown in (39b) and (40) by the habitual form *ps-á-va-t* and the reduplicated *ps-á-vá-va-t*, respectively.

(57) a. vy-děl-*a*-t^{PF}
       out-make-TH-INF
       'to earn'

    b. vy-děl-*á*-va-t^{IPF}
       out-make-TH-SI-INF
       'to earn'   (Czech)

(58) a. chod-*i*-t^{IPF}
       walk-TH-INF
       'to walk'

    b. chod-*í*-va-t^{IPF}
       walk-TH-HAB-INF
       'to walk repeatedly'   (Czech)

However, there are differences between phonological effects of the two markers. The habitual marker lengthens the preceding vowel but does not induce transitive palatalization in contrast to the secondary imperfective suffix. Consider the following examples, with the root *pros*, which is palatalized by the theme -*i*- in (59a)–(59b) but is not affected in (59c)–(59d).

Petr Biskup

(59) a. vy-*pros*-i-t$^{\text{PF}}$
out-beg-TH-INF
'to beg'

b. vy-*proš*-ova-t$^{\text{IPF}}$
out-beg-SI-INF
'to beg'

c. *pros*-i-t$^{\text{IPF}}$
beg-TH-INF
'to beg'

d. *pros*-í-va-t$^{\text{IPF}}$
beg-SI-SI-INF
'to beg repeatedly' (Czech)

This different behavior possibly results from a specific templatic properties of secondary imperfective verbs in Czech, which must weigh three morae without the prefix (see Scheer 2003, Caha & Scheer 2008, Caha & Ziková 2016 for templatic properties of Czech verbal forms). In fact, this is what we expect if the imperfectivizing suffix and the habitual marker are two different elements representing distinct pieces of structure that enter into relations with differently complex constituents.

Moreover, the Czech habitual marker does not induce the vowel gradation in the root (with transitive palatalization) in contrast to the imperfectivizing marker. Compare *chod-í-va-t* 'to walk repeatedly' from (58b) with the Russian *cháž-iva-ť* 'to walk repeatedly' in (56b) and with (60), which contains the /o/-/a/ alternation induced by the imperfective suffix.

(60) a. vy-*tvoř*-i-t$^{\text{PF}}$
out-make-TH-INF
'to make'

b. vy-*tvář*-e-t$^{\text{IPF}}$
out-make-SI-INF
'to make' (Czech)

Given the differences just discussed, I conclude that the imperfectivizing suffix and the habitual suffix are not identical elements. Yet, there can be one underspecified vocabulary item that spells out both elements, as shown in (61).

(61) -*yva*- ↔ [ipf]

82

According to this rule, *-yva-* (which represents allomorphs of the habitual and the imperfectivizing suffix) is inserted into a morphosyntactic context specified as imperfective. That is, *-yva-* can realize the habitual and the imperfectivizing head, which both have the imperfective feature (for more discussion, see §3). The syntactic, semantic and phonological differences between the two suffixes then result from the fact that they represent distinct pieces of the morphosyntactic structure and consequently enter into relations with different elements.

To conclude this section, the linearized structure with the four aspectual markers and their morphological aspect effects looks like (62).

(62) $[[[SP_{higher} [[SP_{lower} [_v SEML [LP [\sqrt{root}]^{PF/IPF}]^{PF}]^{PF}]^{PF} SI]^{IPF}]^{PF} Asp]$
      $HAB]^{IPF}$

Note that it is an overall picture that does not take into account selection properties and particular incompatibilities of the markers.

2 Aspect separated from the four aspectual markers

We have seen that the aspectual interpretation is determined by several elements, which can have opposite aspectual effects (perfective versus imperfective). The discussion of the four markers and their morphological aspect effects showed that the morphological aspect value of a predicate can change in the course of its derivation. That is, each new aspect marker adds a new aspect layer to the preceding derivation that covers the preceding aspect values. Recall that we have seen that the morphological aspect is determined by the last attached aspectual morpheme. I will call it *Morphological Aspect Generalization* (MAG); consider (63).

(63) *Morphological Aspect Generalization*
      The morphological aspect is determined by the last attached aspectual morpheme.

I also showed that in certain cases aspectual markers do not change the morphological aspect interpretation. These facts are not new; see e.g. Karcevski (1927), Isačenko (1962), Zinova & Filip (2015) and Tatevosov (2020). Given these facts, we need a mechanism that can *inspect* all the relevant aspectual morphemes and can determine which of them is the final one.

The ideal candidate is the operation Agree. Given that it can establish a relation between the probe and the goal at a distance, it is suitable for cases where the interpretation is separated from the element that triggers it.

Petr Biskup

Tatevosov (2011) argues that prefixes are not morphological exponents of the perfective aspect. His argument is based on the fact that Russian stem nominalizations are aspectless although they are formed from prefixed stems. In other words, if prefixes were not dissociated from the perfective meaning, Russian *-nie* nominals would have to be interpreted as perfective. According to Pazelskaya & Tatevosov (2008) and Tatevosov (2011), Russian stem nominalizations include the projection with the secondary imperfective suffix at the most. As discussed in §1.4, Czech stem nominalizations also contain higher superlexical prefixes and the aspectual projection. Thus, the structures of the two languages differ in the presence/absence of higher superlexicals and the aspectual projection (i.e. the presence/absence of the aspectual interpretation), as shown in my notation in (64) and (65).

(64) $[[[[\text{SP}_{\text{lower}} [_v \text{SEML} [\text{LP} [\sqrt{\text{root}}]^{\text{PF/IPF}}]^{\text{PF}}]^{\text{PF}}]^{\text{PF}} \text{SI}]^{\text{IPF}} \text{N/T}] n]$ Russian

(65) $[[[[\text{SP}_{\text{higher}} [[\text{SP}_{\text{lower}} [_v \text{SEML} [\text{LP} [\sqrt{\text{root}}]^{\text{PF/IPF}}]^{\text{PF}}]^{\text{PF}}]^{\text{PF}} \text{SI}]^{\text{IPF}}]^{\text{PF}} \text{Asp}]$ N/T] n]   Czech

Now I will extend the separation argument to the semelfactive marker. Since Russian nominalizations generally disallow the presence of the semelfactive $-n(V)$- and Czech stem nominalizations (with or without SEML) always have the morphological aspect, we cannot construct a direct argument with aspectless nominals containing the semelfactive $-n(V)$-.[14] Recall that I argued in §1.3 that the semelfactive suffix spells out the verbalizing head v, as do other theme elements; consider (64) and (65) again. Given this and the fact that the aspectual projection occurs outside the projection with the imperfectivizing suffix (and also higher than projections with the *-n-/-t-* suffix and the nominalizing suffix in Russian, as shown in (64)), it is obvious that the semelfactive marker is separated from the perfective aspect. Below I will show that the semelfactive marker is also separated from the aspectual projection by the projection of Voice, which introduces the agent argument.

Note that it would not be reasonable to postulate another aspectual projection with the perfective interpretation specific to the semelfactive $-n(V)$- because of the reason of language economy and because of universality of the clausal hierarchy. Moreover, given that the perfectivity effect of the semelfactive $-n(V)$- is real – see the periphrastic future test in (66) and (67) – the analysis of the semelfactive marker cannot be based only on its inner aspect properties.

---

[14]The question of exactly how the presence of Asp licenses the presence of the semelfactive marker in Czech, I leave for future research.

(66) a.   budet krič-a-ťIPF
          will   shout-TH-INF
          'it/(s)he will shout'

    b. * budet krik-nu-ťPF
          will   shout-SEML-INF
          Intended: 'it/(s)he will shout out'   (Russian)

(67) a.   bude bod-a-tIPF
          will point-TH-INF
          'it/(s)he will stab'

    b. * bude bod-nou-tPF
          will point-SEML-INF
          Intended: 'it/(s)he will stab'   (Czech)

Romanova (2004), Tatevosov (2015) and Mueller-Reichau (2020) argue for Russian that the imperfectivizing suffix merges inside the verbal domain. Thus, the secondary imperfective marker, too, is dissociated from its interpretation because the aspectual head responsible for the imperfective interpretation is located in a higher position above *v*P. According to Biskup (2020) – who uses a scope argument like the one in Tatevosov (2015) – scope facts with the Czech cumulative *na-* also suggest that the position of the imperfectivizing suffix is below the projection with the agentive argument. The same point can be done with the distributive prefix *po-*.

Concretely, cumulative *na-* and distributive *po-* can quantify over an object, as shown by the grammatical plural (non-quantized) object in (68a). The ungrammaticality of the quantized, singular object jablko 'apple' shows that the prefix *na-* is indeed cumulative and the prefix *po-* distributive. In contrast, the prefixes cannot quantify over an agentive subject, as demonstrated in (68b), where the plural subject is ungrammatical. Only if the object is plural, non-quantized, the sentence is grammatical, as demonstrated in (68c). This goes hand in hand with the fact that when we want to quantify over the agentive subject, the argument structure (including case properties) of the verb needs to be manipulated and the reflexive element must be added in the case of the cumulative *na-*, as shown in (68d).[15]

---

[15]Also compare the following examples with 'self' and the cumulative/saturative *na-*, which can quantify over the subject.

(i) a.   na-begat'-sja
          on-run-self
          'to have one's fill of running'   (Russian)

Petr Biskup

(68)   a. po-/na-s-bír-a-t^PF   {jablka / *jablko}
          DIST-/CUM-with-take-SI-INF apples apple
          distributive: 'to pick apples/*apple one after another'
          cumulative: 'to pick amount of apples/*apple'

      b. * Sousedi   po-/na-sbírali   jablko.
          neighbors DIST-/CUM-picked apple
          Intended distributive: 'Neighbors one after another picked an apple.'
          Intended cumulative: 'Amount of neighbors picked an apple.'

      c. Sousedi   po-/na-sbírali   jablka.
          neighbors DIST-/CUM-picked apples
          distributive: 'Neighbors picked apples one after another.'
          cumulative: 'Neighbors picked amount of apples.'

      d. Sousedi   se nasbírali jablek   do sytosti.
          neighbors self picked   apples.GEN.PL to one's.fill
          'Neighbors had their fill of picking apples.'   (Czech)

Given that the perfective *nasbírat* is derived by attaching the cumulative *na-* and the distributive *po-* to the stem after the secondary imperfective suffix, the example suggests that higher superlexical prefixes like the cumulative *na-* and the distributive *po-* merge below the head introducing the agent and above the imperfectivizing suffix in Czech. Consequently, in the light of the fact that the aspectual projection is above the projection introducing the agent (e.g. Babko-Malaya 2003, Filip 2005, Błaszczak & Klimek-Jankowska 2012, Gribanova 2015), it is possible to conclude that the imperfective interpretation is separated from the imperfectivizing suffix.

At the same time, if it is correct that higher superlexical prefixes merge below the projection with the agent (VoiceP), we also have an argument for separating prefixes from the perfective interpretation occurring in the aspectual projection.

The following examples show that stem nominalizations like the Russian *-nie* nominals and the Czech *-ní* nominals can have an agent. The nominals can co-occur with an agent-oriented modifier, as in (69a) and (70a), and can be modified by an agentive *by*-phase, as shown in (69b) and (70b).

(69)   a. umyšlennoe prestuplenie
        deliberate   delict
        'a wilful delict'

---

        b. na-běhat se
           on-run   self
           'to have one's fill of running'   (Czech)

    b. soveršenie   prestuplenija licom…
       perpetration delict.GEN.SG person.INSTR.SG
       'a perpetration of the delict by a person'   (Russian)

(70) a. úmyslné poškození
       deliberate damage
       'a malicious damage'

    b. spáchání   trestného   činu   osobou…
       perpetration criminal.GEN.SG act.GEN.SG person.INSTR.SG
       'a perpetration of the delict by a person'   (Czech)

Now let us combine it with the fact that Russian stem nominalizations are aspectless (as discussed in §1.4). Applying the containment argument again, we conclude that (at least in Russian) the aspectual projection is indeed above VoiceP, as shown in (71).

(71) $[[[\text{SP}_{\text{higher}} [[\text{SP}_{\text{lower}} [_v \text{ SEML } [\text{LP } [\sqrt{\text{root}}]^{\text{PF/IPF}}]^{\text{PF}}]^{\text{PF}}]^{\text{PF}}\text{SI}]^{\text{IPF}}]^{\text{PF}} \text{ Voice}]$
     Asp]

Kwapiszewski (2021) argues for the position of the secondary imperfective suffix below Voice and in this way also for separating the imperfectivizing suffix from the morphological aspect in Polish. He builds on Baker & Vinokurova (2009) and draws a parallelism between English nominals in *-er* and Polish agent/instrument *-acz/-arka* nominals. He shows that Polish *-acz/-arka* nominalizations can contain the imperfectivizing suffix but do not embed the Voice projection since they do not allow the relevant modifiers.

    The same argument can be done for the Czech counterpart: *-č* nominals (Russian does not have this form of nominals). The animate as well as the inanimate nominal contain the imperfectivizing suffix but do not allow agent-oriented modifiers, as demonstrated in (72).

(72) a. (*úmyslný) vy-jedn-a-va-č   (*, aby zabránil válce)
        deliberate out-one-TH-SI-NMLZ   so.that prevent war
        Intended: 'someone who (deliberately) negotiates (in order to avoid a
        war)'

    b. o-vlad-a-č   (*osobou)   (*s cílem měnit programy)
        about-rule-SI-NMLZ person.INSTR.SG with goal switch channels
        Intended: 'a remote control (used by a person) (for switching TV
        channels)'   (Czech)

Petr Biskup

Thus, in Czech, too, such nominalizations include the projection with the secondary imperfective suffix but are structurally smaller than VoiceP and by transitivity, also smaller than AspP. Beside separating the imperfective suffix from the imperfective interpretation, it also argues for the claim that prefixes are separated from the perfective interpretation in the aspectual projection. Because of the presence of the imperfectivizing suffix, at least lexical and lower superlexical prefixes are expected to be able to occur in this type of nominalizations. This seems to be correct, given the prefixed examples in (72).

If Baker & Vinokurova (2009) are correct in that agentive nominalizing morphemes like *-er* are nominal versions of the Voice head (having meanings similar to morphemes of Voice heads) that combine with the same complements as Voice does, then the order of the morphemes itself can be taken to mean that the projection of Voice is higher than the projection of the secondary imperfective suffix. The point is that the imperfectivizing suffix is always closer to the root than the agentive nominalizing morpheme.

It is possible to extend this reasoning to other agent nominalizations, e.g. to nominals ending in *-tel'* in Russian, *-tel* in Czech (and *-ciel* in Polish) and to Russian nominals with the suffixes -(*l'*)*ščik* and *-čik*, which are counterparts of the Czech *-č* discussed above. Such agent nominalizations can contain the imperfectivizing suffix and the suffix is always closer to the root than the agentive morpheme, independently of whether the nominal is inanimate (instrument), as in (73a), or animate, as in examples (73b) and (74). [16]

(73) a. pere-gruž-a-tel'
         over-load-si-NMLZ
         'a loader'

      b. ras-se-va-l'ščik
         apart-sow-si-NMLZ
         'a sorter'   (Russian)

(74) o-šetř-ova-tel
      about-spare-si-NMLZ
      'a keeper'   (Czech)

The consequences for dissociating prefixes and the secondary imperfective suffix from the corresponding morphological aspect interpretation are identical to those in the case of *-acz/-arka* and *-č* nominalizations discussed above.

---

[16]The underlying theme vowel /i/ brings about the palatalization of the root consonant /z/ in (73a); compare: *peregruzit'* 'to transfer'.

The current analysis with AspP above VoiceP, as discussed wrt. (71), goes against analyses like Zdziebko (2017: 571, 585), who argues that in Polish, the agentive VoiceP is placed above the aspectual projection(s). According to a reviewer, data like (75) suggest that in Polish, VoiceP is also higher than HabP since the habitual *-yw-* is inside the passive *-n-*.

(75) Ta melodia jest / była grywana   w wielu rozgłośniach radiowych.
      this melody is   was played.HAB in many stations   radio.
      'This melody is/was played in many radio stations.'   (Polish)

However, I assume that *-n-* in fact projects a participial phrase, as in Biskup (2016) and Biskup (2019: Chapter 4). PartP then includes HabP. An argument for HabP above VoiceP could be based on the fact that stem nominalizations can be agentive but cannot contain the habitual morpheme, like the Russian *pisy-vanie 'writing' and the Czech *psávání 'writing' in (45b). Since Polish habitual nominalizations like *pisywanie* 'writing' are grammatical, they can also contain HabP.[17]

Since the nominalizations under discussion typically refer to an instrument or an agent repeatedly performing the event expressed by the verb stem (they often contain the imperfectivizing morpheme, as in (72)–(74)), they are incompatible with the semelfactive suffix. Specifically, they conflict with the *cardinality one* property of the semelfactive morpheme, as defined in (37).

The next structural prediction is that the nominalizations under discussion cannot include the habitual marker for it is located above the aspectual projection. This prediction seems to be correct since e.g. the Czech National Corpus, SYN 8 (Křen et al. 2019) contains no agent nominalization that have the habitual marker and ends in *-vatel*.

Let us now consider the separation of the morphological aspect interpretation from the habitual marker. The habitual suffix is special. First, in contrast to the other aspectual markers, it occurs above the aspectual projection, as argued in

---

[17]In addition, given the reasoning in §1.4 that HabP is above AspP, the 'be' auxiliary in constructions like (i) cannot be placed in AspP, contrary to Błaszczak & Klimek-Jankowska (2012) and Błaszczak et al. (2014). As to the Russian habitual *igryvať* 'to play repeatedly', it is standardly claimed that such forms are colloquial and used only in the past tense (see §1.4 again).

(i) a. Jan będzie grywać w różnych lokalach w Londynie.   (Polish)
         Jan will   play.HAB in different pubs   in London.
     b. Jan bude hrávat v různých hospodách v Londýně.
         Jan will play.HAB in different pubs   in London.
         Both: 'Jan will play in London in various pubs.'   (Czech)

Petr Biskup

§1.4. Second, in contrast to the other markers, it does not reverse the morphological aspect value of the predicate to which it adjoins. Because of the second property, it in actuality does not have to be in a syntactic relation with the aspectual head. It suffices when it imposes the imperfective requirement on its complement. Moreover, given this selection property and the specific quantificational meaning of the marker, the habitual suffix can be treated as semantically independent from the aspectual head, which encodes the inclusiveness relation between the event time and the reference time.[18] Furthermore, since there are forms with the morphological aspect interpretation that exclude the habitual marker – recall the Czech stem nominalizations from §1.4 –, I conclude that the habitual marker can be separated from the aspectual phrase as well.

3 Deriving the morphological aspect value

As stated in the beginning of the preceding section, the operation Agree is very suitable for cases where a certain interpretation is separated from the element bringing out the interpretational effect. In our case, it is about perfective versus imperfective effects triggered by the four aspectual markers. For this reason, we need an interpretable unvalued aspect feature on the aspectual head and valued features on the aspectual markers. The feature on the aspectual markers (either perfective or imperfective) can value the unvalued feature on the head Asp and in this way, it can bring about the appropriate inclusiveness relation between the event time and the reference time.

   In the current proposal, I follow the Agree analysis by Biskup (2020) and assume that the secondary imperfective marker has an uninterpretable aspect feature with the imperfective value (recall the imperfectivizing effect of this suffix from §1.2). In contrast, since prefixes perfectivize the base predicate, as we saw in §1.1, they bear an uninterpretable aspect feature with the perfective value. The same also holds for the semelfactive marker because it also has the perfective effect, as discussed in §1.3. With respect to the habitual head, I argued in the preceding section that it has an imperfective selection feature and that it does not have to enter into an Agree relation with the aspectual head. However, the habitual head bears the imperfective aspect feature, which ensures that the marker *-yva-* can spell out it in accordance with the rule (61).

   If we make the standard assumption that lexical prefixes merge in the complement position of the root (e.g. Ramchand 2004, Svenonius 2004, Gehrke 2008,

---

[18]For the specific aspectual operators, see footnote 12.

Biskup 2019), then the (non-linearized) hierarchy with the four aspect markers and their aspect features looks like (76).

(76) $[_{\text{HabP}}\ \text{HAB}_{\text{ipf}}\ [_{\text{AspP}}\ \text{Asp}_{\text{asp-F:[ ]}}\ [_{\text{VoiceP}}\ \text{Voice}\ [_{\text{SPP}}\ \text{SP}_{\text{pf}}\ [_{\text{SIP}}\ \text{SI}_{\text{ipf}}\ [_{\text{SPP}}\ \text{SP}_{\text{pf}}$
      $[_{v\text{P}}\ \text{SEML}_{\text{pf}}\ [_{\sqrt{\text{P}}}\ \sqrt{}\ [_{\text{PP}}\ \text{LP}_{\text{pf}}]]]]]]]]]$

Assuming that morphemes are structurally heads, lexical prefixes head a prepositional phrase, the semelfactive marker heads the vP projection, superlexical prefixes head their own projection SPP and the habitual suffix heads the habitual projection. Superlexical projections can be iterated and occur either lower or higher than the projection of the imperfectivizing morpheme SIP.

The Agree analysis can successfully deal with the generalization MAG, that is, with the fact that the morphological aspect value is determined by the last attached aspectual morpheme. Specifically, using the standard operation of downward Agree, the last – structurally, the highest – aspectual marker can be determined on the basis of minimality, i.e. the structural distance from the probing aspectual head. The aspect feature of the closest marker will then value the unvalued aspect feature of the aspect head. Since only downward Agree is used, the habitual marker – occurring in a higher structural position – is not visible for the probing aspectual head. This, however, does not pose a problem because the marker cannot change the morphological aspect value, as already discussed above.

If it is correct that the verb moves to the head Asp, as argued by Gribanova (2013, 2015) for Russian, we receive the syntactic structure in Figure 1. Concretely, when the unvalued feature of the aspectual head probes, the complex verbal head is located in Voice. To determine the closeness of aspectual affixes and their features, I employ the concept of dominance. It is the head to which the moving element adjoins that projects, as demonstrated in the abstract structure in Figure 1. Since this head dominates the adjoined head, its features (among others, its valued aspect feature) are closer to the c-commanding aspectual head than the features of the adjoined head.

The complex Voice head in Figure 1 contains the following markers with their aspect features: a lexical prefix (the preposition), a lower superlexical prefix, the secondary imperfective suffix and a higher superlexical prefix. Therefore, the structure can represent predicates like the Russian *po-pere-za-pis-yva-t'* 'to re-record for a while'. The delimitative prefix *po-* merges in the higher superlexical position and the repetitive *pere-* merges in the lower superlexical position, i.e. below the secondary imperfective suffix *-yva-*. The lexical prefix *za-* is represented by the preposition in Figure 1. What is crucial here, is that the delimitative *po-*
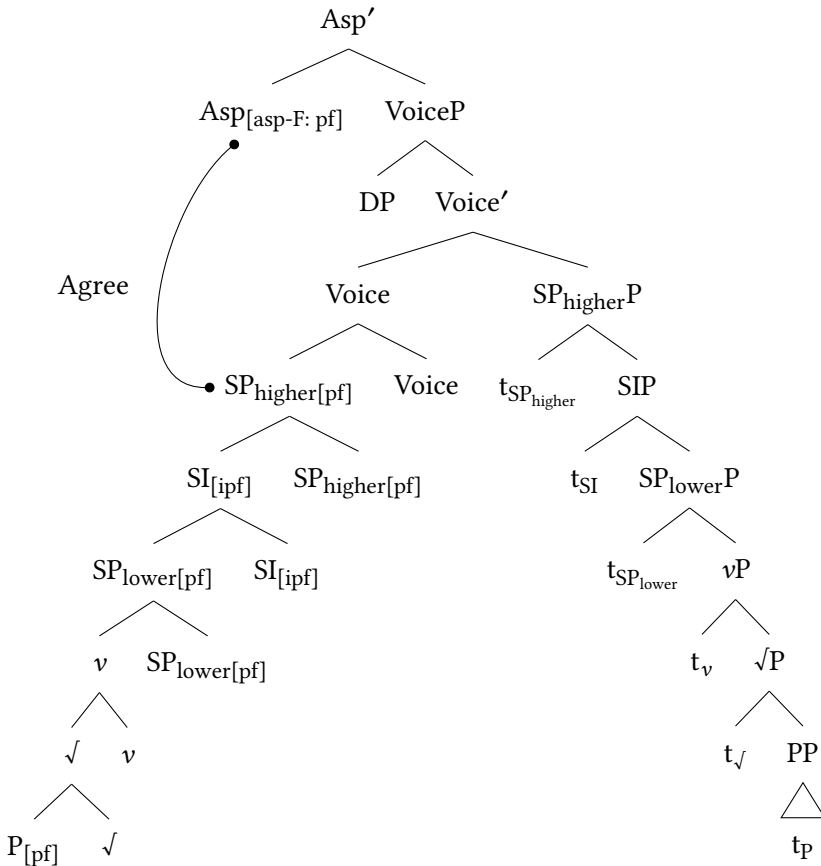
Petr Biskup

Figure 1: The derivation of the perfective morphological aspect

projects its perfective feature and dominates the SI constituent headed by *-yva-* with its imperfective aspect feature. Hence, it is the perfective feature of the delimitative *po-* that is the closest aspect feature and values the unvalued aspect feature on Asp. Consequently, the predicate is interpreted as perfective.

Nothing changes on the result, if the lower superlexical prefix is missing like in the perfective Russian example *po-vy-talk-iva-t'* 'to push out one after another' from §1.2. The distributive *po-*, with its perfective aspect feature, spells out the higher SP in Figure 1 and it is again the closest element to the aspectual head.

In contrast, if a single superlexical prefix merges in the lower SP position like in the Czech predicate *do-plét-a-t* 'to complete knitting' in (17b), the imperfective feature of the imperfectivizing suffix will be the closest aspect feature to Asp. Consequently, the imperfective operator will be used for the aspectual head.

It is obvious from the discussion that there can be aspectual markers with valued, uninterpretable aspect features that do not enter into an Agree relation (recall also the habitual head, which is not c-commanded by the probing Asp and bears a valued, uninterpretable imperfective feature). To cope with this issue, I assume that for the semantic interface, only unvalued features (but not uninterpretable features) are offending. Concretely, the uninterpretable property of a feature just signals that the feature should not be interpreted at the semantic interface (cf. Zeijlstra 2009). In other words, the interpretable versus uninterpretable property can indicate where (i.e. which occurrence of) the feature should be interpreted in the structure.

In the case of predicates containing a lexical prefix and the imperfectivizing suffix like the Russian *za-rabat-yva-ť* 'to earn' in (9b) and the Czech *vy-proš-ova-t* 'to beg' in (16b), we also receive the imperfective aspect because the mother SI node, with its imperfective feature, unambiguously dominates the P element (lexical prefix); consider the structure in Figure 1 again.

If only a lexical prefix attaches to the predicate, as in *na-kle-i-ť* 'to stick on' in (1b) and *vy-chov-a-t* 'to raise' in (2b), the aspectual head probes the whole way down in the complex Voice head and finally finds the only available aspect feature on P. This brings about the perfective interpretation. Obviously, the same result is obtained if a superlexical prefix is added to the lexical one, as in the Russian *pere-vy-poln-i-ť* 'to overfulfill' in (7a) and the Czech *pře-vy-chov-a-t* 'to re-educate' in (8a). There, however, it is the perfective feature of the superlexical prefix that values the aspectual head.

Since lexical prefixes merge in the complement of the root and then adjoin to it, it must be the root that projects its features in the complex verbal head. From this and the fact that lexical prefixes perfectivize the base predicate, it follows that the root cannot have an imperfective aspect feature. For this reason, I assume that the morphological aspect of simplex verbs is derived by a default mechanism. Specifically, if the probing aspectual head does not find an aspect feature in its c-command domain, it will receive the imperfective aspect value when it is sent to the interfaces (see Preminger 2014 for the claim that the operation Agree can fail). Note that this proposal is in line with the standard approach to Slavic aspect, which takes imperfectivity to be the default aspect value (see e.g. Jakobson 1932, 1956, Comrie 1976, Nübler et al. 2017). As to the root of the exceptional perfective simplex predicates like the Rusian and Czech *kupiť*/*koupit* 'to buy' and *dať*/*dát* 'to give', it bears a perfective feature, which is found by the probing aspectual head. Concerning bi-aspectual verbs, I assume that their root can optionally have the perfective feature (in addition to applying the default mechanism resulting in imperfectivity) until the aspect value of the predicate is settled.

Petr Biskup

With respect to the semelfactive marker, it was shown in §1.3 that the suffix combines with prefixes but does not co-occur with the secondary imperfective suffix and the habitual marker. Given that the semelfactive marker also bears an aspect feature and spells out the verbalizing head v, its perfective feature will value the aspect feature of Asp in the case of lexically prefixed predicates like the Russian *vs-krik-nu-t'* 'to give a scream' in (34) and, of course, in the case of unprefixed semelfactive verbs like *krik-nu-t'* 'to shout out' in (23b), which were discussed in §1.3.

On the contrary, in the case of superlexically prefixed semelfactive verbs like the Czech *na-prask-nou-t* 'to crack partially' in (35), it will be the perfective feature of the superlexical prefix that values the aspectual head (independently of whether it is a lower or a higher superlexical prefix) since any SP projected by a superlexical prefix always dominates v.

As discussed in sections §1.3 and §1.4, Russian and Czech stem nominalizations differ in the complexity of their structure, specifically, in the presence or absence of higher superlexical prefixes and the aspectual projection. In the case of Czech *-ní* nominals – which can contain higher superlexicals and have the morphological aspect – the morphological aspect value on the aspectual head will be derived as described above. In the case of Russian *-nie* nominals there is no Agree operation because they are aspectless and include the projection with the imperfectivizing marker at the most, plus the projection with the suffix *-n-/-t-* and the nominalizing projection nP; see (64) again. Here, the assumption that the uninterpretability of features just signals whether or not the appropriate (instance of the) feature should be interpreted at the semantic interface is applicable. This reasoning applies to all forms that lack the aspectual projection but contain an aspectual marker with an aspect feature, e.g. to the root nominalizations discussed in §1.3, which can include a lexical prefix.

The proposal in Figure 1 derives the correct order for all morphemes except superlexical prefixes. Given that prefixes display a peculiar behavior more generally, I assume that they also have weak prosodic properties which force them to linearize to the left (see e.g. Caha & Ziková 2016, who argue for a proclitic character of short verbal prefixes in Czech, and Biskup et al. 2011, who discuss differences between prefixed verbs and particle verbs in German and argue that in prefixed verbs the prepositional phonological word is weak in contrast to particle verbs).

4 Conclusions

I have argued that the four aspectual morphemes (prefixes, the secondary imperfective suffix, the semelfactive marker and the habitual suffix) are not exponents of the morphological aspect in Russian and Czech; they just work as a trigger of the corresponding aspectual interpretation. However, this is not to say that the aspectual markers are meaningless. They have their own meaning, which can be inner aspectual, as proposed e.g. for the semelfactive suffix in §1.3. I have shown that the morphological aspect value is determined by the last attached aspectual marker. The aspect value, I have derived by means of the operation Agree, using the concept of closeness based on dominance relations in the moved verbal head. The last-attached aspectual marker is the closest element with a valued aspect feature.

Abbreviations

ACC	accusative	IPF	imperfective
ATT	attenuative	LP	lexical prefix
COMP	completive	NMLZ	nominalizing affix
CUM	cumulative	NOM	nominative
DA	degree achievement	PART	participle
DEL	delimitative	PF	perfective
DIST	distributive	REP	repetitive
EXC	excessive	SEML	semelfactive
HAB	habitual	SI	secondary imperfective
INC	inceptive	SP	superlexical prefix
INF	infinitive	TH	theme (vowel)

Acknowledgments

References

Arsenijević, Boban. 2006. *Inner aspect and telicity: The decompositional and the quantificational nature of eventualities at the syntax-semantics interface.* Leiden: Leiden University. (Doctoral dissertation).

Petr Biskup

Babko-Malaya, Olga. 1999. *Zero morphology: A study of aspect, argument structure and case.* New Brunswick, NJ: Rutgers University. (Doctoral dissertation).

Babko-Malaya, Olga. 2003. Perfectivity and prefixation in Russian. *Journal of Slavic Linguistics* 11(1). 5–36. https://www.jstor.org/stable/24599703.

Baker, Mark C. & Nadya Vinokurova. 2009. On agent nominalizations and why they are not like event nominalizations. *Language* 85(3). 517–556. DOI: 10.1353/lan.0.0144.

Berger, Tilman. 2009. Anmerkungen zur Produktivität der tschechischen Iterativa. In Lenka Scholze & Björn Wiemer (eds.), *Von Zuständen, Dynamik und Veränderung bei Pygmäen und Giganten. Festschrift zum 60. Geburtstag von Walter Breu*, 25–43. Bochum: Universitätsverlag Dr. N. Brockmeyer.

Biskup, Petr. 2016. Prefixed adjectival participles. *Linguistica Brunensia* 64(1). 7–26. https://hdl.handle.net/11222.digilib/135447.

Biskup, Petr. 2019. *Prepositions, case and verbal prefixes: The case of Slavic.* Amsterdam: John Benjamins. DOI: 10.1075/la.255.

Biskup, Petr. 2020. An agree analysis of the morphological aspect in Slavic. Manuscript, Charles University, Prague. https://home.uni-leipzig.de/biskup/.

Biskup, Petr, Michael Putnam & Laura C. Smith. 2011. German particle and prefix verbs at the syntax-phonology interface. *Leuvense Bijdragen* 97. 106–135. DOI: 10.2143/LB.97.0.2977249.

Biskup, Petr & Gerhild Zybatow. 2015. Verbal prefixation in Slavonic: A minimalist approach. In Peter O. Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word-formation: An international handbook of the languages of Europe. (HSK 40.2)*, 1492–1515. Berlin: De Gruyter Mouton. DOI: 10.1515/9783110246278-042.

Błaszczak, Joanna, Patrycja Jabłońska, Dorota Klimek-Jankowska & Krzysztof Migdalski. 2014. The riddle of 'future tense' in Polish. In Philippe De Brabanter, Mikhail Kissine & Saghie Sharifzadeh (eds.), *Future times, future tenses*, 165–204. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199679157.003.0008.

Błaszczak, Joanna & Dorota Klimek-Jankowska. 2012. Futures in Polish and Slovenian: 'A hole in a sock' theory. In Alexander Podobryaev (ed.), *Formal Approaches to Slavic Linguistics 20: The second MIT Meeting 2011*, 17–32. Ann Arbor, MI: Michigan Slavic Publications.

Borer, Hagit. 2005. *The normal course of events: Structuring sense, Vol. II.* Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199263929.001.0001.

Caha, Pavel & Tobias Scheer. 2008. The syntax and phonology of Czech templatic morphology. In Andrei Antonenko, John Bailyn & Christina Y. Bethin (eds.),

Formal Approaches to Slavic Linguistics 16: The Stony Brook Meeting 2007, 68–83. Ann Arbor, MI: Michigan Slavic Publications.

Caha, Pavel & Marketa Ziková. 2016. Vocalic alternations in Czech prefixes: Evidence for prefix movement. *Acta Linguistica Hungarica* 63(3). 331–377. DOI: 10.1556/064.2016.63.3.3.

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

Egg, Markus. 2018. Semelfactives. *Oslo Studies in Language* 10(2). 65–81. https://www.journals.uio.no/index.php/osla.

Esvan, Francois. 2007. *Vidová morfologie českého slovesa*. Praha: Nakladatelství Lidové noviny.

Filip, Hana. 1993. On genericity: A case study in Czech. In *Proceedings of the 9th Annual Meeting of BLS 19*, 125–142. DOI: 10.3765/bls.v19i1.1516.

Filip, Hana. 2005. On accumulating and having it all: Perfectivity, prefixes and bare arguments. In Henk Verkuyl, Henriëtte de Swart & Angeliek van Hout (eds.), *Perspectives on aspect*, 125–148. Dordrecht: Springer. DOI: 10.1007/1-4020-3232-3_7.

Filip, Hana & Gregory N. Carlson. 1997. Sui generis genericity. In *Proceedings of the Twenty-First Annual Penn Linguistics Colloquium, Vol. 4*, 91–110. Philadelphia: The University of Pennsylvania.

Gehrke, Berit. 2008. *Ps in motion: On the semantics and syntax of P elements and motion events*. Utrecht: LOT. https://www.lotpublications.nl/ps-in-motion-ps-in-motion-on-the-semantics-and-syntax-of-p-elements-and-motion-events.

Gribanova, Vera. 2013. Verb-stranding verb phrase ellipsis and the structure of the Russian verbal complex. *Natural Language & Linguistic Theory* 31(1). 91–136. DOI: 10.1007/s11049-012-9183-3.

Gribanova, Vera. 2015. Exponence and morphosyntactically triggered phonological processes in the Russian verbal complex. *Journal of Linguistics* 51(3). 519–561. DOI: 10.1017/S0022226714000553.

Grzegorczykowa, Renata, Roman Laskowski & Henryk Wróbel. 1984. *Gramatyka współczesnego języka polskiego. Morfologia*. Warszawa: PWN.

Isačenko, Alexander V. 1962. *Die russische Sprache der Gegenwart. Teil 1. Formenlehre*. Halle (Saale): Max Niemeyer.

Jabłońska, Patrycja. 2007. *Radical decomposition and argument structure*. Tromsø: CASTL/University of Tromsø. (Doctoral dissertation).

Jakobson, Roman. 1932. Zur Struktur des russischen Verbums. In *Charisteria gvilelmo methsio qvinqvagenario a discipulis et circuli linvistici pragensis sodalibus oblata*, 74–84. Prague: Cercle Linguistique de Prague.

Petr Biskup

Jakobson, Roman. 1956. Shifters, verbal categories and the Russian verb. In Roman Jakobson (ed.), *Selected writings II: Word and language*, 130–147. The Hague: Mouton.

Karcevski, Serge. 1927. *Système du verbe russe*. Praha: Plamja.

Klimek-Jankowska, Dorota & Joanna Błaszczak. 2021. *The interaction of idioms and aspect in Polish.* Manuscript, University of Wrocław. www.researchgate.net/publication/348433153.

Klimek-Jankowska, Dorota & Joanna Błaszczak. 2022. The status of secondary imperfectivization in Polish: Evidence from VP idioms. *Journal of Slavic Linguistics* 30(FASL 29 extra issue). 1–19. http://ojs.ung.si/index.php/JSL/article/view/99.

Kopečný, František. 1962. *Slovesný vid v češtině*. Praha: Nakladatelství Československé akademie věd.

Křen, Michal, Václav Cvrček, Tomáš Čapka, Anna Čermáková, Milena Hnátková, Lucie Chlumská, Tomáš Jelínek, Dominika Kováříková, Vladimír Petkevič, Pavel Procházka, Hana Skoumalová, Michal Škrabal, Petr Truneček, Pavel Vondřička & Adrian Zasina. 2019. *Český národní korpus (SYN 8)*. Praha: Ústav Českého národního korpusu, FF UK. http://www.korpus.cz.

Kuznecov, Petr S. 1953. *Istoričeskaja grammatika russkogo jazyka*. Moskva: Nauka.

Kwapiszewski, Arkadiusz. 2021. Secondary imperfective is below voice: Evidence from *-acz/-arka* nominals and adjectival *-ąc* participles. Manuscript, University of Oxford. https : / / www . researchgate . net / publication / 352134509 _ Secondary_imperfective_is_below_Voice_Evidence_from_agentinstrument_ nominals_and_adjectival_active_participles.

Łaziński, Marek. 2020. *Wykłady o aspekcie polskiego czasownika*. Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.

Łazorczyk, Agnieszka. 2010. *Decomposing Slavic aspect: The role of aspectual morphology in Polish and other Slavic languages*. Los Angeles, CA: University of Southern California. (Doctoral dissertation).

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

Lunt, Horace G. 2001. *Old Church Slavonic grammar*. Berlin & New York: De Gruyter Mouton.

Markman, Vita. G. 2008. On Slavic semelfactives and secondary imperfectives: Implications for the split 'AspP'. In *University of Pennsylvania Working Papers in Linguistics 14(1)*, 255–268. https://repository.upenn.edu/pwpl/vol14/iss1/20.

Matushansky, Ora. 2009. On the featural composition of the Russian back yer. In Gerhild Zybatow, Uwe Junghanns, Denisa Lenertová & Petr Biskup (eds.), *Studies in formal slavic phonology, morphology, syntax, semantics and information structure: Proceedings of FDSL 7*, 397–410. Frankfurt am Main: Peter Lang.

Mueller-Reichau, Olav. 2020. Default aspect based on state change. *Rhema* 1. 90–105. DOI: 10.31862/2500-2953-2020-1-90-105.

Nandris, Grigore & Robert Auty. 1969. *Handbook of Old Church Slavonic*. London: The Athlone Press.

Nübler, Norbert. 2017. Iterativnost. In Petr Karlík, Marek Nekula & Jana Pleskalová (eds.), *Czechency − Nový encyklopedický slovník češtiny*. https://www.czechency.org/slovnik/ITERATIVNOST.

Nübler, Norbert, Petr Biskup & Susan Kresin. 2017. Vid. In Petr Karlík, Marek Nekula & Jana Pleskalová (eds.), *Czechency − Nový encyklopedický slovník češtiny*. https://www.czechency.org/slovnik/VID.

Paducheva, Elena V. 2015. *Glagoly byt' i byvat': istorija i sovremennost'*. Materialy konferencii DIALOGUE 2015. http://www.dialog-21.ru/digests/dialog2015/materials/pdf/PaduchevaEV.pdf.

Paslawska, Alla & Arnim von Stechow. 2003. Perfect readings in Russian. In Artemis Alexiadou, Monika Rathert & Arnim von Stechow (eds.), *Perfect explorations*, 307–362. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110902358.307.

Pazelskaya, Anna & Sergei Tatevosov. 2008. Otglagol'noe imja i struktura russkogo glagola. In Sergei G. Tatevosov & Vladimir A. Plungyan (eds.), *Issledovanija po glagol'noj derivacii*, 348–379. Moskva: Jazyki slavjanskich kultur.

Petr, Jan (ed.). 1986. *Mluvnice češtiny 1: Fonetika - fonologie - morfonologie a morfemika - tvoření slov*. Praha: Academia.

Preminger, Omer. 2014. *Agreement and its failures*. Cambridge, MA: MIT Press.

Procházková, Věra. 2006. *Argument structure of Czech event nominals*. MA thesis, Tromsø: University of Tromsø.

Ramchand, Gillian. 2004. Time and the event: The semantics of Russian prefixes. In Peter Svenonius (ed.), *Nordlyd 32(2): Slavic prefixes*. Tromsø: University of Tromsø. DOI: 10.7557/12.72.

Romanova, Eugenia. 2004. Superlexical versus lexical prefixes. In Peter Svenonius (ed.), *Nordlyd 32(2): Slavic prefixes*. Tromsø: University of Tromsø. DOI: 10.7557/12.69.

Romanova, Eugenia. 2006. *Constructing perfectivity in Russian*. Tromsø: University of Tromsø. (Doctoral dissertation).

Rothstein, Susan. 2004. *Structuring events: A study in the lexical semantics of aspect*. Malden, MA: Blackwell.

Petr Biskup

Scheer, Tobias. 2003. The key to Czech vowel length: Templates. In Petr Kosta, Joanna Błaszczak, Jens Frasek, Ljudmila Geist & Marzena Żygis (eds.), *Investigations into formal Slavic linguistics: Contributions of the Fourth European Conference on Formal Descriptions of Slavic Languages (FDSL IV) / Potsdam, November 2001*, 97–118. Frankfurt am Main: Peter Lang.

Schoorlemmer, Maaike. 1995. *Participial passive and aspect in Russian*. Utrecht: Utrecht University. (Doctoral dissertation).

Smith, Carlota S. 1991. *The parameter of aspect*. Dordrecht: Kluwer.

Švedova, Natal'ja Ju. (ed.). 1980. *Russkaja grammatika*. Moskva: Nauka.

Svenonius, Peter. 2004. Slavic prefixes inside and outside VP. In Peter Svenonius (ed.), *Nordlyd 32(2): Slavic prefixes*, 205–253. University of Tromsø. DOI: 10.7557/12.68.

Szucsich, Luka. 2014. Restriktionen bei mehrfacher Prä- und Suffigierung. In Hagen Pitsch (ed.), *Linguistische Beiträge zur Slavistik. XXI. JungslavistInnen-Treffen in Göttingen 2012*, 199–217. München: Otto Sagner.

Taraldsen Medová, Lucie & Bartosz Wiland. 2019. Semelfactives are bigger than degree achievements. *Natural Language & Linguistic Theory* 37(4). 1463–1513. DOI: 10.1007/s11049-018-9434-z.

Tatevosov, Sergei. 2008. Intermediate prefixes in Russian. In Andrei Antonenko, John F. Bailyn & Christina Y. Bethin (eds.), *Formal Approaches to Slavic Linguistics 16: The Stony Brook Meeting 2007*, 423–445. Ann Arbor, MI: Michigan Slavic Publications.

Tatevosov, Sergei. 2011. Severing perfectivity from the verb. *Scando-Slavica* 57(2). 216–244. DOI: 10.1080/00806765.2011.631782.

Tatevosov, Sergei. 2013. Mnozestvennaja prefiksacija i eë sledstvija (Zametki o fiziologii russkogo glagola) [Multiple prefixation and its consequences (Notes on the physiology of Russian verb)]. *Voprosy jazykoznanija* 3. 42–89.

Tatevosov, Sergei. 2015. Severing imperfectivity from the verb. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebova (eds.), *Slavic grammar from a formal perspective: The 10th anniversary FDSL conference*, 465–494. Frankfurt am Main: Peter Lang. DOI: 10.3726/978-3-653-05335-7.

Tatevosov, Sergei. 2020. On the aspectual architecture of Russian. Manuscript, Lomonosov Moscow State University. http://otipl.philol.msu.xn--ru%20%20aspectless_verb_3-m49l.1.pdf.

Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2). 143–160.

Vinogradov, Viktor V., Evgenija S. Istrina & Stepan G. Barchudarov. 1952. *Grammatika russkogo jazyka*. Moskva: Izdatel'stvo Akademii nauk SSSR.

Wiemer, Björn & Ilja A. Seržant. 2017. Diachrony and typology of Slavic aspect: What does morphology tell us? In Walter Bisang & Andrej Malchukov (eds.), *Unity and diversity in grammaticalization scenarios*, 239–307. Berlin: Language Science Press. DOI: 10.5281/zenodo.823224.

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

Zdziebko, Sławomir. 2017. On the structure and interpretation of Polish passives. *Acta Linguistica Academica* 64(4). 563–617. DOI: 10.1556/2062.2017.64.4.4.

Zeijlstra, Hedde. 2009. Dislocation effects, uninterpretable features, functional heads, and parametric variation: Consequences of conflicting interface conditions. In Kleanthes K. Grohmann (ed.), *Interphases: Phase-theoretic investigations of linguistic interfaces*, 82–114. Oxford: Oxford University Press.

Zinova, Yulia & Hana Filip. 2015. The role of derivational history in aspect determination. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebova (eds.), *Slavic grammar from a formal perspective: The 10th anniversary FDSL conference*, 595–609. Frankfurt am Main: Peter Lang. DOI: 10.3726/978-3-653-05335-7.

Chapter 4

Degree achievements from a Slavic perspective

Mojmír Dočekal[a], Lucia Vlášková[a] & Maria Onoeva[b]

[a]Masaryk University [b]Charles University

The evaluative behaviour of degree achievements (e.g., *cool, widen, lengthen, dry*) has been a puzzling problem for many linguists. The currently standard theory (Kennedy & Levin 2008) treats them as degree expressions based on different types of scales, which in turn influence the resulting evaluative or non-evaluative interpretation. While it may account for English, this theory faces empirical problems when confronted with cross-linguistic data. In this paper, we present an experiment on Russian exploring if verbal prefixes influence the (non-)evaluative interpretation of degree achievements. It follows from the results that prefixation is at least as important as the underlying scales for the cases we studied, which empirically challenges the scalar theory.

Keywords: degree achievements, evaluativity, prefix, Russian, Slavic, experimental evidence

1 Introduction

The current paper describes the relationship of evaluativity inferences of adjectival degree achievements with Slavic verbal morphology, namely verbal prefixes. We also report the results of an experiment testing said relationship in Russian degree achievements.

Degree achievements such as *increase* or *age* are typically analysed as verbs where the argument undergoes a positive scalar change, e.g., in the sentence *The river widened*, the degree of the river's width undergoes a positive change (= increases) along some relevant dimension (= width). A large group of degree

achievements consists of verbs derived from gradable adjectives, such as English *widen*$_V$ from *wide*$_A$ or *empty*$_V$ from *empty*$_A$. These deadjectival degree achievements will be the focus of the current paper and the experiment on Russian.

It is common to analyse gradable adjectives via the formal semantics notion of an underlying scale. The underlying scales can differ with regard to their openness. A scale is open when there are no endpoints specified; this leads to relative gradable adjectives, where the standard of comparison that is needed to license the positive form of such an adjective is supplied via the context of utterance. Hence, as an example, it will take different absolute lengths to be considered *a long desk* and *a long boat*.

On the other hand, a scale with at least one endpoint gives rise to an absolute gradable adjective: the upper-bounded adjectives have the maximum endpoint specified, the lower-bounded ones have the minimum, and closed-scale adjectives have both endpoints. The standard of comparison used in positive forms is then taken to be the specified endpoint of the given scale. Therefore, context does not play the same role as in the relative adjectives and there should be no difference in the degrees of dryness in *a dry desk* and *a dry boat*.

This division is supported by the different patterning of modifiers with different types of adjectives, as shown below. We take *almost* and *slightly* as examples of modifiers that are licensed only in particular situations: (i) for *almost*, the scale in question has to have the maximum endpoint specified, hence the acceptability with upper-bounded and closed-scale adjectives; whereas (ii) for *slightly*, it is the other way around – the scale needs a minimum endpoint, as in lower-bounded and closed-scale adjectives. Naturally, the scale with no specified endpoints does not accept either of the mentioned modifiers.

1. relative adjectives: **almost long*, **slightly tall*

2. absolute adjectives

    2.1 upper-bounded: *almost dry*, **slightly clean*

    2.2 lower-bounded: **almost dirty*, *slightly wet*

    2.3 closed-scale: *almost opaque*, *slightly transparent*

The scale typology will be important also while discussing degree achievements. Below, we will argue, following Kennedy & Levin (2008) among others, that the underlying scale of the adjective remains in the meaning of the derived degree achievement and influences its telicity and evaluativity behaviour.

This paper is structured as follows: §1.1 and §1.2 present an overview of the degree achievement research, as well as the currently standard scalar theory. In §2,

we turn to Slavic degree achievements with a focus on their prefixation pattern. §3 reports the experiment testing the evaluative inferences of Russian degree achievements. Finally, §4 summarises the article.

1.1 Telicity and evaluativity behaviour of degree achievements

Degree achievements are a puzzling group with regard to their telicity and evaluativity behaviour, as was first noted by Dowty (1979). Moreover, these two notions have been often confused in the previous literature due to the misunderstandings in the terminology. This section aims to delineate the two aspects and clarify the terminology used in this paper.

We understand TELICITY as a property of verb phrases that denote an action or an event with a specific endpoint. Let us first look at the telicity pattern in motion verbs as a basis of the later comparison to degree achievements.

According to the standard telicity test of the acceptability of the adverbial phrase *for/in an hour*, the predicate *walked* in (1a) is atelic (having no specific endpoint). However, when an argument is added, e.g., *to the pub*, the whole predicate *walked to the pub* in (1b) becomes telic and licenses the adverbial *in an hour*. The telic event is maximal in a sense that it reaches its goal, so the VP *walked to the pub* describes such events where its agens ends in the pub. Thus, motion verbs can change their telicity according to the supplied arguments.

(1)   a. John walked {for/*in} an hour.   ATELIC
      b. John walked to the pub {*for/in} an hour.   TELIC

On the other hand, as shown by (2), English degree achievement *cool* is ambiguous between the atelic interpretation (plausibly, the sentence would be true in such a situation where some decrease of the temperature in the tea occurred) and the telic interpretation (the most probable scenarios verifying the telic reading would be such where the tea reached the room's temperature). Moreover, the ambiguity seems not to be related to change of the argument like in (1).

(2)   The tea cooled {for/in} one hour.

Furthermore, what we refer to as EVALUATIVITY (following Brasoveanu & Rett 2018), is a property of (deadjectival) degree achievements whose corresponding adjectives instantiate a degree above a particular standard. In other words, a degree achievement is evaluative if it implies its base adjective in its positive form,

Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

as in (3a); and non-evaluative, if the implication does not hold, as in (3b).[1]

(3)   a.   The tea cooled in an hour. ⤳ The tea is cool.   EVALUATIVE
       b.   The tea cooled for an hour. ⤴̸ The tea is cool.   NON-EVALUATIVE

In English, the telic predicates indicated by the adverbial *in an hour* usually correspond with the evaluative interpretation, and vice versa, the atelic predicates indicated by *for an hour* are usually non-evaluative, although this is not always the case. However, as §2 will show, Slavic degree achievements can differentiate between the two notions on a more visible level. Nevertheless, the situation in Slavic languages is complicated by the fact that in degree achievements two notions of maximalization coincide and also interact: degree maximalization (called evaluativity in our article) and event maximalization (as described by Krifka 1992, Filip 2008 a.o.). More about it in §2.

1.2 Accounts of degree achievements

The pattern presented above lead some researchers (most notably Abusch 1986) to claim that all degree achievements are ambiguous between the evaluative and non-evaluative reading. But this is empirically incorrect, as was noticed by other linguists soon thereafter. Consider first the upper-bounded degree achievements *quieten*, *darken* and *ripen* in (4) from Kearns (2007: ex. 36–38). If all degree achievements were ambiguous between the evaluative and non-evaluative interpretation, the non-evaluative interpretation should warrant the acceptability of the continuation *but it wasn't A* (= the base adjective). The usual conclusion drawn from data like this is that English upper-bounded degree achievements strongly prefer the evaluative reading (see Hay et al. 1999, Kennedy & Levin 2008 a.o.) and that the ambiguity behaviour of English degree achievements is more an exception than a rule.

(4)   a.   The room quietened in a few minutes #but it wasn't quiet.
       b.   The sky darkened in an hour #but it wasn't dark.
       c.   The fruit ripened in five days #but it wasn't ripe.

The same point can be concluded from the lower-bounded degree achievements, since they seem to prefer the non-evaluative reading, which we illustrate with (5)

---

[1]Other terms used in the literature have been *positive* and *telic* for EVALUATIVE readings; and *comparative* and *atelic* for NON-EVALUATIVE readings. However, it is important to differentiate between telicity and evaluativity, hence the separate terminology in this paper.

from Davies (2009), where the most salient interpretation is that the hands are only partially wet. The general conclusion, then, seems to be that again, for lower-bounded degree achievements, the ambiguity treatment is empirically wrong.

(5) Wet your hands with warm water and mix the dough with your hands.

(COCA)

Finally, turning to relative degree achievements, despite examples like (2), they seem to strongly incline to the non-evaluative interpretation, as shown by the examples from Kennedy & Levin (2008: ex. 6) repeated here as (6).

(6) a. The gap between the boats widened {for/*in} a few minutes.
     b. The recession deepened {for/*in} several years.

To conclude, the current default theory of degree achievements (Hay et al. 1999, Kennedy & Levin 2008, Kennedy 2012), which is constructed in a way that is naturally reflecting the reported English contrasts, could be succinctly summarised as follows: (i) relative degree achievements tend to be interpreted as non-evaluative; (ii) lower-bounded degree achievements are, by default, interpreted as non-evaluative; (iii) upper-bounded degree achievements receive mostly evaluative interpretations; (iv) closed-scale degree achievements lead, by default, to the evaluative interpretation. This more or less summarizes the empirical landscape of English degree achievements but it is an open question how much the scalar theory is adequate for cross-linguistic data.[2]

Let us now introduce the mechanics of the standard scalar approach. It is based on analysing an adjective as a measure function of the type $\langle e, d \rangle$, returning the degree of an object on a scale along the relevant dimension. The measure function is then type shifted to a property of objects with the morphologically null element **pos** (first introduced by Kennedy 1997), which also supplies the contextual standard needed for the interpretation of relative adjectives.

Turning now to degree achievements, the scalar approach models them as a measure of change function, as seen in (7). It is built on top of the "regular" measure function and returns the degree of change on the appropriate scale that

---

[2]During the revision of our article there appeared a new work which very nicely covers degree achievements from the cross-linguistic picture, Martínez Vera (2021). Thanks to one of the anonymous reviewers for pointing us the article. In a future work we would be more than happy to integrate our findings with Martínez Vera (2021) but since Martínez Vera (2021) and our work build upon slightly different theoretical assumptions, such an integration would be non-trivial and alas is beyond the scope of this article.

Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

the particular object underwent during the event. The core of its meaning is a difference function $\mathbf{m}_\Delta$, which returns the difference between the degree at the initial and the final phase of the event ($_\Delta$-notation signals the difference function). Note also the difference scale $m^\uparrow$, ranging from the standard of comparison to the correlate, which represents a common ground between the analyses of degree achievements and comparative forms of adjectives. This means, that the difference scale *wide*$^\uparrow$ in (8), for example, would be the common meaning core of *widen* and a comparative form *wider than*: $^\uparrow$ measures object on a scale provided – scale of width in (8). Finally, we follow Henderson (2013) in extending Kennedy & Levin's (2008) notation, which allows the verbal measure function to access its arguments via theta-roles, as reflected in (7) and (8) – Θ are then in the particular sentences substituted for the individual theta-roles.

(7) *Measure of change*
     For any measure function $\mathbf{m}$, $\mathbf{m}^\Theta_\Delta = \lambda e\big[\mathbf{m}_{m^\uparrow(\Theta(e))(init(e))}(\Theta(e))(fin(e))\big]$

(8) $\lambda e\big[\text{WIDE}_{\text{WIDE}^\uparrow(\Theta(e))(init(e))}(\Theta(e))(fin(e))\big]$

The measure of change function in (8) is then type-shifted into a property of events, again via the morphologically null element **pos**. The application is exemplified by (9) for relative degree achievements, and (10) for absolute ones. The standard of comparison (**stnd**) is supplied on the basis of the Interpretive Economy principle in (11) from Kennedy & Levin (2008). In the case of absolute degree achievements, maximising the contributions of the elements means using the lexicalised endpoint of the underlying scale as the standard of comparison, e.g., the maximum endpoint of the upper-bounded *dry* in (10): the truth conditions then specify that there was an event e and the agens of the event dries over the course of e in a way which exceeds the standard for drying. On the other hand, the relative *long* in (9) has an open scale without an endpoint, so the **stnd** needs to be supplied via context and the event e exceeds any contextually provided degree d.

(9) The shadow of the tree lengthened.
     $[\![(9)]\!] = \exists e\big[\text{LONG}^{ag}_\Delta(e) \geq \mathbf{stnd}(\text{LONG}_\Delta) \wedge ag(e) = \sigma x.^*\text{SHADOW}(x)\big]$

(10) The shirt dried.
     $[\![(10)]\!] = \exists e\big[\text{DRY}^{ag}_\Delta(e) \geq \mathbf{stnd}(\text{DRY}_\Delta) \wedge ag(e) = \sigma x.^*\text{SHIRT}(x)\big]$

(11) *Interpretive Economy* (Kennedy & Levin 2008: ex. 18)
     Maximize the contribution of the conventional meanings of the elements of a sentence to the computation of its truth conditions.

2 Degree achievements in Slavic

Let us now turn to the data in focus: the Slavic degree achievements. The first important observation comes from the morphosyntactic realisation of Slavic degree achievements that is different from English. The majority of Slavic degree achievements seems to be perfective, prefixed verbs. This is supported by the data obtained from the national corpora of Czech (Křen et al. 2015), Slovak (SNK 2020) and Russian (RNC 2003–2020). For each language, we elicited three representative degree achievements and three other (transitive, unergative, and unaccusative) verbs and compared the proportions of prefixed vs unprefixed tokens within them. We ran the Fisher's test and concluded from the results (Czech: p-value $< 2.2e^{-16}$, OR $\approx$ 10.6; Slovak: p-value $< 2.2e^{-16}$, OR $\approx$ 9.5; Russian: p-value $< 2.2^{-16}$, OR $\approx$ 10.9) that throughout these Slavic languages, the degree achievements are approximately 10 times more probable to be prefixed than the other verb types. A full account of Slavic degree achievements would, of course, have to integrate the grammatical aspect as well, and compare imperfective vs perfective degree achievements, but statistics like this provide a good argument to start analysing Slavic degree achievements from perfective, prefixed verbs, as is the case of the current paper.

In the rest of the article, we focus on the prefixed Slavic degree achievements (for reasons mentioned above) but let us make some preliminary notes concerning the interaction of grammatical aspect with the scalar component of Slavic degree achievements. We acknowledge that such notes are nothing more than first steps in a full story which would integrate event and degree maximalization and that our notes cannot show appropriate respect to the enormous Slavic aspectual literature. But be it as it may, we follow Filip (2008) in her treatment of imperfective degree achievements as non-maximal. And that seems to hold even if the degree achievements are derived from upper-bounded scales. Consider (12) with the imperfective degree achievement *schnout* 'to dry' with the lexical scale based on the upper-bounded scale of the adjective *suchý* 'dry'. In this case, the non-maximal (atelic) interpretation of the imperfective aspect leads to the non-evaluative interpretation of the degree achievement and since the same is true for secondary imperfective version of the same verb, it seems probable that the decisive factor for imperfective degree achievements is the grammatical aspect which can override the scalar information. If we would apply the evaluativity test introduced in §1.1, it would yield the non-evaluativity (truth of (12) doesn't imply truth of the base adjective *suchý* 'dry' in the positive form). This is also the claim of Filip (2008) which (we believe) points in the right direction for the

Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

Slavic imperfective degree achievements but of course calls for a proper empirical verification.

(12) Prádlo schlo   dvě hodiny.
      laundry was-drying two hours
      'The laundry was drying for two hours.'   ATELIC/NON-EVALUATIVE

Turning now to the perfective degree achievements data, let us compare the prototypical English example presented in (2) with its Czech counterpart in (13). The different readings of the English degree achievement *cool* would be unambiguously expressed – depending on the particular prefix – by the following Czech predicates: the prefix *vy-* in (13a) yields the evaluative reading, which would be true in a situation where the tea reached, e.g., the room temperature or the temperature suitable for drinking. On the other hand, the prefix *o-* in (13b) distinctly signalises the non-evaluative reading, which would be verified by any decrease of the tea's temperature. Moreover, the native speakers of Czech would infer that a Czech sentence corresponding to English *The tea was cool* would follow only from (13a), not (13b).

(13) a. Čaj vy-chladl   za hodinu.
          tea from-cooled in hour
          'The tea cooled completely in an hour.'   EVALUATIVE
      b. Čaj o-chladl   za hodinu.
          tea around-cooled in hour
          'The tea cooled slightly in an hour.'   NON-EVALUATIVE

Notice, however, that the adverbial test we used in (13) classifies both sentences as telic, which corresponds with the fact that both prefixed verbs are perfective (here, we follow the standard approach to the relationship between the grammatical and the lexical aspect in Slavic languages, exemplified by Brecht 1985 a.o.). Despite the fact that both *vy-chladl* 'cooled completely' in (13a) and *o-chladl* 'cooled slightly' in (13b) are classified as telic, we can clearly see that in Czech (and generally in Slavic languages), the verbal morphology distinguishes the (non-)evaluative interpretation according to the prefix that is used. Notice as well that the evaluativity classification fits nicely with the standard theory's (Kennedy & Levin 2008, Kennedy 2012) emphasis on the core of the adjectival meaning that unites degree achievements and their corresponding adjectives.

As discussed above, Slavic languages allow disambiguation of such degree achievements like English *cool* via different prefixes. But even more importantly, in some cases, the prefixes can override their default interpretation. By way of

example, the English upper-bounded degree achievement *dry* is predicted by the standard theory to be evaluative by default. But the Czech perfective example in (14) shows that depending on the nature of prefix, the degree achievement {*o-/vy-*}*schnout* 'dry' can be interpreted either as non-evaluative in (14a) or evaluative in (14b). This pattern is general: for all types of degree achievements, absolute or relative, we can construct both evaluative and non-evaluative versions by various prefixes.[3] So, next to the imperfective degree achievements, as in (12), non-evaluative perfective degrree achievements can be found too. The data and theory that aims at explaining this Slavic degree achievement pattern can be found in Dočekal & Vlášková (2021).

(14)   a.   Dřevo o-schlo,   ale pořád bylo většinou vlhké.
           wood around-dried but still   was mostly   wet
           'The wood dried slightly, but it was still mostly wet.'

       b.   Dřevo vy-schlo,   #ale pořád bylo většinou vlhké.
           wood from-dried but was   still mostly   wet
           'The wood dried completely, #but it was still mostly wet.'

In this article, we focus on the empirical properties of Slavic degree achievements and test them experimentally, but let us note that, semantically, Slavic prefixation of degree achievements resembles the English degree modifiers like *completely* or *partially*. As Kennedy & Levin (2008) notice while discussing their example (29) repeated bellow as (15), such degree modifiers can override the default interpretation of closed-scale degree achievements like *fill*. The default interpretation is supported with the degree modifier *completely* in (15a) but coerced to the non-evaluative interpretation with *partially* in (15b).

(15)   a.   The basin filled completely in 10 minutes.
       b.   The basin filled partially ??in 10 minutes.

In this respect, Slavic prefixes and English degree modifiers resemble each other semantically, but there are still some important differences: the first is the near obligatory presence of prefixes on Slavic verbs (as noted above); the second concerns the relative degree achievements. Consider (13a) again: the degree achievement is constructed on the open scale, so how can we even attain the evaluative interpretation, when there is no clear scalar boundary to be reached? One reasonable way to understand this theoretically is to propose that at least some degree

---

[3]Inspired by Zwarts (2005), we categorise verbal prefixes according to their cumulativity (bounded/unbounded nature) into evaluative and non-evaluative (in his terminology, telic vs atelic, respectively).

Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

achievements are variable with respect to their scales and in case like (13) they allow both relative and bounded scale. While in English the difference between the scales would be left for the context, Slavic languages can signal the nature of the scale morphologically. And because that, Slavic relative degree achievements can get the evaluative interpretation with the right kind of prefixes. Again here we seem to be following Filip (2008) when she claims that perfective degree achievements (at least with bounded scales) are always maximal in terms of the event structure and by default also evaluative but their evaluativity can be contextually overriden. Our experimental research can be then understood as a search for morphological clues determining the factors which Filip (2008) claims to be contextual.

To summarise this section: once we move beyond the territory of English degree achievements and focus on Slavic, we seem to see two sources of the (non-)evaluative interpretation: (i) the scalar lexical information inherited from the source adjectives; (ii) the degree modifiers and their contribution to the evaluative profile of the degree achievement. And this leads us to the research question behind our experiment, formulated in (16). It is clear that both factors (nature of the scale and the prefixation type) play a role, but only a controlled experiment can give us hints about their relative strength.

(16)   What are the factors of the evaluative interpretation in the case of Slavic degree achievements?

3 Experiment

In order to find out what the factors of the evaluative interpretation of Slavic degree achievements is, prefixes or adjectival scales, we conducted an experiment on Russian. We put the research question above reformulated into three subquestions in (17).

(17)   a.   How much does the lexical semantics of Russian degree achievements influence their evaluativity?
       b.   How much does the prefix of Russian degree achievements affect their evaluativity?
       c.   Which of the two factors is stronger (at least in terms of statistics)?

The measuring of the experimental results can give us at least partial answers to the questions above. And the most interesting question is the third one: such a

question is also not answerable by native speakers' intuition, that can otherwise give reasonable hints in case of the two previous sub-questions.

This section is structured as follows: we briefly describe the design of the experiment in §3.1, present its outcome in §3.2 and analyse the results in §3.3. The experiment was carried out as a part of a Master's thesis of one of the authors. Therefore, the following section borrows from Onoeva (2021).

3.1 Design

The experiment was completed by 165 native speakers, but the data of three of them were excluded due to low reliability discovered via their filler ratings. The experiment was a coherence acceptability task. The subjects evaluated how justified is a reasoning from indirect speech containing a degree achievement to a sentence containing an adjective in a positive form on a Likert scale from 1 'completely unacceptable' to 5 'completely acceptable'. The design was 2 × 2 with 4 conditions. Each participant saw 8 items and 8 fillers. A total of 16 stimuli was randomised for each participant. L-Rex platform by Starschenko & Wierzba (2021) was chosen for hosting.

The degree achievements tested in the experiment are present in Table 1. The absolute/relative adjectival distinction was used, thus, we divided the degree achievements into two groups. Then, we found the evaluative and non-evaluative prefixes for each verb. Whether the prefixes contribute total or partial reading was decided based on the judgements of the author of the experiment who is a native speaker of Russian. We were looking for the verbs which allow both types of prefixes, otherwise they were not suitable.

Table 1: The lists of the adjectives and DAs used in the experiment

adjective	eval. DAs	non-eval. DAs
relative		
gorjačij 'hot'	*razo-greť*	*po-greť*
nizkij 'low'	*s-niziťsja*	*po-niziťsja*
bednyj 'poor'	*o-bedneť*	*po-bedneť*
korotkij 'short'	*u-korotiť*	*pod-korotiť*
absolute		
suxoj 'dry'	*vy-soxnuť*	*pod-soxnuť*
polnyj 'full'	*na-polniť*	*po-polniť*
mokryj 'wet'	*vy-močiť*	*po-močiť*
čistyj 'clean'	*vy-čistiť*	*po-čistiť*

Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

The items always consisted of two sentences: the first one in indirect speech with a degree achievement, (18a) and (19a), the second with its core adjective in a positive form, (18b) and (19b). As mentioned above, the absolute/relative adjectival distinction was used. In (18), there is an example of the verb derived from an absolute adjective *suchoj* 'dry', while in (19), *gorjačij* 'hot' is relative. When it comes to the prefixes, these are *vy-* 'out' and *razo-* 'from' contributing the total reading in the given examples, then *pod-* 'under' and *po-* 'along, on' providing only partial.

(18) a. Detektiv Smit s   mesta prestuplenija soobščil svoemu kollege
          Detective Smith from scene crime   reported his   colleague
          detektivu Džonsonu, čto rubaška na sušilke   {vy-soxla,
          detective Johnson   that shirt   on drying-rack out-dried
          pod-soxla}.
          under-dried
          'Detective Smith reported to his colleague detective Johnson from
          a crime scene that a shirt dried on a drying rack.'

      b. Detektiv Džonson rešil,   čto rubaška byla suxaja.
          Detective Johnson concluded that shirt   was dry
          'Detective Johnson concluded that the shirt was dry.'

(19) a. Detektiv Smit s   mesta prestuplenija soobščil svojemu kollege
          Detective Smith from scene crime   report his   colleague
          detektivu Džonsonu, čto ubityj   prjamo pered smerťju
          detective Johnson   that murdered just   before death
          {razo-grel, po-grel} edu.
          from-hot   on-hot food
          'Detective Smith reported to his colleague detective Johnson from
          a crime scene that a murdered man warmed food right before his
          death.'

      b. Detektiv Džonson rešil,   čto eda v moment prestuplenija
          Detective Johnson concluded that food in moment crime
          byla gorjačaja.
          was hot
          'Detective Johnson concluded that food was warm at the time of the
          crime.'

We tested whether the subjects interpret the meaning of a particular degree achievement as evaluative (then the continuation with the positive form of an

adjective should be acceptable for them) or as non-evaluative (in which case the continuation should be rejected). In other words, we used the evaluative criterion discussed above in form of a coherence acceptability task. Generally, the expectation was that the speakers will accept the evaluative prefix with the absolute degree achievements better than other types of degree achievements.

We used the same structure with a verb in the first sentence and a corresponding adjective or past participle in the second for the fillers. They were also divided into two sets: good (4–5 ratings expected) and bad (1–2 ratings expected). The verbs in the good fillers were always perfective, e.g., *postroiť* 'to built' or *vypiť* 'to drink out', therefore, the participants could conclude that the second sentence was completely acceptable, while in the bad set, all the verbs were imperfective, e.g., *čitať* 'to read' or *pisať* 'to write', so they should be unacceptable in the given contexts.

3.2 Results

It was expected that the degree achievements with the evaluative prefixes should be accepted more, as they denote the finite state reading which should be equal to the meaning of the corresponding adjectives in their positive form. However, from a descriptive statistics of the experiment presented in Table 2 and Figure 1, it follows that this was not always the case.

Table 2: Measures of central tendency

item	mean	median
absolute + non-evaluative	3.01	3
absolute + evaluative	3.96	4
relative + non-evaluative	2.80	3
relative + evaluative	2.90	3

The degree achievements derived from the absolute adjectives with the evaluative prefixes were accepted better in comparison with the non-evaluative ones, whereas there is no big difference in acceptability of the relative degree achievements. With the aim of checking what happened inside the classes and to get a detailed view, we also looked at each item separately, see Figure 2.

The degree achievements from the absolute class (grey filled boxes) fall under the expected pattern: the verbs with the non-evaluative prefixes (black framed boxes) have lower acceptability rates than the verbs with the evaluative ones,
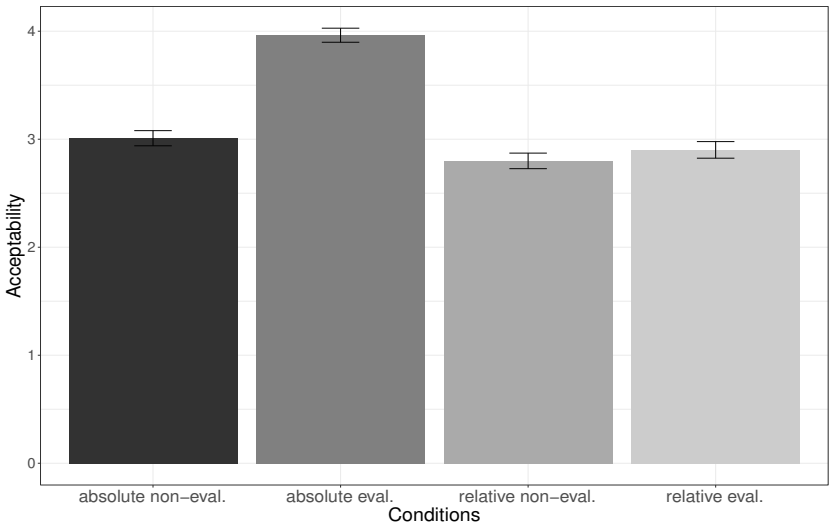
Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

Figure 1: Standard error graph

which are favoured in general. Nevertheless, the non-evaluative variants of *čistyj* 'clean' and *mokryj* 'wet' climbed higher than the other two and have the same medians as their evaluative counterparts.

When it comes to the relative class (white filled boxes), it is clear that *gorjačij* 'hot' was placed on top of it. Even though its evaluative variant *razogreť* 'heat up' was definitely liked better, non-evaluative *pogreť* 'heat up' also has the median rating 4. The degree achievements based on *nizkij* 'low' and *bednyj* 'poor' correspond to the expected pattern, but their acceptability was lower in general. A curious thing happened to *korotkij* 'short': both verbs were rated relatively low, but according to the mean ratings, what we considered to be the non-evaluative variant *podkorotiť* 'shorten', was slightly better accepted than evaluative *ukorotiť* 'shorten'.

We analysed the experimental data in a mixed-effects linear model with subject and item intercept+slope random effects using the LME4 package (Bates et al. 2015) in R (R Core Team 2021). The explanatory variables were conditions DA-CLASS (values: relative and absolute), PREFIX (values: evaluative, non-evaluative) and their interaction. The dependent variable was the subject's rating. The reference levels were absolute and non-evaluative for the conditions DACLASS and PREFIX respectively.

We found a strong positive effect of the evaluative prefixes: t-value = 11.437, $p <$ 0.001, which was the strongest effect recorded. Next, there was also a negative

4 Degree achievements from a Slavic perspective
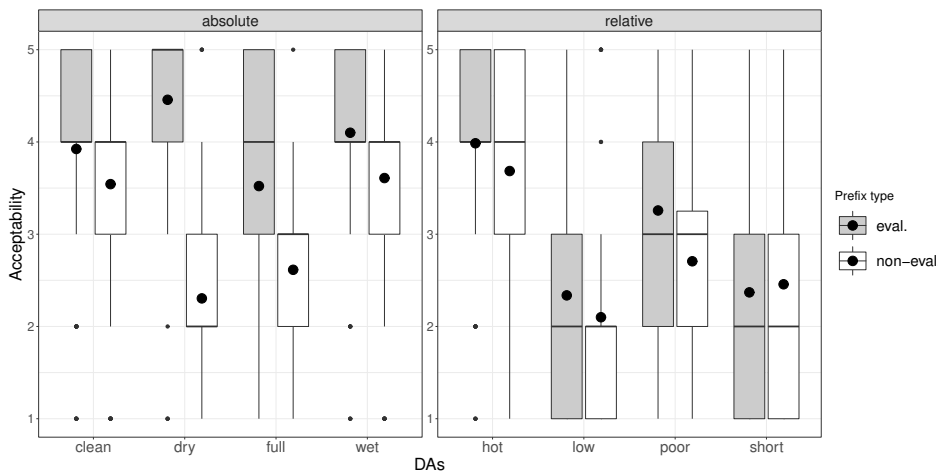
Figure 2: Box plot graph for each degree achievement

effect of the relative degree achievement class: t-value $= -2.318$, $p < 0.05$. And a negative interaction of the relative degree achievement class by the evaluative prefixes was found: t-value $= -6.652$, $p < 0.001$. The coefficients are reported in Table 3.

Table 3: Linear mixed model

	Estimate	SE	t-value	p-value
(Intercept)	2.96587	0.21794	13.609	<0.001
DAClassRELATIVE	−0.20988	0.09056	−2.318	0.02
PREFIXEVAL	1.04047	0.09097	11.437	<0.001
DAClassRELATIVE:PREFIXTEVAL	−0.85185	0.12807	−6.652	<0.001

3.3 Discussion

Now we can answer the research questions, for convenience repeated in (20):

(20) a. How much does the lexical semantics of Russian degree achievements influence their evaluativity?

  b. How much does the prefix of Russian degree achievements affect their evaluativity?

  c. Which of the two factors is stronger?

Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

The descriptive statistics and the model give some answers to both first and second question. Firstly, the negative effect of the relative degree achievement (DAClassrelative) class shows that in Russian, the lexical semantics of the degree achievements clearly affect their non-evaluative interpretation. The subjects judged the inference to the positive form of the corresponding adjective as less acceptable in items with relative degree achievements (which is already predicted by the standard theory).

Secondly: the strongest effect (the positive effect of the evaluative prefix: prefixeval) seems to show that at least in the material we tested the nature of the prefix was a stronger factor than the nature of the scale (see also Dočekal & Vlášková 2021). But of course it is not straighforward to translate strength of the statistic effects into the linguistic theory, so we don't want to jump to too hasty conslusion. Nevertheless, the most intriguing is the last question: simply comparing the strength of the main effects indicates that prefixation (at least for the verbs we tested) is the more important factor. But the interaction between the two factors also shows that the picture is not that clear: the negative interaction seem to be a reflex of the observed pattern in judgements – the evaluative prefix (which improves the acceptance with absolute degree achievements) plays significantly smaller role in the case of relative degree achievements.

Why do the speakers have problems accessing the evaluative interpretation with relative degree achievements is a very important question and the standard theory gives an answer: it is because relative degree achievements do not have scalar boundaries. But the answer faces some difficulties when we look at the absolute degree achievements where the prefix clearly plays the most important role and overrides the lexical information. Theoretical conclusions which can be drawn from the results of our experiment are divergent. One possibility would be to claim that some degree achievements are able to be linked with both relative and bounded scales and the nature of the prefix then determines the scale: if the degree achievement is prefixed with a non-evaluative prefix and it can be associated with both relative and upper-bounded or closed scale, the degree achievement would choose the relative scale (and the reverse pattern for the evaluative prefix).[4] But there is still the interaction effect which (simply put) tells us that it is easier (for subjects) to un-maximize the absolute degree achievements via some non-evaluative prefix but the reverse strategy: to maximize relative degree achievements is much harder. At this stage of work we simply report this asymmetry and offer some ideas above but a real theoretical description of what is going on is left for a future work.

---

[4]Thanks to the one of our anonymous reviewers for pointing us importance of this possibility.

118

4 Summary

In our article, we summarised the Slavic degree achievements data, which pose an empirical problem for the standard theory. More importantly, we reported the results of our experiment, which basically gives us some preliminary answers to the research questions in (17)/(20). Namely, the evaluative profile of Slavic degree achievements is related both to the lexical semantics (the nature of the scale, as predicted by the standard theory) and to the prefixes which modify the degree achievements. The nature of the prefix is, as it appears from the experiment, the more important factor, at least for the the absolute degree achievements, and for the relative ones the effect is real too but its impact is smaller.

But of course, as usually in the problem solving cycle, the answers we got from the experiment just mean starting another cycle of research questions, experiments and their analysis. Let us list some of the open questions which naturally appear: (i) Why do absolute and relative degree achievements show different sensitivity to prefixes? (ii) Is there some semantic (or other) criterion that distinguishes the evaluative prefixes from the non-evaluative ones? (iii) Why are some degree achievements perfectly fine without any prefix attached, while the others require it to be felicitous?

One possible answer to the first open question is the following: the relative degree achievements allow the evaluative interpretation (signalled via prefixation) only if they allow scalar variability as suggested above. This hypothesis can be tested in an experiment measuring both scalar variability of a particular relative degree achievement and its openness for evaluative prefixation. The second question is more theoretic in nature and some possible answers to it are given in Dočekal & Vlášková (2021), but see also Filip (2008) or Martínez Vera (2021) for a more general perspective; again, the differing theoretical routes are good candidates for experimental testing. The third question is a more general one without a clear answer, but a possible route here would be experimentally targeting Slavic imperfective degree achievements and their evaluativity behaviour. In the end, it seems that we ended up with more open questions than we started with, but that is (hopefully) a promise for a fruitful future work.

Acknowledgements

Mojmír Dočekal, Lucia Vlášková & Maria Onoeva

References

Abusch, Dorit. 1986. *Verbs of change, causation and time*. Tech. rep. Center for the Study of Language & Information. Stanford University. 86–50.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.

Brasoveanu, Adrian & Jessica Rett. 2018. Evaluativity across adjective and construction types: an experimental study. *Journal of Linguistics* 54(2). 263–329. DOI: 10.1017/S0022226717000123.

Brecht, Richard D. 1985. The form and function of aspect in Russian. *Issues in Russian morphosyntax* 10. 9–34.

Davies, Mark. 2009. The 385+ million word Corpus of Contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics* 14(2). 159–190. DOI: 10.1075/ijcl.14.2.02dav.

Dočekal, Mojmír & Lucia Vlášková. 2021. Evaluativity of degree achievements and verbal prefixes: Evidence from Czech morphology. *Acta Linguistica Academica* 68(4). 536–552. DOI: 10.1556/2062.2021.00427.

Dowty, David R. 1979. *Word meaning and Montague grammar: the semantics of verbs and times in generative semantics and in Montague's PTQ* (Studies in Linguistics and Philosophy 7). Dordrecht: Reidel. DOI: 10.1007/978-94-009-9473-7.

Filip, Hana. 2008. Events and maximalization: The case of telicity and perfectivity. In Susan D. Rothstein (ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect* (Linguistik Aktuell/Linguistics Today 110), 217–256. Amsterdam: John Benjamins. DOI: 10.1075/la.110.10fil.

Hay, Jennifer, Christopher Kennedy & Beth Levin. 1999. Scalar structure underlies telicity in "degree achievements". In Tanya Matthews & Devon Strolovitch (eds.), *Semantics and linguistic theory 9*, 127–144. Ithaca, NY: CLC Publications. DOI: 10.3765/salt.v9i0.2833.

Henderson, Robert. 2013. Quantizing scalar change. In Todd Snider (ed.), *Proceedings of the 23rd semantics and linguistic theory conference*, 473–492. DOI: 10.3765/salt.v23i0.2674.

Kearns, Kate. 2007. Telic senses of deadjectival verbs. *Lingua* 117. 26–66.

Kennedy, Christopher. 1997. Comparison and polar opposition. In Aaron Lawson (ed.), *Semantics and linguistic theory 7*, 240–257. Ithaca, NY: CLC Publications. DOI: 10.3765/salt.v7i0.2782.

Kennedy, Christopher. 2012. The composition of incremental change. en. In Violeta Demonte & Louise McNally (eds.), *Telicity, Change, and State*, 103–121. Oxford University Press. DOI: 10.1093/acprof:oso/9780199693498.003.0004. (3 April, 2020).

Kennedy, Christopher & Beth Levin. 2008. Measure of change: the adjectival core of degree achievements. In Louise McNally & Christopher Kennedy (eds.), *Adjectives and adverbs: syntax, semantics and discourse*, chap. 7, 156–182. Oxford: Oxford University Press.

Křen, M., V. Cvrček, T. Čapka, A. Čermáková, M. Hnátková, L. Chlumská, T. Jelínek, D. Kováříková, V. Petkevič, P. Procházka, H. Skoumalová, M. Škrabal, P. Truneček, P. Vondřička & A. Zasina. 2015. *SYN2015: Reprezentativní korpus psané češtiny*. Praha: Ústav Českého národního korpusu, FF UK. https://korpus.cz.

Krifka, Manfred. 1992. Thematic Relations as Links between Nominal Reference and Temporal Constitution. In Ivan A. Sag & Anna Szabolcsi (eds.), *Lexical Matters*, 29–53. Stanford, CA: CSLI Publications.

Martínez Vera, Gabriel. 2021. Degree achievements and maximalization: a cross-linguistic perspective. *Glossa: a journal of general linguistics* 6(1). 94. 1–28. DOI: 10.16995/glossa.5883.

Onoeva, Maria. 2021. *Inferential behavior of degree achievements: An experimental study*. Masaryk University, Faculty of Arts, Brno. (MA thesis). https://is.muni.cz/th/p5uq5/.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/.

RNC. 2003–2020. *Russian National Corpus – Nacionaľnyj korpus russkogo jazyka*. Moskva: Institut russkogo jazyka imeni V.V. Vinogradova, RAN. https://ruscorpora.ru/new/index.html.

SNK. 2020. *Slovenský národný korpus – prim-9.0*. Bratislava: Jazykovedný ústav Ľ. Štúra, SAV. https://korpus.juls.savba.sk.

Starschenko, Alexej & Marta Wierzba. 2021. *L-Rex Linguistic rating experiments [software], version beta*. GNU General Public License v3.0. https://github.com/2e2a/l-rex/.

Zwarts, Joost. 2005. Prepositional aspect and the algebra of paths. *Linguistics and Philosophy* 28(6). 739–779. DOI: 10.1007/s10988-005-2466-y.

Chapter 5

"True" imperfectivity in discourse

Berit Gehrke

Humboldt-Universität zu Berlin

By taking into account the broader discourse structure, I show that a standard imperfective (IPFV) semantics can also account for cases in Russian where IPFV forms describe actually completed events, thereby refuting an analysis of such forms as "fake" IPFVs with a perfective (PFV) semantics. The proposed account captures the general intuition that the use of the IPFV is conditioned by a particular discourse structure, in which the event described is already part of the common ground, and the IPFV sentence elaborates on this event, zooming in on a narrower reference time. The proposal also has repercussion for definitions of the PFV and encourages us to take a closer look also at the role of PFV beyond the sentential level.

Keywords: Russian aspect, imperfective, perfective, discourse, general-factual, presupposition

1 Introduction

Cross-linguistically, the perfective aspect (PFV) is assumed to involve the event time (or situation time) being included in the reference time (or topic/assertion time) (e.g. Klein 1995: for Russian), resulting in an external perspective on a given event, or PFV predicates denote whole events (e.g. Filip 1999, Altshuler 2014: for Russian). With the imperfective aspect (IPFV), on the other hand, the reference time (topic/assertion time) is taken to be included in the event time (situation time), giving rise to an internal perspective on a given event, or IPFV predicates denote partial events. In addition, there is a common intuition that completed events involve PFV semantics. The notion of a "completed event" in this context is usually just an intuitive notion and never properly defined. Nevertheless, this intuition is commonly thought to be problematic for Russian, in which IPFV

forms appear in descriptions of (intuitively) completed events, most famously in the so-called general-factual use, and this has led Grønn (2015) to claim that the Russian ɪᴘғᴠ is a "fake" ɪᴘғᴠ in these contexts. In particular, he argues that ɪᴘғᴠ forms in these contexts have a ᴘғᴠ semantics, thereby giving up on the otherwise attractive idea that (here: Russian) ɪᴘғᴠ forms have a uniform ɪᴘғᴠ semantics.

In this paper, I will argue that there is no "fake" ɪᴘғᴠ in Russian but that a uniform semantics for ɪᴘғᴠ forms succeeds if we take into account the discourse structure in which these forms occur. §2 provides background information on Russian aspect, characterises general-factual uses of the ɪᴘғᴠ, and discusses prominent accounts of the semantics of ɪᴘғᴠ that also aim at dealing with general-factuals. In §3, I will call into question the analytical move to take the intuition of event completion at the sentence level as a basis for analysing ɪᴘғᴠ forms as involving ᴘғᴠ semantics, by showing that event non-completion is neither a necessary nor sufficient condition for the use of ɪᴘғᴠ forms, and moreover, that event completion is not a necessary or sufficient condition for the use of ᴘғᴠ forms, either. In §4, I will demonstrate how we can still work with a "proper" ɪᴘғᴠ semantics for the given ɪᴘғᴠ forms when we take into account the discourse structure in which these forms occur. §5 concludes.

2 Grammatical aspect in Russian

This section provides background information on grammatical aspect in Russian, the canonical and non-canonical readings of the ɪᴘғᴠ, in particular factual ones, and outlines recent proposals with a focus on how they deal with factual ɪᴘғᴠs.

2.1 Background on Russian aspect morphology

Like all Slavic languages, Russian has a grammatical category aspect. This means that a given verb form is either ɪᴘғᴠ or ᴘғᴠ. Identical lexical meaning can be expressed by ɪᴘғᴠ and ᴘғᴠ verb forms, and there is the common assumption that many verb(form)s come in aspectual pairs. The received view is that one type of aspectual pair is derived from simple ɪᴘғᴠs by so-called "empty" prefixes; see (1).

(1) a.   ɪᴘғᴠ *pit'* > ᴘғᴠ *vy-pit'*   'to drink'
     b.   ɪᴘғᴠ *risovat'* > ᴘғᴠ *na-risovat'*   'to draw'

Another type of aspectual pair involves a suffix deriving an ɪᴘғᴠ from a ᴘғᴠ; cf. (2).

(2) a. PFV *pro-dat'* > IPFV *pro-da-va-t'*   'to sell' (lit. through-give)

   b. PFV *ot-kryt'* > IPFV *ot-kry-va-t'* 'to dis-cover, open' (lit. from-cover)

   c. PFV *dat'* > IPFV *da-va-t'*   'to give'

Given that such suffixes most often attach to already prefixed verbs (but not always, see (2c)), such derivations are descriptively labeled SECONDARY IMPER-FECTIVES (SI). There are other types of aspectual pairs, which I set aside for now, namely suppletive pairs that – at least from a synchronic point of view – are not morphologically transparent. I will also set aside (im)perfectiva tantum, which do not appear in aspectual pairs (arguably due to the lexical semantics of the predicates involved) (see, e.g., Isačenko 1962), as well as biaspectual verbs, for which the aspectual semantics is determined by context (see, e.g., Janda 2007).

   We can already see from these few examples that there is no uniform morphology for (I)PFVs in Russian: IPFVs can appear without any aspectual affixes, such as those in (1) (SIMPLE IPFVs) or they can appear with a suffix and often also a prefix, such as those in (2) (SIs); PFVs can contain a prefix, such as those in (1), (2a), and (2b), or they can lack aspectual affixes altogether, such as the one in (2c). Nevertheless, native speakers clearly have an intuition what it means for a given verb form to be IPFV or PFV, and there are also diagnostics for (I)PFV forms. For example, only IPFV verb forms can derive a periphrastic future tense form (the future auxiliary in combination with the IPFV infinitive) (3a); phase verbs like *begin, start, continue, stop, finish* only combine with IPFV infinitives (3b).

(3) a. Ja budu   {*pročitat' / čitat'}   knigu.
        I will.1SG read.PFV read.IPFV book.ACC
        'I will read a/the book.'

   b. Ja načinaju   {*pročitat'   / čitat'}   knigu.
        I start.IPFV.PRS.1SG read.PFV.INF read.IPFV.INF book.ACC
        'I am starting to read a/the book.'

The diagnostics are illustrated in (3) only for aspectual pairs with simple IPFVs and prefixed PFVs, but what is said here extends to other aspectual pairs as well (see, e.g., discussion in Isačenko 1962, Borik 2002).

2.2 Canonical and non-canonical readings of the Russian IPFV

There are two "canonical" readings (or two groups of readings) that Russian IPFV forms give rise to; these readings are canonical because such readings are commonly attested for IPFV forms cross-linguistically (see e.g. Deo 2009). The first

canonical IPFV reading is a process/durativity reading, which for example is the reading expressed by the English Progressive, an instance of IPFV. This reading is illustrated for Russian in the main clause of (4).

(4) Kogda ja vošla,   moj   brat   čital   knigu.
     when I in-went.PFV my.NOM brother.NOM read.IPFV book.ACC
     'When I came in, my brother was reading a book.'

The second canonical reading is that of iterativity/habituality, illustrated in (5).

(5) Ona každyj den' otkryvaet okno.
     she every day opens.SI window.ACC
     'She opens the window every day.'

This is not a reading that the English Progressive expresses primarily but it is a reading that IPFV forms in some other languages with grammatical aspect can give rise to. In Russian, whenever an event happened more than once (or potentially more than once), i.e. whenever the reference does not involve a single event, the IPFV has to be used.[1]

There are also non-canonical IPFV readings in Russian, i.e. readings that IPFV forms give rise to that are not common IPFV readings cross-linguistically, and outside of Slavic they might not even be attested. One family of such readings falls under the label GENERAL-FACTUAL (*obščefaktičeskoe*, after Maslov 1959), where IPFV forms can appear in contexts with typical PFV meanings, namely when referring to bounded "completed" events.[2] The literature on Russian aspect distinguishes at least two subtypes of the general-factual IPFV, the existential type (Padučeva 1996, Grønn 2004) and what Grønn calls the presuppositional type ("actional" in Padučeva 1996).

The EXISTENTIAL IPFV is illustrated in (6) (corpus example from Grønn 2004).

(6) Ne bylo   somnenij,   čto ja prežde vstrečal ee.
     not was.3SG.N doubt.GEN.PL that I before met.SI her
     'There was no doubt that I had met her before.'

---

[1]A notable exception to this rule is the so-called vivid-exemplifying use of a PFV present tense form in habitual contexts that are clearly marked as such (see Zaliznjak & Šmelev 2000). I will set such cases aside.

[2]However, the traditional literature also discerns subtypes of the general-factual with intuitively non-completed events; I will come back to this in §3.3.

Here the speaker asserts that he had a meeting with a female person in the past, and meetings in the past intuitively involve completed events that actually happened (at some time in the past). Nevertheless, we find an IPFV form here to describe such a meeting. More generally, the existential IPFV can be paraphrased as 'There has been/is etc. (at least) one event of this type.' (following the idea that existential IPFVs involve event types or kinds; see Mehlig 2001, 2013, Mueller-Reichau & Gehrke 2015). So in this case the paraphrase would be 'There was at least one event of the type "meet her".'

In this paper, I will not discuss the existential IPFV in detail, but I assume that the reason why an IPFV form is used in existential contexts has to do with the fact that the event is not necessarily a single event and that we are dealing with potential iterativity (labeled *kratnost'* '(lit.) multiple-ness' in Padučeva 1996). As stated in the beginning of this section, iterativity is one of the canonical readings of the Russian IPFV, so an account of the existential IPFV can build on an account for why the IPFV appears in iterative contexts (e.g. in terms of unbounded event plurality, as in Ferreira 2005, Altshuler 2014). This also means that a semantic account of the PFV in Russian somehow has to build in a restriction to single events, rather than just the external perspective on an event.

The PRESUPPOSITIONAL IPFV is illustrated in (7) (example from Glovinskaja 1982).

(7) Zimnij   Dvorec   stroil   Rastrelli.
    winter.ADJ.ACC palace.ACC built.IPFV Rastrelli.NOM
    'It was Rastrelli who built the Winter Palace.'

The presuppositional IPFV (at least with telic predicates) is probably the most noteworthy mismatch between event completion and aspect usage in Russian. In our example at hand we are dealing with a single event that happened in the past, namely the building of the Winter Palace in Saint Petersburg (which hosts the Hermitage). It is a known fact that this event took place only once and that it was completed, because we can see the result in front of us. It is also known when this event happened. Nevertheless, an IPFV verb form is used to describe this event.

The presuppositional IPFV is used when it is already clear from the context that the event in question exists (this is why Grønn labelled it presuppositional), and the sentence in which the IPFV form appears provides further information about this event. A suitable paraphrase is therefore 'The (already mentioned or contextually retrievable) event was/is etc. such and such.' In our example, this means that context presupposes the existence of the event 'build Winter Palace',

and the new information is that the architect of the building was Rastrelli. This use of the IPFV often goes hand in hand with a particular information structure, which is also evident in our example (and in the English translation I provided, a cleft construction): What is presupposed or backgrounded appears sentence-initially (the building of the Winter Palace) and the new information in focus is Rastrelli, in sentence-final position, resulting in a non-canonical OVS order.

In the following, I will outline the conditions under which this use of the IPFV arises, building on Grønn (2004) (who, in turn, heavily builds on empirical generalisations in the Russian literature, e.g. Glovinskaja 1982, Padučeva 1996).

2.3 Presuppositional IPFVs: Grønn (2004)

Let us look at another example from Grønn (2004) to discuss empirical generalisations about presuppositional IPFVs, namely the chess example in (8).

(8) Sdelav   ėtot   xod   [...], ja [predložil   nič'ju]$_{antecedent}$. [...]
    made.PFV.AP this.ACC move.ACC   I offered.PFV draw.ACC
    Navernjaka, černye   deržatsja   [...], no mne   ne
    probably   blacks.NOM hold-back.IPFV   but I.DAT not
    xotelos'   načinat' sčetnuju   igru,   [poėtomu]$_F$ ja
    wanted.IPFV.REFL begin.IPFV calculating.ACC game.ACC therefore   I
    i   [predlagal nič'ju]$_{anaphora}$.
    and/also offered.SI draw.ACC
    'Having played this move, I offered a draw. Black can probably hold on,
    but I didn't want to get involved in heavy calculations, and for this
    reason, I offered a draw.'   (after Grønn 2004: 207; my glosses)

In this example, the first sentence introduces a new event in the PFV (*predložil nič'ju* 'offered a draw'). The following discourse elaborates on the reason for offering a draw, and the last part of it states that for this reason (*poėtomu*) the draw was offered. This second mentioning of the event (offering a draw) is now described with an IPFV verb form (*predlagal*, the aspectual partner of *predložil*), and this is an instance of the presuppositional IPFV. The verb in this case is deaccentuated (see also Padučeva 1996), focus (indicated by the subscript F) is on some other constituent, in this case on *poėtomu* 'for this reason'. Grønn argues that the deaccentuation of the verb leads to the event given by the verb being backgrounded and to its prior instantiation being presupposed.

Following Geurts & van der Sandt (1997), Grønn (2004) treats presuppositions as anaphora that are either directly bound in the discourse, as in (8) (the an-

tecedent for the IPFV *predlagal* is the PFV *predložil* in the first sentence of the example), or contextually derivable, as in (9).

(9) Dlja bol'šinstva znakomyx   vaš   [ot"ezd]$_{(pseudo-)antecedent}$
     for majority   acquaintants.GEN your.NOM departure.NOM
     stal   polnoj   neožidannost'ju   ... Vy
     became.PFV full.INSTR unexpectedness.INSTR you.NOM
     [uezžali]$_{anaphora}$ v Ameriku   [ot   čego-to, k čemu-to ili že
      away-drove.SI in America.ACC from what-TO to what-TO or PRT
     prosto voznamerilis' spokojno provesti   tam buduščuju
     simply decided.PFV   calmly   spend.INF.PFV there future.ADJ.ACC
     starost']$_F$?
     old-age.ACC
     'For most of your friends your departure to America came as a total
     surprise ... Did you leave for America for a particular reason or with a
     certain goal, or did you simply decide to spend your retirement calmly
     over there?'   (after Grønn 2004: 207f.; my glosses)

In this example we do not have a direct finite PFV antecedent to the presuppositional IPFV *uezžali* 'departed'; instead, a nominalisation based on a related verb, *ot"ezd* 'departure', serves as what Grønn labels pseudo-antecedent in the previous discourse. Again, the presuppositional IPFV verb form is deaccentuated and focus lies on the questions for the reasons for the departure.

To illustrate Grønn's account of the presuppositional IPFV let us look at his analysis of (10) (attributed to Forsyth 1970).

(10) V ėtoj porternoj ja [...] napisal   pervoe ljubovnoe pis'mo.
      in this tavern I   wrote.PFV first.ACC love-.ACC letter.ACC
      Pisal   [karandašom]$_F$.
      wrote.IPFV pencil.INSTR
      'In this tavern I wrote my first love letter. I wrote it with pencil.'

Grønn's DRT analysis of the VP of the second sentence of (10) is given in (11).[3]

(11) $\lambda e[x \mid \text{INSTRUMENT}(e, x), \text{PENCIL}(x)]_{[\mid \text{WRITE}(e)]}$

---

[3]DRT is the abbreviation of Discourse Representation Theory (see Kamp & Reyle 1993). Grønn employs a linear notation for Discourse Representation Structures (DRSs), where discourse referents are written on the left-hand side, before | (in a traditional DRS they appear at the top of the DRS), and the conditions on these discourse referents are listed to the right of |, separated by commas (which in a different notation can be translated as conjunctions).

Berit Gehrke

Grønn argues that the VP is divided into background and focus (following Krifka 2001), where backgrounded material is turned into a presupposition, following the Background/Presupposition Rule in Geurts & van der Sandt (1997). In Grønn's DRT analysis, backgrounded material is subscripted in the DRS, so in this example the writing event itself is backgrounded and presupposed in the discourse. This VP gets further embedded under Aspect and Tense, which is where my proposal will differ from Grønn's proposal, but up to this point I will follow his account of presuppositional IPFVs.

What is the semantics of the (I)PFV then? In the following, I will discuss various proposals in light of how they deal with existential and presuppositional IPFVs.

2.4 The semantics of Russian aspect: Some proposals

As outlined in the introduction, common approaches to the semantics of Russian aspect treat it as a relation between reference/assertion time and some other temporal interval (e.g. Klein 1995, Schoorlemmer 1995, Borik 2002, Paslawska & von Stechow 2003, Grønn 2004, 2015, Ramchand 2008, Tatevosov 2011, 2015) or as an event predicate modifier, in the opposition of total vs. partial events (e.g. Filip 1999, Altshuler 2014). The most common approach is to provide a positive definition only of the PFV and to treat the IPFV as (semantically) "unmarked" (−PFV or ±PFV), but some approaches also provide a positive definition of the IPFV. One of the main motivations for treating the IPFV as unmarked is precisely the general-factual IPFV. Most agree that PFV forms always express a uniform PFV meaning, for example that the event time is included in the reference time. There is more disagreement with respect to the question whether IPFV forms come with a uniform IPFV meaning. Setting aside explicitly modal definitions of the IPFV, such as Arregui et al. (2014), who argue that different IPFV readings come about due to different modal bases, let me outline four representative types of proposals.

Borik (2002) argues that the meaning of the IPFV is the negation of the positive definition of the PFV, as illustrated in (12).

(12)   a. $S \cap R = \varnothing \,\&\, E \subseteq R$ PFV
       b. $\neg(S \cap R = \varnothing \,\&\, E \subseteq R) = S \cap R \neq \varnothing \lor E \nsubseteq R$ IPFV

The PFV is defined as a conjunction of two conditions that have to be met (12a): The speech time S must not overlap with the reference time R, and the event time E is included in the reference time. Negating this conjunction leads to a disjunction for the IPFV in (12b): Speech time and reference time overlap, or the

event time is not included in the reference time. This disjunction captures what Borik labels the "progressive" reading of the IPFV (when the event time is not included in the reference time) as well as what she labels the "present perfect" reading, which is essentially the existential IPFV reading outlined in the previous section (speech time and reference time overlap). Borik explicitly sets habitual and iterative readings of the IPFV aside, but we could assume that this can be added along the lines of other proposals in the literature. What is problematic for her account, though, is that it leaves the presuppositional IPFV unaccounted for.

Grønn (2004) and Altshuler (2014) provide weak positive definitions for the IPFV that get pragmatically/contextually strengthened in different directions. Building on Klein (1995), Grønn (2004) argues that the IPFV involves the event time overlapping with the reference time ($e \circ t$). This weak semantics gets pragmatically strengthened to a "proper" IPFV (the reference time is included in the event time), or to an actual PFV semantics (the event time is included in the reference time), which, he argues, happens in the case of factual IPFVs. Grønn takes into account the role of information structure to characterise the contexts in which strengthening happens in one or the other direction.

Altshuler (2014) provides the definition of the IPFV in (13), according to which the IPFV denotes an event e' that is a stage of an event e that exists in world w (where the current world of e' is $w*$) and that has the property P.[4]

(13) IPFV $\rightsquigarrow \lambda P \lambda e' \exists e \exists w [\text{STAGE}(e', e, w^*, w, P)]$

A stage of an event is defined as in (14), building on Landman's (1992) definition of the English Progressive.[5]

(14) $[\![\text{STAGE}(e', e, w^*, w, P)]\!]^{M,g} = 1$ iff (a)–(d) hold:

   a. the history of $g(w)$ is the same as the history of $g(w^*)$ up to and including $\tau(g(e'))$

   b. $g(w)$ is a reasonable option for $g(e')$ in $g(w*)$

   c. $[\![P]\!]^{M,g}(e, w) = 1$

---

[4]I render Altshuler's (2014) original formalisations, which use indirect translation. Otherwise, I use direct translations in this paper, and where not directly relevant I omit worlds and assignment functions.

[5]Note that with respect to the condition in (14d), Altshuler (2014) deviates from Landman (1992) and defines the English Progressive as a proper part relation, as he views this to be the crucial difference between Russian (part-of-relation) and English (proper-part relation). Landman, on the other hand, employed the weaker part-relation for the Progressive.

Berit Gehrke

      d. $g(e') \subseteq g(e)$

This is essentially an account of IPFV events as denoting partial events, and to capture what it means for an event to be a partial event (and notably also to capture the imperfective paradox), the definitions of stages and histories of events in (14a)–(14c) are needed. For our purposes, however, the essential part of the definition is given in (14d), according to which the event description in question is part of or equals the whole event. Altshuler argues that this can get pragmatically strengthened to a proper part meaning for the ongoing IPFV ($g(e') \subset g(e)$), or it can get strengthened to $g(e') = g(e)$, which essentially says that the partial event is identical to the whole event. In particular this last type of strengthening gives rise to the presuppositional IPFV reading. Altshuler does not address existential IPFVs (but see Altshuler 2012), but again this use arguably follows from a full account of habituality and iterativity. He argues that the use of IPFV for habitual event descriptions is captured by assuming a theory of plural events, following Ferreira (2005).

    Finally, Grønn (2015) departs from his earlier work and proposes that IPFV forms can express both IPFV (the reference time is included in the event time) and PFV semantics (the event time is included in the reference time), as in (15).

(15)   a. $[\![\text{PFV}]\!] = \lambda t \lambda e [e \subseteq t]$

          b. $[\![\text{IPFV}_{\text{ongoing}}]\!] = \lambda t \lambda e [t \subseteq e]$

          c. $[\![\text{IPFV}_{\text{factual}}]\!] = \lambda t \lambda e [e \subseteq t]$ "Fake" IPFV

Grønn calls the IPFV that has the same semantics as the PFV in (15c) a "fake" IPFV. The existence of IPFV$_{\text{factual}}$ alongside the PFV, he argues, leads to an aspectual competition. In the default case the PFV appears but in certain contexts, he argues, the IPFV$_{\text{factual}}$ wins the competition. This gives rise to the presuppositional IPFV in cases where narrative progression is to be avoided (under the assumption that the PFV always leads to narrative progression). The existential IPFV appears when the reference time is too large for the perfective semantics to be informative.

    Grønn's (2015) account essentially gives up on the idea that the Russian IPFV can have a uniform semantics. Altshuler's (2014) account provides a weak semantics for the IPFV. Both delegate the role of distinguishing between different IPFV readings to pragmatics and to the context. In this paper, I will equally take into account the role of context, but I will explore how far we can take a strong positive definition of the IPFV while still accounting for the occurrence of the presuppositional IPFV. In particular, I will argue that we can stick to a "proper" IPFV semantics, as opposed to a weak semantics or even a PFV semantics, if we take

the discourse and information structural cues into account. First, however, I will show that taking the intuitive notion of event completion as a crucial indicator for the right formal account of the semantics of aspect in Russian is misleading.

3 The focus on event completion is misleading

As discussed in the previous section, the fact that intuitively completed events can be described by IPFV forms has led to semantic accounts of the IPFV that give it a rather weak semantics (Grønn 2004, Altshuler 2014) or even argue that it can express both PFV and IPFV meanings (Grønn 2015). In this section, I will show that event non-completion is indeed neither a necessary nor a sufficient condition for an IPFV form to arise, just as we would expect from an account like Grønn's, which takes the intuitive notion of event completion as its starting point. We have already discussed factual IPFVs in the previous section, and further contexts to be addressed here involve chains of foregrounded events in habitual contexts and in the historical present, as well as the "annulled result" reading, which is sometimes considered a subtype of the factual IPFV. However, I will also show that event completion (as an intuitive notion) is neither a necessary nor a sufficient condition for a PFV form to arise. This is the case with PFV forms with the prefixes *po-* and *pro-*, as well as with the last event in a unique chain of foregrounded events. If event completion is taken as a key notion or intuition behind the definition of the PFV these examples are problematic. Instead, I will argue that the intuitive notion of event completion is not useful, at least not at the sentence level, since at this level we are interested in the particular description of events and make assertions that hold during particular reference time intervals, without making any claims about the actual events being completed or not. If we compare this with the nominal domain, we can also have complete entities, e.g. chairs and tables, but we can also choose to describe only parts of these in a particular sentence. The intuitive notion of event completion can still be relevant at the discourse level, however, and this is precisely what I will argue for in this paper. A main conclusion from this section will be that the discourse structure plays a crucial role for the choice of aspect in Russian (see also Altshuler 2012).

I will first discuss the use of IPFVs with completed events, then move on to the use of PFVs with non-completed events. At the end of the section, I will point out that general-factual readings also arise in the absence of intuitively completed events, which shows that giving factual IPFVs a PFV semantics will not work in these cases. What all these examples aim to show is that in contexts in which the IPFV occurs despite the intuitition that the event is completed, other than the

Berit Gehrke

factual IPFV, there is an explanation for the use of the IPFV that still falls within a "proper" IPFV semantics. It is only for factual IPFVs that authors like Grønn (2015) depart from such a semantics. This conclusion will serve as a point of departure for §4, in which I will argue that also these can be accounted for with a "proper" IPFV semantics.

3.1 IPFV with completed events

Let us take a look at (16) (discussed in Gehrke 2002, 2022).

(16) Ona   prixodila ko mne každyj den', a ždat'   ee   ja
      she.NOM to-went.SI to me every day and wait.INF.SI her.GEN I
      načinal s   utra.   [...] Za   desjat' minut ja sadilsja   k
      began.SI from morning.GEN   within ten   minutes I down-sat.SI to
      okoncu i   načinal prislušivat'sja, ne stuknet   li vetxaja
      window and began.SI listen.INF.SI   not clatters.PRES.PFV PRT old.NOM
      kalitka.
      gate.NOM
      'She came to me every day, and I started waiting for her from morning
      onwards. Within ten minutes [of her arrival] I sat next to the window
      and started listening whether the gate clatters.'

      (from Bulgakov, *Master i Margarita*)

The whole passage in (16) is explicitly marked as habitual by *každyj den'* 'every day' in the first sentence. There are four foregrounded events (*prixodila* 'arrived', *načinal ždat'* 'started to wait', *sadilsja* 'sat down', *načinal prislušivat'sja* 'started to listen'), out of which at least two (the first and the third) are intuitively completed, before the other two start. Nevertheless, these verb forms are IPFV (SIS) and the PFV would even be infelicitous in this context.[6] However, these IPFVs are generally not treated as cases of "fake" IPFV because the common explanation for the occurrence of the IPFV here is that habituality requires IPFV forms. I do not want to dispute this explanation, I just want to point out that event completion does not play a crucial role here for the choice of aspectual form.

Similarly, event completion does not seem to play a role in passages in the historical present. The historical present is a stylistic device in narratives, and in

---

[6]Note that other Slavic languages might be different in this respect. For example, in a Czech translation of (16), the third form is translated with a PFV verb (habituality in this language does not require IPFV), and this might indicate that event completion does play a bigger role here. For further discussion of differences in aspect usage between Russian and Czech see Gehrke (2002, 2022); for a description of cross-Slavic differences in general, see Dickey (2000).

these contexts Russian cannot use PFV forms (with the caveat mentioned in fn. 1). One such example is given in (17).

(17) […] les   končilsja,   neskol'ko kazakov vyezžajut   iz nego
         forest.NOM end.PFV.PST some   cossacks out-ride.SI.PRS out it
      na poljanu, i   vot, vyskakivaet   prjamo k nim moj
      on field   and there out-jump.SI.PRS directly to him my.NOM
      Karagez;   vse   kinulis'   za nim s   krikom […]
      Karagez.NOM all.NOM.PL rush.PFV.PST after him with shout
      'The forest ended, a few cossacks are riding out of it into the field, and
      there my Karagez jumps out directly towards them. They all rushed
      after him with a shout.'
              (from Lermontov, *Geroj našego vremeni*; discussed in Galton 1976: 25)

In this example there is again a chain of completed events, in particular the riding out of the forest (*vyezžajut*) and the jumping out (*vyskakivaet*), as a reaction to the first event, but these are nevertheless described with IPFV forms. Again, nobody calls these forms "fake" IPFVs, instead an alternative explanation is provided for why the historical present is incompatible with a PFV semantics (e.g. that a true present tense semantics is incompatible with the event time being part of the reference time).[7]

Finally, let us look at the example in (18) (after Smith 1991/1997: 311), which illustrates the use of the IPFV where the result is "annulled".

(18) K vam kto-to   prixodil.
      to you someone came.SI
      'Someone came to you.' (The person is not there anymore.)

In this example there is an intuitively completed event, and the IPFV is used to signal that the result state of this event (someone being there) does not hold anymore at the time of utterance. While Grønn (2004, 2015) subsumes cases like these under the notion of factual IPFVs and therefore would also treat them as "fake" IPFVs,[8] it is again clear that the role that these IPFVs play in discourse is crucial and we might want to look at an alternative explanation for the use of IPFV in such contexts in Russian.

---

[7]See Anand & Toosarvandani (2019) for a recent account of the historical present, which is incompatible also with the Progressive in English even in contexts where an ongoing event is described.

[8]Treatments of such cases as a type of general-factual IPFV can also be found in the Slavistic traditional literature; e.g. Padučeva (1996) calls this meaning *dvunapravlennoe obščefaktičeskoe* 'bi-directed general-factual', especially with motion verbs, as in (18).

Berit Gehrke

3.2 PFV with non-completed events

Let me then move on to PFV forms that can be used to describe non-completed events. It is well-known that in chains of foregrounded single events Russian requires PFV verb forms for reference time movement (in the sense of Kamp & Reyle 1993) (see also Borik 2002). This is also true for the last event in the chain, even if this event is not necessarily completed, as illustrated in (19).

(19) No v tot že mig   vspomnil   svoj   dom   i
      but in this PRT moment remembered.PFV his.REFL.ACC house.ACC and
      gor'ko {zaplakal   / *plakal}.
      bitterly ZA-cried.PFV   cried.IPFV
      'But at that moment he remembered his home and wept bitterly.'
                              (grammatical version from http://skazbook.ru/vodyanoi)

In this example the crying starts right after the remembering, but the crying itself does not necessarily have to be completed. In all likelihood we are just witnessing the beginning of the crying here. While some authors try to reason that the actual event described is precisely the onset and not the crying itself and that this warrants the use of the PFV (see, for instance, Ramchand 2008), descriptions and intuitions about such ingressive events suggest that the event in focus is the crying itself, including its process, not so much its onset, and that intuitively this event is not or at least does not have to be completed. Nevertheless the PFV is and has to be used. Furthermore, the example in (20) (discussed in Dickey 2000: 224 and attributed to Švedova & Trofimova 1983) shows that several such PFV verbs with the ingressive prefix *za-* in a row can be intepreted as "actions beginning simultaneously".

(20) Fljagin   vyšel: Čto   tut načalos'! Zagudeli,
      Fljagin.NOM out-went.PFV what.NOM then began.PFV ZA-hooted.PFV.PL
      zavorčali,   zakričali.
      ZA-grumbled.PFV.PL ZA-shouted.PFV.PL
      'Fljagin went out. And what began then! They started hooting,
      grumbling and shouting.'

What all these examples show is that event (non-)completion is not (necessarily) decisive for the choice of (I)PFV in a given sentence and should therefore not play the central role in formal semantic accounts of (I)PFV, at least not at the sentence level. Instead we need to pay closer attention to the discourse structure and to the role that (I)PFV forms play in discourse.

3.3 General-factual IPFV without completed events

Finally, merely treating factual IPFVs as "fake" IPFVs with a PFV semantics is missing an important insight from the Russian traditional linguistic literature (e.g. Glovinskaja 1981, Padučeva 1996). In particular, this literature discusses different subtypes of factual IPFVs, including some that appear with intuitively "incomplete" events. For example, Padučeva (1996) differentiates between resultative factual uses (the cases of existential IPFVs we have discussed so far), bi-directed factual uses (of the type in (18)), as well as non-resultative (*nerezul'tativnoe*) and atelic (*nepredel'noe*) factual IPFVs.[9] The latter two are illustrated in (21).

(21) a. Ja ugovarival   ee vernut'sja.
          I convinced.SI her return.INF.PFV
          'I convinced (tried to convince) her to return.'   (Padučeva 1996: 22)

      b. Ja vas   ljubil.
          I you.ACC loved.IPFV
          'I loved you.'   (Padučeva 1996: 32)

In the non-resultative factual IPFV in (21a) it remains open whether the speaker succeeded in convincing the person referred to by 'her', which could be made explicit by adding 'tried to' to the translation. The atelic factual IPFV in (21b), in turn, is the famous first line of a poem by Puškin, which continues with *ljubov' ešče, byt' možet, v duše moej ugasla ne sovsem* 'it is possible that in my soul this love is not yet completely extinguished', and this continuation makes explicit the effect of the atelic factual IPFV: it remains open whether the state described still holds at the moment of utterance. Both types share with the "resultative" factual IPFV (which for Padučeva involves existential IPFVs) that the time in the past at which these events or states held is not specific and that the relation to the current time of utterance is unclear; the first example furthermore involves potential iterativity.

These examples are usually ignored in the formal literature, because the more extraordinary situation seems to be where a (presumably) single "completed" event is referred to with an IPFV form. However, they still constitute a different IPFV "reading" than process or habituality, and we would want to know more about these readings rather than just treating one subset of factual IPFVs as "fake", thereby ignoring these other cases that share important similarities. Calling factual IPFVs "fake" IPFVs and giving them the same semantics as PFV is missing the point.

---

[9]Recall that she treats presuppositional IPFVs as distinct from other factual IPFVs, under the label *akcional'noe* 'actional'.

Berit Gehrke

How can we account for the semantics of factual IPFVs then? While I will leave existential IPFVs aside and assume that an account for the use of IPFV in habitual and iterative contexts and the requirement of a single event for the PFV will play a role here, the following section will provide an explicit account of presuppositional IPFVs that employs a standard IPFV semantics and takes into account information structural cues and the discourse. Event completion will be shown not to play a role at the sentence level, but at the discourse level the intuition of event completion will still be captured.

4 A discourse semantic account of presuppositional IPFVs

As the previous section showed, Russian aspectual forms play a crucial role in discourse (see also Altshuler 2012), which can easily be overlooked if one simply stays at the sentential level. Following Grønn (2004), I assume that presuppositional IPFVs are anaphorically linked to a previously introduced event in the ideal case, or that the presupposition that the event is already given in the context has to be accommodated. Combining this with discourse semantic accounts, such as Kamp & Reyle (1993) and Lascarides & Asher (1993), this means that a presuppositional IPFV introduces an eventive discourse referent that is identified with another eventive discourse referent already introduced in previous discourse, in parallel to the treatment of individual pronouns and definite descriptions in the nominal domain in, e.g., DRT. In terms of discourse relations that hold between events, in the case of presuppositional IPFVs we are intuitively dealing with ELABORATION. In Lascarides & Asher's (1993) system of rhetorical relations between events described in two clauses α and β, where the former precedes the latter, Elaboration holds when β's event is part of α's. So at this point Altshuler's (2014) partitive semantics is more promising than Grønn's (2004) weak IPFV semantics as mere temporal overlap or even Grønn's (2015) PFV semantics. Altshuler himself suggests in his discussion of the example in (10) ((97) in Altshuler 2014: 769) that Elaboration is the discourse relation involved and that pragmatic strengthening of the part relation to an equal-relation leads to both events being identical. In this paper, I propose to go a step further and work with a proper part semantics from the start, thereby abandoning the need for pragmatic strengthening. Instead, I will argue that event identity follows from the information structural cues, along the lines of what was proposed in Grønn (2004).

4.1 First attempt

As an empirical point of departure for illustrating how a proper part semantics coupled with standard discourse semantic assumptions will account for the presuppositional IPFV I will use data from a corpus study with Olga Borik (Borik & Gehrke 2018). In this study we show that IPFV past passive participles (PPPs) in Russian, which are often claimed not to exist (at least from a synchronic point of view), are attested in corpora, and that they can be given a compositional semantics and are not just frozen forms. The corpus study results indicate important restrictions though: First, there are no secondary IPFV PPPs, and second – more importantly for our purposes – there are no IPFV PPPs with a process meaning. Our hypothesis was that IPFV PPPs are always factual, and we particularly focussed on presuppositional IPFV PPPs, like the one in (22) (from Borik & Gehrke 2018).

(22) Čto kasaetjsa platy   deneg,   to plačeny byli naličnymi
      what concerns payment money.GEN so paid.IPFV were in-cash
      šest'   tysjač   rublej […]
      six.NOM thousand Rubles
      'What concerns the payment: 6000 Rubles were paid in cash.'

In this example, the payment event is first introduced by a nominalisation (*plata* 'payment'), and the IPFV PPP in the main clause links back to this already introduced event. The marked word order and the most natural way to read this example also indicate a marked information structure: the paying event appears in the beginning of the sentence and is backgrounded, focus lies on the sentence-final subject and (possibly also) on the modifier ('6000 Rubles (in cash)').

Let us work with a proper part semantics for the IPFV and build on independently motivated and received assumptions about discourse semantics. A first attempt, employing a linear notation of DRT (recall fn. 3) but leaving the division into background/presupposed and focused material implicit, is in (23).

(23) $[e_1, e_2, t, n, x \mid \text{PAYMENT}(e_1), \text{PAY}(e_2), e_2 = e_1,$
      $\text{THEME}(e_2, x), 6.000\text{R}(x), \text{IN-CASH}(e_2), t \subset \tau(e_2), t < n]$

The DRS keeps track of various discourse referents and conditions on these, as follows. *Plata* 'payment' is an event nominal that introduces the event discourse referent e_1. Since it is a non-finite (i.e. tenseless) verb form, I assume that there is no reference time and no temporal trace related to it; I will get back to this.[10] The

---

[10]The temporal trace of an event is represented as $\tau(e)$, following Krifka (1998).

Berit Gehrke

event described by the ɪᴘꜰᴠ PPP is represented by e_2, and this event description is treated like a definite description that is anaphorically linked to e_1 ($e_2 = e_1$), along the lines of the DRT treatment of definite descriptions in the nominal domain.[11] The new information in focus is about e_2, and since e_2 is identical to e_1 it is also about e_1: the theme of e_2 is '6.000 Rubles' and this was payed 'in cash' (treated as an event modifier). Following Kamp & Reyle (1993), the semantic contribution of past tense is that it introduces a reference time interval t that is before now ($t < n$). The crucial condition now is that we analyse ɪᴘꜰᴠ with a proper part semantics, which I treat as a temporal relation: the reference time interval t is properly included in the run time of e_2 ($t \subset \tau(e_2)$).

If we still wanted to capture the intuition that the actual paying event was completed, at least in the overall discourse, this analysis does not succeed, because the antecedent (or pseudo-antecedent) for the factual ɪᴘꜰᴠ is not a finite verb form but a nominalisation. In the next section, I will make a second attempt, in order to see if we can remedy this potentially intuitive shortcoming.

4.2 Second attempt

If we wanted to directly capture the intuition that in the overall discourse the event referred to by the nominalisation is completed, we would have to reconstruct a ᴘꜰᴠ semantics for the nominalisation, along the lines of (24).

(24) $[e_1, e_2, t_1, t_2, n, x \mid \text{PAYMENT}(e_1), \text{PAY}(e_2), \text{THEME}(e_2, x),$
        $6.000\text{R}(x), \text{IN-CASH}(e_2), e_2 = e_1, \tau(e_1) \subset t_1, t_2 \subset \tau(e_2), t_2 < n]$

What is new now is that we add a new discourse referent t_1 to the DRS, which serves as a reference time for e_1 (the event discourse referent introduced by the nominalisation). We furthermore reconstruct a ᴘꜰᴠ semantics for this nominalisation, since this would represent our intuition that the event is completed: the run time of e_1 is properly included in the reference time t_1 ($\tau(e_1) \subset t_1$).

However, we now face new problems. Since nominalisations are non-finite, t_1 is not related to n; intuitively it is before n, but this would be a second reconstruction. Furthermore, without this reconstruction, we do not know how t_1 and t_2 are related (with it, it will work as in §4.3). More generally, we do not know whether we want to associate nominalisations with temporal traces to begin with – this might at most make sense for complex event nominals (in the

---

[11]I assume that due to the information structure involved a prior step involves Grønn's (2004) account for the VP domain, as outlined in §2; in this section I already take this step for granted and outline the following step in which information structural cues have already been resolved.

sense of Grimshaw 1990) but not necessarily for nominalisations in general. It is also not clear why we would associate nominalisations with a particular aspect semantics; intuitively we want a PFV semantics here because intuitively the event is completed. However, Russian nominalisations do not come in aspectual pairs, which could be taken as evidence for nominalisations lacking a functional projection associated with Aspect (AspP), as argued, for instance, by Schoorlemmer (1995). So why associate them with (I)PFV semantics to begin with?

I do not think our first two attempts at a formalisation should make us want to give up on the idea that we can have an IPFV semantics for factual IPFVs in a given sentence, while still capturing the overall intuition at the discourse level that the actual event was completed. I think it rather shows that in the cases where we have to accommodate a discourse referent, as in the case with nominalisations, if we follow Grønn's (2004) reasoning, we will also have to accommodate more information that is otherwise contributed by tense and aspect. A full-fledged theory of accommodation would have to address this, but I will not attempt to do this in this short contribution.[12] Instead, in the following I will explore what happens if in the discourse we do have a PFV antecedent that explicitly provides the antecedent for the factual IPFV.

4.3 The account: The zooming-in function of presuppositional IPFVs

In order to work with an example with a finite PFV antecedent for the presuppositional IPFV, I constructed an example that is not attested in the corpus, unlike (22), but which is still a fully acceptable discourse, namely (25).[13]

---

[12]Olav Mueller-Reichau (p.c.) suggests that the completedness intuition might be captured by assuming that presupposed entities are whole entities (unless there is evidence to the contrary), because they are listed as items on file cards.

[13]This is not to say that there are no such examples in the corpus, it is just that presuppositional IPFVs quite often require accommodation rather than true antecedents, so I wanted to address the general issue of how do deal with accommodation. An example from the corpus with a PFV antecent and an analysis that works just like (25b) is the following.

(i) a. I   tak napisano,   čto mnogie   rasplakalis'   – krovju
       and so written.N.SG.PFV that many.NOM started-crying.PFV blood.INSTR
       duši   pisano.
       soul.GEN written.N.SG.IPFV
       'It was written so that many started to cry, it was written with the blood of the soul.'

    b. $[e_1, e_2, t_1, t_2, n, x \mid \text{WRITE}(e_1), \tau(e_1) \subset t_1, t_1 < n, \text{WRITE}(e_2),$
       $\text{BLOOD-OF-SOUL}(x), \text{INSTRUMENT}(e_2, x), e_2 = e_1, t_2 \subset \tau(e_2), t_2 < n]$

Berit Gehrke

(25)   a.   Zaplatili.   Plačeny byli naličnymi šest'   tysjač   rublej.
            paid.3PL.PFV paid.IPFV were in-cash   six.NOM thousand Rubles
            'They paid. It was paid 6.000 Rubles in cash.'

      b. $[e_1, e_2, t_1, t_2, n, x \mid \text{PAY}(e_1), \tau(e_1) \subset t_1, t_1 < n, \text{PAY}(e_2), \text{THEME}(e_2, x),$
            $6.000\text{R}(x), \text{IN CASH}(e_2), e_2 = e_1, t_2 \subset \tau(e_2), t_2 < n]$

Under the analysis in (25b), there is a paying event e_1, introduced by the PFV verb form in the first sentence: its run time, $\tau(e_1)$, is properly included in the reference time t_1 (the semantics of PFV), which is before n(ow) (the semantics of past tense). The analysis for the second sentence does not differ from the second attempt: The presuppositional IPFV PPP introduces a second paying event e_2, which is anaphorically linked to e_1, i.e. $e_2 = e_1$. The new information about this event is that its theme is 6.000 Rubles and it was paid in cash. The IPFV semantics specifies that there is a second reference time, t_2, which is properly included in the run time of the event, $\tau(e_2)$, and past tense indicates that this reference time is before the time of utterance.

At this point, a proponent of the "fake" IPFV analysis might object and say that the IPFV semantics for e_2 in the second sentence still does not directly capture that the paying event was completed. This is indeed true, but only at the sentence level. However, it follows from the discourse structure as a whole: Event completion information is already given in the first sentence about e_1 (its run time falls within the first reference time t_1). Since e_2 equals e_1, the actual event of paying remains completed. Furthermore, the second reference time, t_2, is properly included in the run time of e_2, and therefore it is also properly included in the run time of e_1 (since e_2 is identical to e_1). By transitivity, t_2 must also be properly included in the first reference time, t_1. The effect of the presuppositional IPFV, then, is that it is used to zoom in on a narrower reference time within a bigger reference time; the link between the two reference times t_1 and t_2 is only indirect, via the events involved, but it can still be made. The assertion that the sentence with the presuppositional IPFV makes, then, is only for part of the bigger reference time and only for part of the actual event, and this is what is captured by the IPFV semantics. This is precisely what we would expect if the event description provided by the presuppositional IPFV merely elaborates on the first event.

There are at least two advantages of this proposal over Grønn's (2015) "fake" IPFV account. First, it can easily be extended to atelic and non-resultative subtypes of the presuppositional IPFV, which are well discussed in the descriptive literature (recall the discussion in §3.3). For Grønn such subtypes would not involve "fake" IPFVs (with a PFV semantics) and would thus not be analysed along the same lines, even though some of these (the presuppositional ones) share the

same information structural properties and anaphoric link to previously intro-
duced events (these events are just not completed, in this intuitive sense). Second,
we maintain a uniform semantics for IPFV verb forms.

The gist of the proposal treats presuppositional IPFVs as a special case of the
ongoing reading of IPFVs, since both involve the reference time being properly
included in the run time of the event. The ongoing reading is analysed as a proper-
part-relation also by Altshuler (2014), but under his account both readings (pre-
suppositional and ongoing) are arrived at only after pragmatically strengthening
the weaker partitive semantics he proposed for the IPFV, and the two readings
end up with a different strengthened semantics since for him the result of prag-
matic strengthening with presuppositional IPFVs is identity of the two events
(recall the discussion in §2.4). In contrast, the current proposal already starts out
with the stronger IPFV semantics, which is the same as under the ongoing rea-
ding; identity of the two events follows from the information structural cues that
build an anaphoric link to the previously introduced (or accommodated) event,
just like what we find with definites in the nominal domain. Thus, by taking the
information structural cues already identified by Grønn (2004) as a point of de-
parture to spell out a discourse semantic account that integrates independently
proposed assumptions about definites and anaphoric relations in discourse, event
identity is the result of the discourse structure and not of pragmatic strengthen-
ing of the IPFV semantics.

5 Conclusion

In this paper I argued that an analysis of factual IPFVs as "fake" IPFVs, assigning
them a PFV semantics, is misguided by the strong focus on event completion. I
claimed that taking the intuitive notion of "completed" events as a central in-
gredient of the semantic definition of the (I)PFV aspect at the sentential level
is misleading because there are numerous mismatches between (I)PFV forms and
(in)complete events in the actual world. Rather, since we are primarily concerned
with the way we describe a given event (with aspectual forms) in a given sentence
and such descriptions can also involve descriptions of parts of events, the intu-
ition of event completion could also be delegated to the level of the discourse. I
argued that by taking into account the discourse structure it is possible to pro-
vide a strong IPFV semantics for presuppositional IPFVs, which therefore turn out
to be "true" IPFVs: they elaborate on a part of a previously introduced event.

There are remaining issues for future research. For one, I have not addressed
other subtypes of the factual IPFV, such as the existential IPFV or the annulled

Berit Gehrke

result cases (if these are indeed subcases). However, I am confident that a full-fledged account of habituality and iterativity, coupled with the single event requirement for PFVs and possibly further discourse semantic considerations, will work for existential IPFVs; also annulled results point to a discourse function. A second area for further investigation derives from the fact that the proposed analysis crucially builds on there being a finite PFV antecedent. What do we do with non-finite antecedents (e.g. nominalisations) which – at least in Russian – do not come in a particular aspect? And finally, how do we handle accommodation, which is similar to bridging in the nominal domain (see discussion in Borik & Gehrke 2018)?

Abbreviations

1	first person	NOM	nominative case
3	third person	PFV	perfective
ACC	accusative case	PL	plural
AP	adverbial participle	PRS	present tense
DAT	dative case	PRT	particle
GEN	genitive case	PST	past tense
IPFV	imperfective	REFL	reflexive
INF	infinitive	TO	specific indefinite marker *-to*
INSTR	instrumental case	SG	singular
F	focus	SI	secondary imperfective
N	neuter	ZA	inchoative/ingressive prefix *za-*

Acknowledgements

References

Altshuler, Daniel. 2012. Aspectual meaning meets discourse coherence: A look at the Russian imperfective. *Journal of Semantics* 29.1. 39–108. DOI: 10.1093/jos/ffr008.

Altshuler, Daniel. 2014. A typology of partitive aspectual operators. *Natural Language & Linguistic Theory* 32.3. 732–775. DOI: 10.1007/s11049-014-9232-1.

Anand, Pranav & Maziar Toosarvandani. 2019. No explanation for the historical present: Temporal sequencing and discourse. *Proceedings of Sinn und Bedeutung* 22(1). 73–90. DOI: 10.18148/sub/2018.v22i1.66.

Arregui, Ana, María Luisa Rivero & Andrés Salanova. 2014. Cross-linguistic variation in imperfectivity. *Natural Language & Linguistic Theory* 32.2. 307–362. DOI: 10.1007/s11049-013-9226-4.

Borik, Olga. 2002. *Aspect and reference time* (LOT Dissertation Series 64). Utrecht: LOT.

Borik, Olga & Berit Gehrke. 2018. Imperfective past passive participles in Russian. In Denisa Lenertová, Roland Meyer, Radek Šimík & Luka Szucsich (eds.), *Advances in formal Slavic linguistics 2016*, 53–76. Berlin: Language Science Press. DOI: 10.5281/zenodo.2546440.

Deo, Ashwini. 2009. Unifying the imperfective and the progressive: Partitions as quantificational domains. *Linguistics and Philosophy* 32.5. 475–521. DOI: 10.1007/s10988-010-9068-z.

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

Ferreira, Marcelo. 2005. *Event quantification and plurality.* Cambridge, MA: MIT. (Doctoral dissertation).

Filip, Hana. 1999. *Aspect, eventuality types and nominal reference.* New York, NY: Garland Publishing.

Forsyth, James. 1970. *A grammar of aspect: Usage and meaning in the Russian verb.* Cambridge: Cambridge University Press.

Galton, Herbert. 1976. *The main functions of the Slavic verbal aspect.* Skopje: Macedonian Academy of Sciences & Arts.

Gehrke, Berit. 2002. *Systemhafte Unterschiede im Aspektgebrauch zwischen dem Russischen und dem Tschechischen.* Berlin: Humboldt-Universität zu Berlin. (MA thesis).

Gehrke, Berit. 2022. Differences between Russian and Czech in the use of aspect in narrative discourse and factual contexts. *Languages* 7(2). 155. DOI: 10.3390/languages7020155.

Geurts, Bart & Rob van der Sandt. 1997. Presuppositions and backgrounds. In Paul Dekker, Martin Stokhof & Yde Venema (eds.), *Proceedings of the eleventh Amsterdam Colloquium, December 17–20, 1997*, 37–42. University of Amsterdam: Institute for Logic, Language & Computation, Department of Philosophy.

Glovinskaja, Marina Ja. 1981. Obščefaktičeskoe značenie nesoveršennogo vida (formy prošedšego vremeni). In Viktor P. Grigor'ev (ed.), *Problemy strukturnoj lingvistiki 1978*, 108–125. Moscow: Nauka.

Berit Gehrke

Glovinskaja, Marina Ja. 1982. *Semantičeskie tipy vidovyx protivopostavlenij russkogo glagola*. Moscow: Nauka.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

Grønn, Atle. 2004. *The semantics and pragmatics of the Russian factual imperfective*. Oslo: University of Oslo. (Doctoral dissertation).

Grønn, Atle. 2015. On (in)definite tense and aspect in Russian. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebowa (eds.), *Slavic grammar from a formal perspective: The 10th anniversary FDSL conference, Leipzig 2013*, 175–196. Frankfurt/M.: Peter Lang.

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

Janda, Laura. 2007. What makes Russian bi-aspectual verbs special? In Dagmar Divjak & Agata Kochanska (eds.), *Cognitive paths into the Slavic domain*, 83–110. Berlin/New York: Mouton de Gruyter. DOI: 10.1515/9783110198799.2.83.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic: Introduction to modeltheoretic semantics of natural language, formal logic and Discourse Representation Theory*. Dordrecht: Kluwer. DOI: 10.1007/978-94-017-1616-1.

Klein, Wolfgang. 1995. A time-relational analysis of Russian aspect. *Language* 71(4). 669–695. DOI: 10.2307/415740.

Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (ed.), *Events and grammar*, 197–235. Dordrecht: Kluwer. DOI: 10.1007/978-94-011-3969-4_9.

Krifka, Manfred. 2001. For a structured meaning account of questions and answers. In Caroline Féry & Wolfgang Sternefeld (eds.), *Audiatur vox sapientiae: A Festschrift for Arnim von Stechow*, 287–319. Berlin: Akademie-Verlag. DOI: 10.1515/9783050080116.

Landman, Fred. 1992. The progressive. *Natural Language Semantics* 1(1). 1–32. DOI: 10.1007/BF02342615.

Lascarides, Alex & Nicholas Asher. 1993. Temporal interpretation, discourse relations and commonsense entailment. *Linguistics and Philosophy* 16(5). 437–493. DOI: 10.1007/BF00986208.

Maslov, Jurij S. 1959. Glagol'nyj vid v sovremennom bolgarskom literaturnom jazyke. In Samuil B. Bernštejn (ed.), *Voprosy grammatiki bolgarskogo literaturnogo jazyka*, 157–312. Moscow: Nauka.

Mehlig, Hans Robert. 2001. Verbal aspect and the referential status of verbal predicates: On aspect usage in Russian who-questions. *Journal of Slavic Linguistics* 9(1). 99–125.

Mehlig, Hans Robert. 2013. Obščefaktičeskoe i edinično-faktičeskoe značenija nesoveršennogo vida v russkom jazyke. *Vestnik Moskovskogo Universiteta Serija 9, Filologija* 4. 19–47.

Mueller-Reichau, Olav & Berit Gehrke. 2015. Event kind formation within the VP: Comparing Russian factual imperfectives and German adjectival passives. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebowa (eds.), *Slavic grammar from a formal perspective: The 10th anniversary FDSL conference, Leipzig 2013*, 367–382. Frankfurt/M.: Peter Lang.

Padučeva, Elena V. 1996. *Semantičeskie issledovanija.* Moscow: Škola 'Jazyki russkoj kul'tury'.

Paslawska, Alla & Arnim von Stechow. 2003. Perfect readings in Russian. In Artemis Alexiadou, Monika Rathert & Arnim von Stechow (eds.), *Perfect explorations*, 307–362. Berlin: Mouton de Gruyter. DOI: 10.1515/9783110902358.307.

Ramchand, Gillian. 2008. Perfectivity as aspectual definiteness: Time and the event in Russian. *Lingua* 118(11). 1690–1715. DOI: 10.1016/j.lingua.2007.03.008.

Schoorlemmer, Maaike. 1995. *Participial passive and aspect in Russian.* Utrecht: Utrecht University. (Doctoral dissertation).

Smith, Carlota S. 1991/1997. *The parameter of aspect.* Dordrecht: Kluwer. DOI: 10.1007/978-94-011-5606-6.

Švedova, Lilija N. & Tamara G. Trofimova. 1983. *Posobie po upotrebleniju vidov glagola dlja raboty s filologami-rusistami.* Moscow: Russkij jazyk.

Tatevosov, Sergei. 2011. Severing perfectivity from the verb. *Scando-Slavica* 57. 216–244.

Tatevosov, Sergei. 2015. Severing imperfectivity from the verb. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebowa (eds.), *Slavic grammar from a formal perspective: The 10th anniversary FDSL conference, Leipzig 2013*, 465–494. Frankfurt/M.: Peter Lang.

Zaliznjak, Anna A. & Aleksej D. Šmelev. 2000. *Vvedenie v russkuju aspektologiju.* Moscow: Jazyki russkoj kul'tury.

Chapter 6

Inflectionless adjectives in Bulgarian as a case of nominal predication

Ekaterina Georgieva

Hungarian Research Centre for Linguistics

This paper deals with the so-called inflectionless adjectives in Bulgarian. Several new empirical observations are made regarding the syntactic distribution, the restrictions on definiteness, and the exclamatory flavour of the noun phrases in which these adjectives occur. The main proposal is that these lexical items are predicates of (nominal) small clauses and that the construction in question does not seem to be limited to these exceptional adjectives. It is argued that both the attributive type and the comparative type of nominal predication are attested in Bulgarian, on a par with English small clauses like *an idiot doctor* and *an idiot of a man*. I outline a syntactic account of these two types of nominal predication, according to which the two types correspond to different structures. I also propose that the semantic and syntactic properties of inflectionless adjectives are best accounted for if we assume that they combine with a null noun.

Keywords: inflectionless adjectives, nominal predication, small clause, null noun, definiteness, Bulgarian

In this paper I discuss a small class of nominal modifiers in Bulgarian, previously referred to as "inflectionless adjectives" (see Halpern 1995, Spencer & Luís 2012, Nicolova 2017, Adamson 2019, 2020, 2022; a.o.). Adjectives in Bulgarian inflect for gender and number, but a small group of loanword adjectives, some of which are borrowings from Turkish, do not. A non-exhaustive list is given in (1) (based on Nicolova 2017: 178 and Adamson 2019 with some additions).

(1)   *serbez* 'bold, insolent', *ursuz* 'crabby, mean', *erbap* 'capable, skillful, cocky', *sert* 'assertive, testy, strong, quick-tempered', *češit* 'weird, crank', *inat* 'stubborn, obstinate' (also used as a noun, with the meaning

'stubbornness, obstinacy'), *kofti* 'bad, shitty', *kurnaz* 'bold, cocky', *piškin* 'experienced, spirited', *mukajat* 'determined, proactive', *pišman* 'fake, feigned, sham'; and also *šik* 'chic', *ekstra* 'perfect', *seksi* 'sexy', *super* 'super', *pop* 'pop', *džaz* 'jazz' (not discussed here)

It should be emphasized that the items in (1) do not likely form a unified group. In this paper, I will focus on the Turkish borrowings (with the exception of *inat* 'stubborn' because of its clearly nouny use) and will show that certain semantic and syntactic properties of the noun phrases containing these items have been left unnoticed: first, the noun phrases with some of these items show restrictions on definiteness, and second, they also have a limited syntactic distribution. The inflectionless adjectives have been discussed in the literature mostly in connection with the placement of the definiteness marker, which is exceptional in comparison to what we find with inflecting adjectives (see Halpern 1995, Spencer & Luís 2012, Adamson 2019, 2020). Although the present paper does not aim to focus on the placement of the definiteness marker, the new empirical data will refine the claims made in the literature about the use of the definiteness marker with inflectionless adjectives, in particular with respect to the interspeaker variation discussed in the earlier studies.

In this paper, I propose that the inflectionless adjective and the noun form a predication structure comparable to well-known cases of nominal predication, e.g., English *an idiot of a doctor* (see Napoli 1989, Kayne 1994; Den Dikken & Lipták 1997; Hulk & Tellier 2000, Doetjes & Rooryck 2003, Casillas Martínez 2003; Den Dikken 2006; Villalba & Bartra-Kaufmann 2010; a.o.). These noun phrases have been referred to as THE QUALITATIVE BINOMINAL NOUN PHRASE (QBNP) in Den Dikken (2006) and Villalba & Bartra-Kaufmann (2010); as "qualitative construction" in Doetjes & Rooryck (2003), and as "*N1/A de N2* affective construction" in Casillas Martínez (2003). Throughout the paper I use the terms nominal predication, nominal small clause and QBNP interchangeably.

I will point out similarities between the noun phrases with inflectionless adjectives such as (2) and the corresponding nominal predication constructions in the languages discussed in the above-mentioned sources. Even more importantly, I will show that nominal predication in Bulgarian is not limited to the closed class of borrowed lexical items listed in (1), as (non-borrowed) nouns can also be used as the first part of the nominal predication, as shown in (3).[1]

---

[1]Possibly, the rather poorly-understood class of compounds with comparative semantics (e.g., *gaitan veždi* 'woollen.braid eyebrow.PL, well-shaped eyebrows') and the so-called 'appositive compounds' also belong here (see Bagasheva 2017). But since most of these examples are quite archaic, I leave them out from the present discussion and focus on more productive patterns like the one illustrated in (3).

(2) Eh, kakva   ursuz žena!
    PRT what.kind crabby woman
    'What a crabby woman!'

(3) a. […] e   leke čovek […]
        be.PRS.3SG stain person
        '[X] is a rotter of a man (lit. a stain person) […]'   [Google search]

     b. văj sega tuj leke čovek na Bolen
        PRT now this stain person of Bolen
        'wow, (and) now this rotter of a man of Bolen's!'   [Google search]

What is common between these examples is that semantically they express an (often negative) evaluation of the referent of the NP. I will also point out certain structural similarities between them, such as the restrictions on definiteness as well as their exclamative flavour. Thus, the phenomenon under consideration cannot possibly be explained with the exceptional properties of the inflectionless adjectives.

The paper is organised as follows: in §1, I first summarize the previous claims made in the literature regarding inflectionless adjectives. In §2, I present novel observations regarding the noun phrases containing inflectionless adjectives. Then in §3, I present an overview of the theoretical analyses of nominal small clauses, based on which I outline a possible analysis of the Bulgarian data in §4. In §5, I conclude and raise some further questions for future research.

1 Previous approaches to inflectionless adjectives

As already mentioned in the introduction, inflectionless adjectives have been discussed in the literature mostly in connection with the placement of the definiteness marker in Bulgarian. As far as empirical data are concerned, the baseline for the placement of the definiteness marker (DEF) is that it attaches to the noun (4a), but if the noun is preceded by adjectival modifiers, then it is placed on the (first) adjective (4b) (for more details see Halpern 1995, Franks 2001, Embick & Noyer 2001, Dost & Gribanova 2006, Harizanov 2014, 2018, Harizanov & Gribanova 2015, Adamson 2019).

(4) a. kniga-ta
        book.F-DEF
        'the book'

Ekaterina Georgieva

    b. nova-ta   kniga
       new.f-DEF book.f
       'the new book'

The adjectives in (1) are exceptional with respect to the placement of DEF: the definiteness marker cannot attach to them (5a), unlike what is observed with regular adjectives (4b). Instead, it skips the adjective and attaches to the noun as in (5b) (Adamson 2019 refers to this as 'skipping'). Importantly, the skipping variant in (5b) is grammatical only for some speakers; for others, DEF cannot be used with inflectionless adjectives at all (see Spencer & Luís 2012).

(5)   a. * sert-ăt   măž
            assertive-DEF man.M
            'the assertive man'

       b. % sert   măž-ăt
             assertive man.M-DEF
            'the assertive man'

Thus, inflectionless adjectives have been brought into the discussion of DEF as they raise two questions: (i) why DEF cannot attach to them, but appears on the noun instead (Adamson 2019, 2020); (ii) why DEF cannot be used at all (for some speakers) (see Halpern 1995). These issues will be addressed from a new perspective in §2.

    Halpern (1995: 165, fn. 22) proposes that inflectionless adjectives form neologistic compounds with the noun. Under a compound analysis, the placement of the definiteness marker on the noun, i.e., the head of the compound, is not surprising. On the other hand, Adamson (2019) argues these are not compounds since the adjectives can be intensified (6), can stand in the comparative form (7), and need not be adjacent to the noun (8).

(6)   mnogo serbez dete
       very   bold   child.N
       'a very bold child'

(7)   po-serbez
       CMPR-bold
       'bolder'

(8)   ? erbap   bălgarsko dete
        stubborn Bulgarian.N child.N
        'a stubborn [sic: capable] Bulgarian child'   (Adamson 2019: 94)

All these diagnostics are taken by Adamson (2019: 94) to indicate that we are dealing with adjectives (or adjective-like modifiers) and, in his view, to falsify the neologistic compound analysis proposed by Halpern (1995). Adamson (2019) proposes that DEF moves postsyntactically to the head of the closest phrase that bears nominal features. Adjectives undergo node-sprouting, as a result of which *a*Infl elaborates the adjective (the M(orphological)W(or)d *a*, to be precise). As this operation precedes the (postsyntactic) Lowering of D (see Embick & Noyer 2001), the definiteness marker ends up on the adjective, as in (4b). In order to account for the inflectionless adjectives, Adamson (2019: 96) proposes that the adjectival heads combining with certain loanword roots bear the diacritic feature [α] and the node-sprouting rule gives no results in the presence of this feature. This is how these *a*Ps are rendered inflectionless. Since the definiteness marker is sensitive to the nominal features present, two possible scenarios arise in the case of the inflectionless adjectives: (i) DEF attempts to attach to the inflectionless adjective and the derivation crashes, (ii) DEF skips the adjective and attaches to the noun instead. The two scenarios are meant to capture the interspeaker variation (recall that (5b) is acceptable only for some speakers, according to Adamson's data). Without going into further detail, I would like to note that this analysis refers to a list of vocabulary items, i.e., it relies on the properties of inflectionless adjectives as specified in the lexicon. As already pointed out in the introduction, the construction in question also occurs with non-loan nouns (cf. (3)), this phenomenon cannot possibly be fully derived from the exceptional features of loanwords.

Furthermore, although I agree with Adamson (2019) in his criticism of Halpern (1995), there are some remarks to be made here. First, example (6) indeed proves that we are not dealing with compounds, as the possibility of using adverbials like *mnogo* 'very, a lot' suggests that a degree phrase (DegP) is present, and more generally, that we are dealing with phrasal modifiers. However, *mnogo* 'very, a lot' also appears with verb phrases in Bulgarian. Additionally, the comparative clitic in Bulgarian can attach to nouns and verb phrases (e.g., *pò măž* 'CMPR man, more of a man/a real man' and *pò običam* 'CMPR love.PRS.1SG, I love more/prefer'). Thus, the data in (6) and (7) do not convincingly prove the adjectival status of these modifiers. As for example (8), the judgments indicate degraded acceptability. I discuss the possibility of other adjectival modifiers interleaved between the inflectionless adjective and the noun in §4.

Putting the issues about the placement of DEF aside, I will show that inflectionless adjectives differ from genuine adjectives both semantically and syntactically with respect to definiteness and syntactic distribution. In the next section, I will present novel empirical data regarding these loans and the noun phrases they appear in.

Ekaterina Georgieva

2 New empirical data

In this section, I will discuss new data on inflectionless adjectives on the basis of which the following generalisations emerge: (i) the noun phrases with inflectionless adjectives split into two groups with respect to definiteness: some of them are compatible with a definite reading, while others are not, (ii) these noun phrases show a limited syntactic distribution, and (iii) they have a strong exclamative flavour. The data presented here were tested with three native speakers, including myself; additionally, corpus examples from the Bulgarian National Corpus [BulNC] are also included.[2]

Let us begin with the use of the definiteness marker. As said in the previous section, this suffix cannot attach to the adjective (for all speakers), but according to the literature, for some speakers, it can attach to the noun as in (9) ("skipping").

(9) % Erbap   žena-ta   (se   obadi).
     skillful woman.F-DEF   REFL   call.PST.3SG
     'The skillful woman called.'   (Adamson 2020)

Firstly, I will argue below that inflectionless adjectives fall into two groups: with some of them, 'skipping' is perfectly fine, while with others it is not. It will be shown that *erbap* 'capable, skillful, cocky' belongs to the latter group. Secondly, I will also demonstrate that the acceptability of the definiteness marker in this group of inflectionless adjectives depends on the definiteness of the noun phrase. This sheds new light on the interspeaker variation reported in the previous literature. Before I proceed with the investigation of the restrictions on definiteness in these noun phrases, let me make an important methodological remark. Most of the examples in literature (Adamson 2019, 2020, Spencer & Luís 2012, Halpern 1995) are not full sentences, but simply Adj+N combinations and the definiteness of the noun phrases is not controlled for. This might have been the reason why certain speakers have accepted the examples, perhaps having in mind one particular reading, while others have rejected them, and this might have given the false appearance of interspeaker variation being present. In order to control for definiteness, I tested the noun phrases with inflectionless adjectives in full sentences. An additional problem is the exclamative flavour of the noun phrases with inflectionless adjectives (see below), which can be also controlled for by using full sentences.

In order to investigate the use of the definiteness marker and the definiteness of these noun phrases, I collected corpus data from BulNC. Based on these data,

---

[2]The corpus contains 1.2 billion words and is available online at: http://search.dcl.bas.bg/. The searches were carried out in May 2021.

it can be shown that noun phrases with inflectionless adjectives may contain a zero article, an indefinite article ('one') or a demonstrative. As for the use of the definiteness marker, it is not attested in the corpus with the inflectionless adjectives listed in (10a), but it is attested with the ones given in (10b).

(10)   a.   *serbez* 'bold, insolent', *ursuz* 'crabby, mean', *erbap* 'capable, skillful, cocky', *sert* 'assertive, testy, strong, quick-tempered', *češit* 'weird, crank', *inat* 'stubborn, obstinate', *piškin* 'experienced, spirited'[3]

       b.   *kofti* 'bad, shitty', *pišman* 'fake, feigned, sham'

Let us take a closer look at the items in (10a) based on native-speaker intuitions. I tested what readings these inflectionless adjectives allow for with various types of determiners in order to obtain a more fine-grained picture of the types of definiteness possible with them and to verify whether the fact that they are not attested with the definiteness marker in the corpus is merely accidental. Below I summarize the judgments.

First, noun phrases with the inflectionless adjectives in (10a) can have an indefinite article ('one'), yielding an indefinite non-specific reading (11). An indefinite specific reading as in (12) is also possible, but it requires a proper context like the relative clause given in parenthesis in the translation line in order to facilitate the specific reading of the noun phrase.[4]

(11)   Obadi   mi   se   edna   ursuz   žena.
      call.PST.3SG to.me REFL one.F crabby woman.F
      'A crabby woman called me.'   [*one*; indefinite non-specific ✓]

---

[3]There are no hits for *mukajat* 'determined, proactive' in the corpus, with any determiner, but based on my native speaker intuitions it belongs to the group in (10a).

[4]Compare (12) with the minimally different (i.a), which, in an out-of-the-blue context, can be used only humorously, i.e., as an epithet. Note that (i.a) is also different from (i.b), which shows that noun phrases without inflectionless adjectives can readily have a specific reading without requiring much contextualization.

(i)   a.   ?/# Tărsja   edna ursuz žena.
            look.for.PRS.1SG one.F crabby woman.F
            'I am looking for a crabby woman.'

       b.   Tărsja   edna žena.
            look.for.PRS.1SG one.F woman.F
            'I am looking for one (specific) woman.'

Ekaterina Georgieva

(12) (?) Hodih   da   tărsja   edna ursuz žena.
        go.PST.1SG CMPR look.for.PRS.1SG one.F crabby woman.F
        'I went looking for a crabby woman (e.g., who had called me the day
        before).'   [*one*; indefinite specific ✓]

Turning to the definiteness marker, we see that its use is highly degraded when
the noun phrase has a definite reading as in (13), but it is acceptable if the noun
phrase is interpreted generically as in (14).[5],[6]

(13) ?? Ursuz žena-ta   pak mi   se   obadi.
        crabby woman.F-DEF again to.me REFL call.PST.3SG
        'The crabby woman called me again.'[7]   [DEF, definite ✗]

(14) Ursuz žena-ta   se   poznava   po pogled-a.
        crabby woman.F-DEF REFL recognize.PRS.3SG by gaze.M-DEF
        'You can recognize a crabby woman by her gaze.'   [DEF, generic ✓]

The contrast between (13) and (14) is important because Halpern (1995), Spencer
& Luís (2012), and Adamson (2019) have claimed that DEF is either ungrammatical
altogether or that it is grammatical but only for some speakers. But what we
observe is that first, there is a contrast between the items in (10a) and (10b) and
second, as far as the group in (10a) is concerned, the grammaticality of DEF is not
a real case of interspeaker variation, as examples (13) and (14) are given different
judgements by the same speakers. Rather, the acceptability depends on the type
of definiteness of the noun phrase.

Spencer & Luís (2012: 129) argue that the impossibility of using DEF with in-
flectionless adjectives cannot be explained with restrictions on definiteness since
demonstratives are licit:

(15) Tazi ursuz žena   pak mi   se   obadi.
        this.F crabby woman.F again to.me REFL call.PST.3SG
        'This crabby woman called me again.'   [DEM ✓]

Indeed, such examples are fully acceptable, even for the speakers who reject (13),
and examples with demonstratives are also attested in the corpus. However, one
remark has to be made. It is well-known that demonstratives differ from definite

---

[5](13) was modelled after (9), which is very degraded to the speakers I have consulted.

[6]The DEF marker in Bulgarian can be used with generic noun phrases, both in the singular and
in the plural (see Nicolova 2017: 165).

[7]This sentence was tested in the following context: 'Yesterday I talked to Maria about the project,
and she really annoyed me. Today...'

articles in both deictic and anaphoric contexts (Lyons 1999, Wolter 2006; a.o.). Additionally, demonstratives may have several discourse/ pragmatic functions. For instance, they may have an indefinite specific reading as in (16a); according to Ionin (2006), in this case specificity is to be explained with noteworthiness. The indefinite use is often subsumed under a broader category, namely, the so-called emotive use of demonstratives, as illustrated in (16b) and (16c) (see Lakoff 1974, Wolter 2006, Potts & Schwarz 2010).

(16)   a. Mary wants to see *this* new movie; I don't know which movie it is, but she's been all excited about seeing it for weeks now.   (Ionin 2006)

       b. *that* mother of John   (Lakoff 1974)

       c. How's *that* throat?   (Lakoff 1974)

Importantly, the demonstrative in (15) cannot be interpreted deictically, i.e., the sentence cannot be uttered felicitously when pointing at someone. According to my intuitions, the referent is interpreted as specific and it must be salient in the discourse (at least on part of the speaker), but it does not need to be unique as with definites. Thus, (15) can also be uttered felicitously if the speaker has several crabby women in mind, but wishes to mention only one of them.

So far I have demonstrated that the grammaticality of the inflectionless adjectives given in (10a) depends on definiteness. These were also unattested with DEF in the corpus. Based on the corpus data, however, we saw that there are two 'outliers', namely, *kofti* 'bad, shitty' and *pišman* 'fake, feigned, sham' in (10b): with these lexical items, the skipping examples are perfectly fine, even with a definite reading. Two corpus examples are given below: (17) is most likely to be interpreted generically (as it combines with a mass noun), but (18) clearly has a definite reading: the NP has a unique referent, previously mentioned in the discourse. According to my native speaker intuitions, these items do not posit the restrictions on definiteness we observed for the ones in (10a).

(17)   […] az se   nasočih   kăm štand-a   s   kofti hrana-ta […]
             I   REFL direct.PST.1SG to   stall.M-DEF with bad food.F-DEF
       'I headed towards the junk food section.'   [BulNC]

(18)   Kofti kopele-to,   radist-ăt,   izpratil   săobštenie-to na
       bad   bastard.N-DEF radio.operator.M-DEF send.PTCP message.N-DEF on
       anglijski, beše   povišen   v staršina.
       English   be.PST.3SG promote.PTCP in sergeant.major
       'That idiot bastard, the radio operator, who (had) sent the message in
       English, was promoted to sergeant major.'   [BulNC]

Ekaterina Georgieva

The second important observation is that the noun phrases with inflectionless adjectives have a limited syntactic distribution.[8] Based on corpus data, it seems that these noun phrases tend to occur in the following syntactic environments: (i) as predicates of copular clauses (19), (ii) in exclamations (20), (iii) with predicates like *dărži se* 'behave, act (like)', *izgležda* 'look like, seem', *izliza* 'turn out (to be)', *okazva se* 'turn out (to be)', *minava (za)* 'be considered (as)', *ostava si* 'remain, to continue to be' (21). These predicates normally select for a predicative complement (a small clause), and more generally, the syntactic environments in (i)–(iii) are similar to each other, as they all express a predication relation, and thus can be subsumed under one more general type, namely, predication.

(19) Marija e   mnogo ursuz / kofti čovek.
      Maria be.PRS.3SG very   crabby bad person.M
      'Maria is a very crabby / bad person.'

(20) a. Eh, kakăv   ursuz / kofti čovek!
         PRT what.kind.M crabby bad person.M
         'What a crabby / bad person!'

      b. ursuz / kofti čovek
         crabby   bad person.M
         '(a) crabby / bad person' *or*
         'What a crabby / bad person!'

(21) a. Izleze   erbap   žena   tja.
         turn.out.PST.3SG capable woman.F she
         'She turned out to be a capable woman.'   [BulNC]

      b. Tja izleze   kofti čovek.
         she turn.out.PST.3SG bad   person.M
         'She turned out to be a bad person.'

In §4, I will argue that inflectionless adjectives stand in a predication relation with the noun, with the two groups of them, (10a) an (10b), exemplifying two different predication structures within the noun phrase, the comparative and the attibutive one, respectively. I will argue that inflectionless adjectives actually combine with a null noun and that this noun phrase functions as the predicate of the nominal small clause. This analysis is supported by the fact that these nominal small clauses are attested not only with the exceptional items traditionally

---

[8]For the inflectionless adjectives in (10a) it comes as no surprise that they have a fairly limited distribution in argument position, as they are incompatible with a definite reading.

referred to as inflectionless adjectives, but also with two noun phrases containing non-loan lexical items as in (22) and (23). Observe also that these two examples show similar restrictions regarding definiteness: the definiteness marker is highly degraded when used with a definite reading, but it is acceptable with a generic reading, as in (23) (compare with (13) and (14), respectively). These constructions also have an exclamative flavour.

(22) ?? Leke čovek-ăt   pak postăpi   užasno.
      stain person.M-DEF again behave.PST.3SG awfully
      'That scoundrel/rotter of a man behaved awfully again.'   [definite ✗]

(23) Leke čovek-ăt   se   poznava   po postăpk-i-te.
      stain person.M-DEF REFL recognize.PRS.3SG by deed-PL-DEF.PL
      'You can recognize a scoundrel/rotter of a man by his deeds.' [generic ✓]

Let us recap the main empirical points presented in this section. First, it was shown that inflectionless adjectives fall into two groups: some of them show restrictions on definiteness, as they are compatible only with the generic use of DEF, but not with the definite one. The other group of inflectionless adjectives do not show such restrictions. These facts further qualified the claims about the use of DEF with these lexical items made in the existing literature. Second, the noun phrases with inflectionless adjectives also show a limited syntactic distribution, being mostly used in predicative contexts. Thirdly, these noun phrases have a strong exclamative flavour. Finally, it was shown that nominal small clauses are also possible with non-loan items; moreover, those show a parallel behaviour with respect to the use of the definiteness marker.

3 Background on nominal predication

In a nutshell, my proposal regarding inflectionless adjectives will be that they stand in a predicational relationship with the noun. I argue that these constructions are comparable to well-studied cases of nominal predication (see Napoli 1989, Hulk & Tellier 2000, Doetjes & Rooryck 2003, Casillas Martínez 2003, Villalba & Bartra-Kaufmann 2010 among others on Romance languages; Den Dikken 2006 on English and Dutch; Den Dikken & Lipták 1997 on Hungarian). Below, I will first provide a summary of the main types of nominal predication and their properties, focusing mostly on English and Spanish, based on the existing literature. I will also summarize the main analytical solutions proposed.

Ekaterina Georgieva

Den Dikken (2006) argues that predication structures in the noun phrase come in two guises: ATTRIBUTIVE NOMINAL PREDICATION and COMPARATIVE NOMINAL PREDICATION, as illustrated below:

(24) a. an idiot doctor, an idiot of a doctor   [attributive]

b. a jewel of a village, an idiot of a man   [comparative]

Den Dikken (2006: 161) points out that the two types of nominal predication are not simply semantic variants to each other, as evidenced by the structural differences between the two types in Italian (examples from Napoli 1989). In the attributive type, which has the meaning that the referent of the complex noun phrase is an ignoramus in his capacity as a doctor, the second noun is bare (25a). In the comparative type, on the other hand, the second noun bears a definite determiner (25b). The meaning of the latter type is that the referent of the complex noun phrase is ignorant as an individual (and just happens to be a doctor by profession). In a similar vein, although the English examples like *an idiot of a doctor* are ambiguous between the two readings, it can be shown that the two noun phrases participating in the nominal predication are obligatorily connected by *of a* in the comparative nominal predication but not in the attributive one, as evidenced by (26) (Den Dikken 2006: 164).

(25) a. quell' ignorante di dottore
       that   ignoramus of doctor
       'that ignoramus (of a) doctor'   [attributive]

b. quell' ignorante del   dottore
       that   ignoramus of-the doctor
       'that ignoramus of a doctor'   [comparative]

(26) a. That idiot (of a) doctor prescribed me the wrong medicine.
                                                             [attributive]

b. That idiot #(of a) doctor just wrecked my car.   [comparative]

What is common between the two types is that in structural terms both are small clauses, i.e., they express a predicational relationship between a subject and a predicate. The difference between them is that they correspond to different syntactic structures according to Den Dikken (2006). In the attributive type, the predicate is in the specifier of the small clause (which is a R(elator)P in his terms), see Figure 1. The comparative type, on the other hand, is different: the predicate is base-generated in the complement position of the small clause, but subsequently undergoes predicate inversion, which derives the surface order, see Figure 2.

6 *Inflectionless adjectives in Bulgarian as a case of nominal predication*
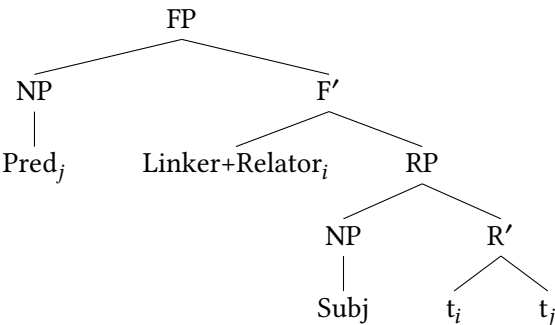
```
              RP
         ┌─────┴─────┐
        NP           R′
         │        ┌───┴───┐
        Pred   Relator    NP
                           │
                          Subj
```
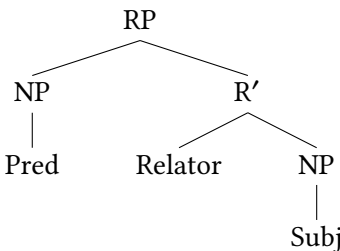
Figure 1: Attributive QBNPs

```
              FP
        ┌──────┴──────┐
       NP             F′
        │        ┌─────┴─────┐
      Pred_j  Linker+Relator_i   RP
                          ┌──────┴──────┐
                         NP             R′
                          │         ┌───┴───┐
                        Subj       t_i     t_j
```

Figure 2: Comparative QBNPs

The trees in Figure 1 and 2 illustrate the main structural difference between the two types of nominal predication: the attributive type is "born" as an inverse predication structure as in Figure 1, while in the comparative type the predicate acquires its surface position via movement as in Figure 2. A further difference concerns the size of the subject and predicate noun phrases (labeled as NP for convenience) and the functional heads connecting them.

Let us take a closer look at the attributive type. Den Dikken (2006: 166–168) argues that structurally, it has two subtypes: *an idiot doctor* and *an idiot of/as a doctor*. Both are small clauses with the structure as in Figure 1, i.e., the predicate is base-generated in the specifier position. In the former case, the subject and the predicate are bare NPs. The small clause is then embedded under a nominal layer (NumP), which derives the external nominal distribution. Attributive nominal predications like *an idiot of a doctor*, on the other hand, contain larger nominal projections as their subparts: both the subject and the predicate are NumPs (as they have an indefinite article) and the small clause is topped off by (a zero) D, giving rise to nominal external syntax. Den Dikken (2006: 166–168) argues that *of* in the attributive nominal predication is a nominal copula that lexicalizes the Relator head.

Ekaterina Georgieva

Turning to the comparative nominal predication, Den Dikken (2006: 175–181) proposes that its predicate starts out in the complement position of the small clause, but undergoes predicate inversion, as the result of which the surface order is derived (see Figure 2). The mechanics of this predicate inversion is the following. First, the Relator moves up to the small clause-external F^0. As a result of this phase-extending movement, the predicate is allowed to move to the specifier of FP; that being a case of A-movement. Den Dikken (2006) argues that the predicate inversion is triggered by the need for licensing an empty head. This empty head, SIMILAR, is part of the predicate; this is how the semantics of comparison is encoded. The comparative type of nominal predication is different from the attributive type with respect to the functional heads connecting the two noun phrases. Den Dikken (2006) argues that in comparative QBNPs the Relator head is spelled out by the spurious indefinite article (based on evidence from Dutch and Hungarian). In addition, comparative QBNPs also feature a Linker spelling out F^0, namely, the nominal copula *of*. The nominal copula *of* is argued to be similar to the obligatory copula in copular inversion constructions (e.g., *I consider the best candidate *(to be) John*). The nominal small clause acquires its outwardly nouny distribution by virtue of being topped off by a NumP layer that harbours the (indefinite) outer determiner in examples like *a jewel of a village*. Moreover, in Den Dikken's (2006) account, both the subject and the predicate are NumPs rather than bare NPs. Hence, comparative nominal predications that would correspond to attributive ones of the type *an idiot doctor* are not possible in English (*a jewel village* is a case of N-N compounding rather than of comparative nominal predication, see Den Dikken 2006: 163–164, 173).

Spanish utilizes two types of nominal predication constructions (see Villalba & Bartra-Kaufmann 2010 for an in-depth discussion). The first one is the so-called *lo-de* construction (27a): the subject of the small clause (*la casa* 'the house') is preceded by an adjective in the neuter and the neuter article *lo*. The second type is the qualitative binominal noun phrase (QBNP) in which two noun phrases participate as the subject and the predicate of the small clause (27b). Both (27a) and (27b) are analysed by Villalba & Bartra-Kaufmann (2010) as small clauses, having the underlying structure of (28a) and (28b), respectively. (The small clause is labeled as XP in their study.)

(27) a. lo   caro   de la   casa
        DEF.N expensive.N of DEF.F car.F
        'the (high degree of) expensiveness of the house' [*lo-de* construction]

    b. el   idiota del   alcalde
       DEF.M idiot.M of.DEF.M mayor.M
       'that idiot of a mayor'   [Spanish QBNP]

(28)   a. $[_{XP}$ $[_{DP}$ *la casa* $]$ $[_{X'}$ X … $[_{AP}$ *car-* $]]]$ = (27a)

        b. $[_{XP}$ $[_{DP}$ *el alcalde* $]$ $[_{X'}$ X $[_{DP}$ *idiota* $]]]$ = (27b)

The *lo-de* construction requires some more explanation. First, the predicate of the small clause is argued to have a more complex structure than what was shown in (28a). As the predicate semantically expresses a high degree quantification, it is argued to contain a DegP on top of the adjectival phrase. Furthermore, the DegP also contains a silent DEGREE head.[9] Finally, the specifier of this phrase hosts a comparative operator. Thus, the structure of the predicate in *lo-de* constructions is as in (29).

(29) $[_{DegP}$ Op $[_{Deg'}$ DEGREE $[_{AP}$ Adj $]]]$

With these assumptions about the structure of the small clause in (28a) in mind, let us proceed to how the surface order of (27a) is derived. This is argued to be the result of three subsequent steps of movement. Firstly, similarly to Den Dikken (2006), Villalba & Bartra-Kaufmann (2010) also assume that the DEGREE head must move to F[0] where it is lexicalized as *de* 'of'. Then, the whole DegP moves to the specifier of the DP-internal FocP, yielding an information-structural partition of the nominal predication construction where the predicate is a focus and the subject is a background topic. The final step is that the operator hosted in SpecDegP moves to SpecDP. Villalba & Bartra-Kaufmann (2010) argue that the exclamatory flavor of the construction arises from the combination of a degree quantificational structure with the definiteness of the Det head: the null degree operator is argued to function like a wh-element. The three movement steps are shown in Figure 3.

    Thus, Villalba & Bartra-Kaufmann's (2010) account is similar to Den Dikken's (2006) analysis as it assumes that the predicate of the nominal small clause undergoes movement, but it crucially argues that this is an A′-dependency, tied in with the information-structural properties of the construction. It should be mentioned that there are also other approaches that assume A′-movement of the

---

[9]This is in a way similar to Den Dikken's proposal that the predicate of comparative nominal predication constructions contains an additional component, i.e., the SIMILAR head, though according to Villalba & Bartra-Kaufmann (2010), only the *lo-de* construction contains a DegP, while Spanish QBNPs do not: in them the evaluative property of the predicate is lexically encoded.
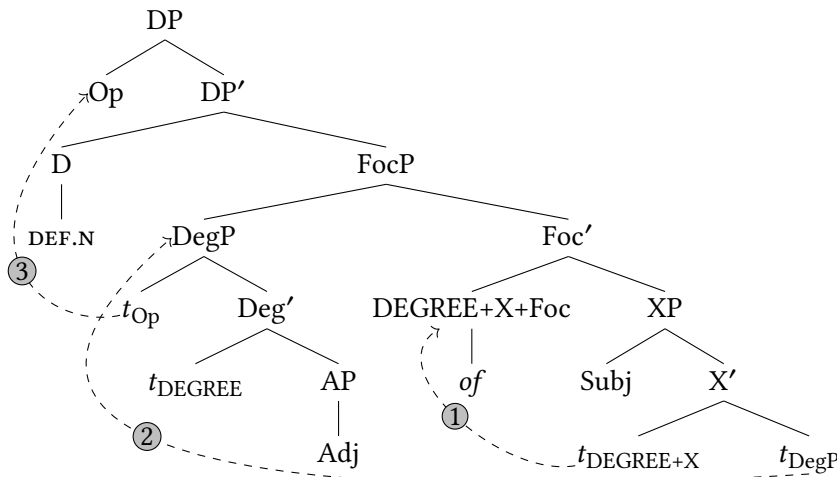
Figure 3: The Spanish *lo-de* construction

predicate, the difference between them being the landing site of the moved predicate: SpecDP (Kayne 1994), SpecCP (Doetjes & Rooryck 2003), or a DP-internal SpecFoc position (Villalba & Bartra-Kaufmann 2010). It is also noteworthy that the idea that the exclamative flavor of the *lo-de* construction is linked to the movement of a null operator is also found in other works: for example, Hulk & Tellier (2000) propose for French QBNPs that the head of the small clause moves because of an affective operator in its predicate.

In sum, it has been shown that nominal predication has two types: attributive and comparative. The latter involves movement of the predicate, but the existing analyses differ as to whether this is a case of A- or A′-movement. Having summarized the main semantic and syntactic properties of nominal predication in English and Spanish, let us turn to the Bulgarian data.

4 Towards an analysis of nominal predication in Bulgarian

In this section, I lay out an analysis of noun phrases with inflectionless adjectives in Bulgarian in terms of nominal predication. Importantly, I suggest that the same analysis can be extended to cover small clauses containing two noun phrases as their subparts, thus, the construction in question is not limited to in-

flectionless adjectives. In the course of this section, I will make the following claims regarding the semantics and the structure of these nominal small clauses:

(30)   a. There is a subject–predicate relationship between the two elements in the complex noun phrase, i.e., we are dealing with a nominal small clause.

   b. Two types of nominal predication are to be distinguished: an attributive and a comparative one.

   c. The attributive one is an inverse predication structure in the sense of Den Dikken (2006); the comparative one involves movement of the predicate.

   d. In both types, the noun phrases in the nominal small clauses are bare NPs.

   e. In both types, the predicate is a noun phrase: the inflectionless adjective modifies a null noun.

In what follows, I will first provide evidence for the subject-predicate relation and for the existence of two types of nominal predication (§4.1). Then I will argue that the predicate of the small clause is a noun phrase in which the inflectionless adjective modifies a null noun (§4.2). Finally, I will discuss the structure of the noun phrases containing nominal predication (§4.3).

4.1 Two types of nominal predication in Bulgarian

I argue that the two elements in noun phrases with inflectionless adjectives stand in a subject-predicate relationship. Support for this comes from the entailments in (31a) (based on Villalba & Bartra-Kaufmann 2010). The continuations given in brackets are perceived as contradictions rather than implicature cancellations.[10]

(31)   a. ursuz   / kofti čovek
         crabby bad person.M
         '(a) crabby/bad man'
         ⇒ 'The man is crabby/bad (#but he's actually not crabby/bad).'

---

[10]In contrast, garden-variety adjectives can be cancelled: the continuation in (i) does not feel like a contradiction but rather as an implicature cancellation.

   (i)   Tja e   krasiva žena.
         she be.PRS.3SG beautiful.F woman.F
         'She is a beaufitul woman (but she's actually not (that) beautiful).'

Ekaterina Georgieva

    b. leke   čovek
        stain.N person.M
        '(a) rotter of a man'
        ⇒ 'The man is a rotter (#but he's actually not a rotter).'

Additionally, as argued by Villalba & Bartra-Kaufmann (2010) for Spanish, the subject-predicate relation is also constrained lexico-semantically, as the Spanish *lo-de* construction cannot contain stage-level predicates, but only individual-level predicates. The lexical items participating in the Bulgarian construction illustrated in (31a) are a closed class, thus, we cannot make a compelling argument based on this parallel. But still, it can be observed that all items in (10a) and (10b) are individual-level predicates.

    Furthermore, I argue that the two types of nominal small clauses distinguished by Den Dikken (2006), namely, the attributive one and the comparative one, are also attested in Bulgarian. Specifically, I propose that the lexical items in (10a) participate in comparative small clauses, while the ones in (10b) are used in attributive nominal predication. The two types can be distinguished semantically when combined with profession-denoting nouns. Comparative nominal small clauses like (32) are more naturally interpreted as 'X is a crabby person in general', rather than 'X is crabby (only) in his capacity of a standard bearer'. On the other hand, in attributive small clauses like (33), the meaning is such that 'X is bad in his capacity of policeman/driver'.

(32) Da ne beše   toja tvoj ursuz bajraktar […]
      COMP NEG be.PST.3SG this.M your.M crabby standard.bearer.M
      'If it wasn't this crabby standard bearer of yours […]'   [BulNC]

(33)   a. kofti policaj
        bad policeman
        'bad policeman (e.g., corrupt)'
    b. pišman šofjor
        fake driver
        'bad driver (e.g., not having a driving license)'

This is also confirmed by the following contradiction test (see den Dikken 2006: 170): the English comparative nominal predication is infelicitous in such a context, while the attributive one is perfectly fine (34). The Bulgarian examples in (35) are parallel to the English ones. This test provides further support for the proposal that both types of nominal small clauses are attested in Bulgarian.

(34) a. That idiot of a doctor is not an idiot (as a person).   [attributive]

     b. # That idiot of a man is not an idiot.   [comparative]

(35) a. Tozi   kofti policaj.M (vsăštnost) ne e   kofti kato
         this.M bad   policeman actually   NEG be.PRS.3SG bad   as
         čovek.
         person.M
         'This bad policeman is actually not bad as a person.'

     b. # Tazi   ursuz žena   (vsăštnost) ne e   ursuz.
         this.F crabby woman.F actually   NEG be.PRS.3SG crabby
         'This crabby woman is actually not crabby.'

Having defended the claims in (30a) and (30b), namely, that we are dealing with nominal small clauses and that these small clauses fall into either the attributive or the comparative type, let us move to their structure.

I adopt the main insight of Den Dikken's (2006) analysis: attributive nominal small clauses are inverse predication structures in which the predicate is base-generated in the specifier position, while comparative nominal small clauses involve movement of the predicate to a higher position in order to derive the surface order. Thus, I am assuming the structures in (36a) and (36b), respectively. (The small clause is labeled as XP, that being the most theory-neutral term, instead of R(elator)P as in Den Dikken 2006.)

(36) Nominal predication in Bulgarian (1st version)

     a. [... [$_{XP}$ Predicate [$_{X'}$ X Subj]]]   [attributive]

     b. [... [$_{FP}$ Predicate$_i$ [$_{F'}$ F [$_{XP}$ Subj [$_{X'}$ X t$_i$]]]]]   [comparative]

The proposed structures account for the semantic differences between the two types: in (36a), the predicate is given an attributive interpretation, as it ascribes a(n additional) property to the referent of the noun phrase, while in (36b), it draws a comparison in such a way that the subject is understood to intrinsically show the property denoted by the predicate and to be identifiable with it.

In the next two subsections, I will discuss the structure of these nominal small clauses in greater detail.

4.2 Inflectionless adjectives combine with a null noun

I propose that the predicates of nominal small clauses in Bulgarian are noun phrases. Furthermore, I argue that inflectionless adjectives in Bulgarian combine

with a null noun. Thus, the QBNPs containing them are actually binominal. This is supported by the fact that nominal predication in Bulgarian is also possible when the predicate of the small clause is a (non-loan) noun as in (23). I propose the following underlying structure for the predicate of both the attributive and the comparative types of nominal predication with inflectionless adjectives (37). For concreteness, I assume that the null noun is semantically roughly equivalent to 'kind/type/quality'.

(37)   [$_{NP}$ Adj [$_{NP}$ NOUN]]

This proposal not only allows us to unify examples like (23) and the ones containing inflectionless adjectives but also to explain several properties of inflectionless adjectives. For example, the fact that they allow for degree modification like *mnogo* 'very', but are outwardly nominal follows from this. In this sense, they are similar to well-known cases of adjectives combining with a null noun like *the rich/the poor*, which also allow for adverbial modification of the adjective (cf. *the very poor*). This has been taken to suggest that the adjective modifies a null noun (see Kester 1996, Giannakidou & Stavrou 1999).

Evidence for positing a null noun comes from the use of the inflectionless adjectives in copular clauses. In Secion 2, I showed that QBNPs are often predicates of copular clauses (example (19) is repeated in (38) for the reader's convenience).

(38)   Marija e   mnogo ursuz   / kofti čovek.
        Maria be.PRS.3SG very   crabby bad person.M
        'Maria is a very crabby / bad person.'

In addition to this, inflectionless adjectives also have what may look like a stand-alone use as predicates of copular clauses. This is illustrated in (39) with *kofti* 'bad, shitty' (cf. (10b)), but it is also possible with the adjectives in (10a).

(39)   Prognoza-ta   za vreme-to   e   mnogo kofti.
        forecast.F-DEF for weather.N-DEF be.PRS.3SG very   bad
        'The weather forecast is very bad.'

One way to approach the example in (39) is to say that this is indeed a standalone use of the adjective, without positing a null noun – which would not be too surprising given that adjectives in Bulgarian can be used as predicates of copular clauses. However, this analysis is insufficient to account for the data in (40). In (40a), we see that the adjective cannot have a standalone use: the noun *čovek* 'person' following it cannot be omitted. At this point, one might wonder if the

standalone use is only possible with inanimates as in (39), but impossible with animates (humans) as in (40a). This analysis, however, is immediately falsified when we look at (40b): the adjective can have a standalone use, even though it refers to an animate (human) subject.

(40) a. Marija e   mnogo kofti #(čovek).
          Maria be.PRS.3SG very bad   person.M
          'Maria is very bad.'

      b. Učitelka-ta   e   mnogo kofti.
          teacher.F-DEF be.PRS.3SG very bad
          'The (female) teacher is very bad.'

Another way to approach these data would be to say that the adjective *kofti* 'bad, shitty' cannot occur in predicative position, as it is well-known that certain adjectives cannot occur as predicates, e.g., *bivš* 'former, ex' in Bulgarian. This could explain (40a), but not the contrast with (39) and (40b). Thus, the explanation cannot possibly be related to animacy or to the attributive/predicative use of the adjective itself.

In order to account for the triplet of data in (39)–(40), I propose the following. The adjective combines with a null noun that has the meaning 'kind/type/quality'. Thus, in (39) and (40b), we are not dealing with a standalone use of the adjective; rather, there is a null nominal modified by it. This also provides an explanation of why (40a) is infelicitous: the sentence underlyingly corresponds to '#Maria is (of) bad KIND/TYPE/QUALITY', which is semantically anomalous. The sentence improves if the noun *čovek* 'person' is present; in this case, I propose that we are dealing with a nominal small clause of the attributive type. That is, the subject of the small clause is *čovek* 'person' and the predicate is the noun phrase with the null noun modified by the adjective. The meaning of the attributive QBNP corresponds to 'Maria is (of) bad KIND/TYPE/QUALITY as a person', which is semantically perfectly fine.[11]

---

[11]At this point, the reader might have started to wonder whether *kofti čovek* 'bad person' should be classified as an attributive QBNP, i.e., 'bad *as a* person', as argued above, or as a comparative one like *ursuz čovek* 'crabby person'. As was shown above, attributive QBNPs are fine in contradiction contexts (cf. (34) and (35)) and this holds for *kofti čovek* 'bad person' in (i), thus verifying that we are dealing with an attributive QBNP:

(i) Marija e   kofti čovek, no e   dobăr učitel.
      Maria be.PRS.3SG bad   person but be.PRS.3SG good teacher
      'Marija is bad as a person, but is a good teacher.'

Ekaterina Georgieva

Thus, I propose that inflectionless adjectives *always* compose with a null noun, this giving rise to their "standalone" use, which is, as I argue, in fact a noun phrase with a null noun. This noun phrase can appear in predicative position in copular clauses as in (39) and (40).

Before I proceed further with the details of my analysis, let me discuss an alternative that has been proposed in the literature. Adamson (2019: 100–103) mentions the standalone nouny use of inflectionless adjectives and discusses two subtypes of this: (a) cases in which the adjective is used as a noun, e.g., *inat* 'stubbornness', for which he claims that the (acategorial) root is directly nominalized by *n*, and (b) cases in which the adjective appears as an appositive to a proper noun (41). In the latter case, he proposes that the (acategorial) root is first categorized by *a* (thus, degree modifiers will be possible), and then a nominal layer with a [+human] *n* is added to further nominalize it: $n \succ a \succ \sqrt{\text{INAT}}$.

(41) Ivan ({mnogo / po-}) inat-ăt
      Ivan very   CMPR stubborn-DEF
      'Ivan, the (very/ more) stubborn'   (Adamson 2019: 102)

The case I discuss above is similar to the second scenario in the sense that the adjective still preserves its properties with respect to modification. But as we see in (39), the referent need not be a human, so it is unlikely that the adjective is nominalized by a [+human] *n*. Besides, in my opinion, the triplet in (39)–(40) cannot be easily explained in an nominalization analysis. Finally, the example in (41) is very degraded for me with the adverbial/ degree modification; the perfectly grammatical variant is when *inat* is modified by the adjective *golemijat* 'big.DEF', which would be an example of the 'direct nominalization' strategy ($n \succ \sqrt{\text{ROOT}}$). This casts doubts whether $n \succ a \succ \sqrt{\text{ROOT}}$ is possible with inflectionless adjectives at all.

Thus, a nominalization analysis cannot sufficiently explain the properties of inflectionless adjectives, which I argue to be derivable from the presence of a null noun. In addition to postulating a null noun that combines with the adjective, I also propose that this noun phrase can be used as the predicate of a nominal small clause of either the attributive or the comparative type. In §1, I showed that inflectionless adjectives split into two groups with respect to the use of the definiteness marker (cf. (10a) and (10b)), and in §4.1, I argued that these two groups correspond to either the attributive or the comparative type of nominal predication. I tentatively submit that it depends on the lexical properties of the adjective whether the noun phrase that contains it (=(37)) can be used in an attributive or in a comparative QBNP.

Furthermore, on the assumption that the null noun the inflectionless adjective combines with is morphosyntactically deficient, i.e., lacking gender features, we would not expect the adjective to show concord with it. As for number features, I assume that the null noun is in the singular (or is specified as [−Pl]). This might seem to be a circular way of explaining why these loanword adjectives do not show inflection, but the deficiency of the null noun could in principle be relevant if we approach the question from yet another angle, namely, why non-loan adjectives do not combine with it: because they require a noun that they can show concord with. Thus, we predict them not to be able to combine with this null noun. The third consequence of the morphosyntactic deficiency of the null noun will become clear when we take a closer at the structure of the noun phrases containing QBNPs.

4.3 The structure of noun phrases with nominal predication

Based on the last two subsections we have arrived at the following structure for nominal small clauses in Bulgarian:

(42) Nominal predication in Bulgarian (2nd version)
    a. [... [$_{XP}$ [$_{NP}$ Adj [$_{NP}$ NOUN]] [$_{X'}$ X Subj]]]   [attributive]
    b. [... [$_{FP}$ [$_{NP}$ Adj [$_{NP}$ NOUN]]$_i$ [$_{F'}$ F [$_{XP}$ Subj [$_{X'}$ X t$_i$]]]]] [comparative]

The structures in (42a) and (42b) raise the following questions: (i) what is the internal structure (and size) of the subject of the small clause; (ii) what functional layers top off the small clause (informally marked by the ellipsis dots above); and (iii) what is the landing site of the predicate in the case of the comparative nominal predication (labeled above as SpecFP).

    The first two questions are somewhat interrelated and can be answered if we compare the Bulgarian examples with English QBNPs. Recall that in the case of English attributive QBNPs like *an idiot doctor*, Den Dikken (2006) argues that the subject and the predicate are bare NPs. I argue that nominal predication in Bulgarian is strikingly similar in this respect: in both the attributive and the comparative types, I propose that the nominal predication consists of bare NPs. This is supported by the fact that interleaved adjectival modifiers are highly degraded, as shown in (43) (*pace* Adamson 2019, cf. example (8) above).

Ekaterina Georgieva

(43) Bălgarsko-to   serbez dete   / ??serbez bălgarsko-to   dete   se
      Bulgarian.N-DEF bold   child.N   bold   Bulgarian.N-DEF child.N REFL
      poznava   po pogled-a.
      recognize.PRS.3SG by gaze.M-DEF
      'You can recognize the bold Bulgarian child by his/ her gaze (lit. The bold
      Bulgarian child is recognized by his/ her gaze).'

The example in (43) is peculiar if we assume that the inflectionless adjective is
a modifier like the adjective 'Bulgarian': in fact, nationality-denoting adjectives
usually precede quality-denoting ones in Bulgarian, thus, the grammatical word
order in (43) is unexpected if we are dealing with regular adjectival modifiers.
However, the word order restrictions fall out naturally if we assume that the in-
flectionless adjective and the noun form a small clause and that this small clause
consists of bare NPs.

This argument can be further strengthened when we look at recursion in nom-
inal predication. The example in (44) is interesting for several reasons. First, it
shows that there are two inflectionless adjectives involved. This might at first
sight be taken to contradict the claim made above that the subject of the small
clause must be a bare NP and cannot be modified by an adjective (as in (43)). But
in (44), *pišman* 'fake, sham' is interleaved between *ursuz* 'crabby' and the noun.
I take this to support the nominal predication analysis from yet another angle:
pišman 'fake, sham' is not a regular adjectival modifier, but participates in an
attributive QBNP. Then, the attributive QBNP acts as the subject of the compara-
tive QBNP. (44) also shows that the the reverse order of the adjectives is degraded.
(According to my intuitions, the meaning would be the same, which is in fact pre-
dicted, since it is lexically determined whether the adjective participates in the
attributive or the comparative type of nominal predication, as I argued above,
cf. the lists in (10a) and (10b)). The meaning and the word order of (44) fit very
nicely with the observations about recursion in QBNPs: N_2 in N_1-*of*-N_2-*of*-N_3 is
forced into an attributive reading, and comparative QBNPs are not recursive, cf.
*that beauty of a jewel of a village.[12] The example in (44) complies with this, as
the outer QBNP is of the comparative type and the inner one is of the attributive
type.

(44) tozi   ursuz pišman šofjor   / ??tozi   pišman ursuz šofjor
      this.M crabby fake   driver.M   this.M fake   crabby driver.M
      'this grump of an idiot driver'

---

[12] I thank Marcel den Dikken for the discussion of recursion and the English example.

Furthermore, I propose that in order to acquire a nominal external distribution, the small clauses are embedded under a nominal layer. This is in unison with the proposals made for languages like English, Dutch, Spanish, and Hungarian. But I would like to propose that nominal predication in Bulgarian takes place very low in the structure, at the *n*P level, which is in sharp difference with QBNPs in English and Spanish. In Bulgarian, further nominal layers can be built up on top on the *n*P (to harbour numerals, possessive pronouns, and demonstratives, cf. example (32)). With these two assumptions in mind, i.e., that the small clause contains bare NPs and is being topped off by *n*P, we can account for the ungrammaticality of interleaved adjectival modifiers. Since QBNPs are formed at the *n*P level, their referent is understood to be a single individual (thus, in a way, it is not surprising that previous accounts, e.g., Halpern 1995, have treated them as nominal compounds).

This analysis also allows us to make an interesting typological observation regarding the structure of nominal small clauses. It is noted by Den Dikken (2006: 168) that apart from 'bare' attributive nominal small clauses in English like *an idiot doctor*, he is not aware of this type of constructions from other languages. Thus, in a way, the Bulgarian nominal small clauses fill a gap in the typology of nominal predication. Attributive QBNPs like *an idiot doctor* are also 'bare' in the sense that they do not contain an overt copula/ linking element between the subject and the predicate of the small clause, in contrast with *an idiot of a doctor* and *a jewel of a village*. Similarly, nominal predication in Bulgarian shows no linking element, and thus conforms to the observation that the size of the subject and the predicate of the nominal small clause correlates with the presence of an overt linking element.

Finally, we need to discuss the landing site of the predicate in comparative QBNPs. I propose that it is precisely Spec*n*P that the predicate moves to. Recall from §3 that there are several proposals on market regarding the position targeted by the movement predicate. One option would be SpecDP, as proposed by Kayne (1994) for the predicate of English QBNPs; similarly, in Villalba & Bartra-Kaufmann's (2010) account, the operator hosted in SpecDegP in the Spanish *lo-de* construction lands in SpecDP (while the predicate moves to a DP-internal SpecFoc position). Additionally, movement to SpecDP has been also proposed for structures like *how tall a man* (see Hendrick 1990). However, movement to SpecDP faces some difficulties in the case of the Bulgarian nominal small clauses. The main problem comes from the order of nominal modifiers. It can be seen from (32), repeated below as (45), that comparative QBNPs can be preceded by demonstratives and pronominal possessive adjectives.

(45) Da   ne beše   toja tvoj   ursuz   bajraktar [...]
      COMP NEG be.PST.3SG this.M your.M crabby standard.bearer.M
      'If it wasn't this crabby standard bearer of yours [...]'   [BulNC]

As demonstratives in Bulgarian are said to be always in SpecDP (either being
base-generated there or being obligatorily moved there, cf. Dimitrova-Vulchanova
& Giusti 1995), the predicate of the small clause cannot possibly move to the very
same position. As the linear order in (45) is Dem > PronPoss > NP$_{pred}$ > NP$_{subj}$,
this suggests that the landing site of the predicate must be lower than the func-
tional projections that harbour demonstratives and pronominal possessors. I pro-
pose that Spec$n$P is an appropriate landing site for the moved predicate. Building
on the intuition in Den Dikken (2006), I argue that the movement has interpre-
tive effects: the subject NP is compared to the predicate NP. Furthermore, it has
been proposed for the Romance languages that the movement is related to the
exclamative flavour and/or the information-structural partition of the nominal
small clause (see Hulk & Tellier 2000, Villalba & Bartra-Kaufmann 2010). It has
been shown in §2 that Bulgarian QBNPs also have an exclamative flavour and
express emphasis on the predicate of the small clause. I would like to tentatively
propose that the movement to Spec$n$P also derives this property of the construc-
tion on the assumption that Spec$n$P can function as a low focus projection in
the nominal domain. As Bulgarian is generally considered to be a split-DP lan-
guage (for example, Dimitrova-Vulchanova & Giusti 1998 postulate a TopP on
top of DP), it is not implausible to assume that noun phrases in Bulgarian con-
tain a functional projection below D that can serve as the landing position of the
predicate of the small clause. Thus, the structures of attributive and comparative
QBNPs in Bulgarian are as shown in Figure 4 and Figure 5, respectively.

    Finally, I would like to briefly address the use of the definiteness marker in
nominal small clauses, as this was the main question discussed in the literature
on inflectionless adjectives, and different proposals have been made for why the
definiteness marker attaches to the noun rather than to the adjective. In my view,
one of the welcome consequences of the structures in Figure 4 and Figure 5 is
that the adjective is "buried" inside the noun phrase of the null noun, that is, it
is not a modifier of the subject NP. Since it is not in the structural position that
regular inflecting adjectives occupy in Bulgarian, it will not be visible for the
definiteness marker to attach to it. Similarly, the null noun itself would also be
invisible (either because of lacking a phonological representation or because of
its morphosyntactic deficiency). In §2, I refined the claims made in earlier stud-
ies, and the two most important conclusions were that the definiteness marker is
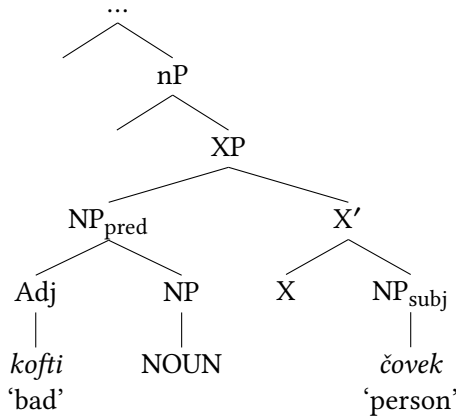fine with attributive QBNPs, while with comparative ones, it is limited to generic
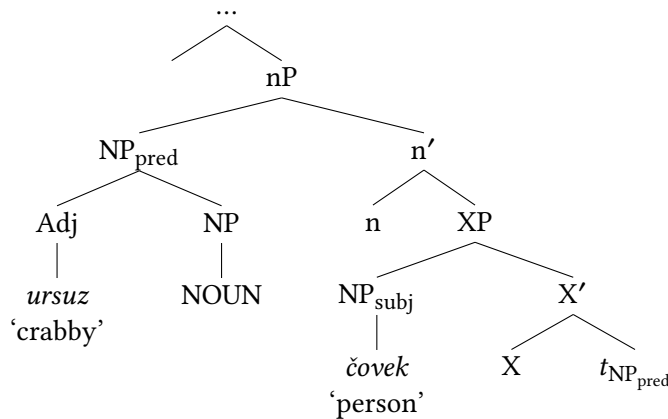
Figure 4: Attributive QBNPs in Bulgarian

Figure 5: Comparative QBNPs in Bulgarian

readings.I would like to submit that the movement of the predicate NP in comparative QBNPs results in the unavailability of definite readings with the definiteness marker. I would tentatively propose that this is due to the fact that the predicate NP is indefinite, which thus precludes the definite reading of DEF for the whole *n*P after the predicate has moved to Spec*n*P. Thus, the presence of a null noun in the structure can explain why inflectionless adjectives are "skipped" by the definiteness marker and also why comparative QBNPs disallow definite readings of DEF.

Acknowledgements

References

Adamson, Luke. 2019. *Derivational trapping and the morphosyntax of inflectionlessness.* Philadelphia, PA: University of Pennsylvania. (Doctoral dissertation).

Adamson, Luke. 2020. *On the transparency of inflectionless adjectives for definiteness marker placement in Bulgarian.* Talk given at the 29th Annual Meeting of Formal Approaches to Slavic Linguistics (FASL 29), May 2020.

Adamson, Luke. 2022. Transparency of inflectionless modifiers for Bulgarian definite marker placement. *Journal of Slavic Linguistics* 30(FASL 29 extra issue). 1–13. https://ojs.ung.si/index.php/JSL/article/view/78.

Bagasheva, Alexandra. 2017. On a subclass of nominal compounds in Bulgarian: the nature of phrasal compounds. In Carola Trips & Jaklin Kornfilt (eds.), *Further investigations into the nature of phrasal compounding*, 81–117. Berlin: Language Science Press. DOI: 10.5281/zenodo.885121.

Casillas Martínez, Luis D. 2003. Gender mismatches in Spanish and French *N1/A de N2* affective constructions: index agreement vs. morphosyntactic concord. In Jong-Bok Kim & Stephen Wechsler (eds.), *The Proceedings of the 9th International Conference on HPSG*, 1–17. Stanford, CA: CSLI Publications.

den Dikken, Marcel. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas.* Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/5873.001.0001.

den Dikken, Marcel & Anikó Lipták. 1997. *Csoda egy nyelv* – Nominal internal predication in Hungarian. In Jane Coerts & Helen de Hoop (eds.), *Linguistics in the Netherlands 1997*, 61–72. Amsterdam: John Benjamins. DOI: 10.1075/avt.14.08dik.

Dimitrova-Vulchanova, Mila & Giuliana Giusti. 1995. Quantified noun phrase structure in Bulgarian. 5(1). 43–65.

Ekaterina Georgieva

Dimitrova-Vulchanova, Mila & Giuliana Giusti. 1998. Fragments of Balkan nominal structure. In Artemis Alexiadou & Chris Wilder (eds.), *Possessors, predicates and movement in the Determiner Phrase*, 333–360. Amsterdam & Philadelphia: John Benjamins. DOI: 10.1075/la.22.12dim.

Doetjes, Jenny & Johan Rooryck. 2003. Generalizing over quantitative and qualitative constructions. In Martine Coene & Yves D'hulst (eds.), *From NP to DP, volume 1: The syntax and semantics of noun phrases*, 277–295. Amsterdam & Philadelphia: John Benjamins. DOI: 10.1075/la.55.13doe.

Dost, Ascander & Vera Gribanova. 2006. Definiteness in the Bulgarian. In Donald Baumer, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 132–140. Somerville, MA: Cascadilla Proceedings Project.

Embick, David & Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32(4). 555–595. DOI: 10.1162/002438901753373005.

Franks, Steven. 2001. The internal structure of Slavic NPs, with special reference to Bulgarian. Ms., Indiana University.

Giannakidou, Anastasia & Melita Stavrou. 1999. Nominalization and ellipsis in the Greek DP. *The Linguistic Review* 16(4). 295–332. DOI: 10.1515/tlir.1999.16.4.295.

Halpern, Aaron. 1995. *On the placement and morphology of clitics.* Stanford, CA: CSLI Publications.

Harizanov, Boris. 2014. *On the mapping from syntax to morphophonology.* Santa Cruz, CA: University of California. (Doctoral dissertation).

Harizanov, Boris. 2018. Word formation at the syntax-morphology interface: denominal adjectives in Bulgarian. *Linguistic Inquiry* 49(2). 283–333. DOI: 10.1162/ling_a_00274.

Harizanov, Boris & Vera Gribanova. 2015. How Across-the-Board movement interacts with nominal concord in Bulgarian. In Helena Aparicio, Gallagher Flinn, Kathryn Franich, Joanna Pietraszko & Tamara Vardomskaya (eds.), *Proceedings from the 49th Annual Meeting of the Chicago Linguistics Society.* Chicago, IL: Chicago Linguistics Society.

Hendrick, Randall. 1990. Operator movement within NP. In Aaron Halpern (ed.), *The Proceedings of the Ninth West Coast Conference on Formal Linguistics*, 249–261. Standord, CA: CSLI Publications.

Hulk, Aafke & Christine Tellier. 2000. Mismatches: agreement in qualitative constructions. *Probus* 12(1). 33–66. DOI: 10.1515/prbs.2000.12.1.33.

Ionin, Tania. 2006. *This* is definitely specific: specificity and definiteness in article systems. *Lingua* 14(2). 175–234. DOI: 10.1007/s11050-005-5255-9.

Kayne, Richard. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Kester, Ellen-Petra. 1996. Adjectival inflection and the licensing of empty categories in DP. *Journal of Linguistics* 32(1). 57–78. DOI: 10 . 1017 / S0022226700000761.

Lakoff, Robin. 1974. Remarks on *this* and *that*. In Michael W. La Galy, Robert Allen Fox & Anthony Bruck (eds.), *Proceedings of the Chicago Linguistics Society 10*, 345–356. Chicago: Chicago Linguistics Society.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9780511605789.

Napoli, Donna Jo. 1989. *Predication theory: A case study for indexing theory*. Cambridge: Cambridge University Press. DOI: 10.1017/cbo9780511519963.

Nicolova, Ruselina. 2017. *Bulgarian grammar*. Berlin: Frank & Timme GmbH.

Potts, Christopher & Florian Schwarz. 2010. Affective *this*. *Linguistic Issues in Language Technology* 3(5). 1–30.

Spencer, Andrew & Ana R. Luís. 2012. *Clitics: An introduction*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139033763.

Villalba, Xavier & Anna Bartra-Kaufmann. 2010. Predicate focus fronting in the Spanish determiner phrase. *Lingua* 120(4). 819–849. DOI: 10.1016/j.lingua.2008. 07.010.

Wolter, Lynsey Kay. 2006. *That's that: The semantics and pragmatics of demonstrative noun phrases*. Santa Cruz, CA: University of California. (Doctoral dissertation).

Chapter 7

Clitic climbing without restructuring in Czech and Polish

Irenäus Kulik

Friedrich Schiller University Jena

Clitic climbing, i.e. the realization of one or more clitics in a syntactic constituent hierarchically higher than the clitics' licensing predicate, has been accounted for in terms of a restructuring approach. The embedded infinitive the clitics are extracted from has been assumed to be structurally deficient – that is, a bare VP. Due to the lack of projections above the lexical V-head, clitics escape the infinitival domain to get their morphosyntactic features licensed in the matrix clause. However, the predictions of the restructuring approach do not withstand a corpus linguistic examination and are falsified by empirical data of Czech and Polish. Clitic climbing cannot be adequately accounted for by syntax proper and alternative accounts have to be taken into consideration seriously. It will be proposed to refer information structure as a feasible explanatory account of clitic climbing.

Keywords: clitic climbing, restructuring, syntax, information structure, Czech, Polish

1 Introduction

CLITIC CLIMBING (CC) is the realization of a pronominal or reflexive clitic in a syntactic constituent hierarchically higher than the licensing predicate. Junghanns (2002a: 66) schematizes CC as in (1a), whereby a constituent α embeds a constituent β. CC is analyzed as movement of the CLITIC (CL) from β to α. (1b) paraphrases the scheme in a theory-neutral way.[1] The gap e co-indexed with CL

---

[1]I will attempt to phrase my arguments in a theory-neutral way, in order not to impose a specific approach on the reader. For instance, I will not opt for a particular analysis of long-distance dependencies in terms of e.g. movement or copy-and-delete.

Irenäus Kulik

captures the fact that CL is linearized in α, but subcategorized for by the verbal predicate in β. In principle, an arbitrary number *n* of phrase boundaries may intervene between α and β (with {$\mathbb{N}_0$} denoting the set of natural numbers including zero).

(1) a. $[_\alpha \dots \text{CL} \dots [_\beta \dots t_{\text{CL}} \dots]]$
     b. $[_\alpha \text{CL}_j [_n [_\beta e_j]]], n \in \{\mathbb{N}_0\}$

There is an extensive body of research literature on CC in Romance, which significantly inspired research on Slavic. In the Italian example (2a), the direct object clitic *lo* 'him.ACC' follows the embedded infinitive *legger(e)* 'read' it is argument of.[2] This is the local or in situ realization of the clitic. CC is found in (2b) with the object clitic being realized before the finite verb of the matrix phrase, which will be also referred to as non-local placement.

(2) a. Martina   vuole$_1$ legger$_2$=*lo*$_2$.   (in situ /local)
         Martina.NOM want.PRS.3SG read.INF=him.ACC
         'Martina wants to read it.'
     b. Martina   *lo*$_2$ vuole$_1$ leggere$_2$.   (CC /non-local)
         Martina.NOM him.ACC want.PRS.3SG read.INF
         'Martina wants to read it.'   (Italian; Spencer & Luís 2012: 163–164)

The remainder of the paper is organized as follows: §2 briefly comments on the syntactic status of clitic pronouns in West Slavic. §3 provides a concise overview of basic clitic climbing properties in West Slavic. §4 is the core of the paper and tests the correlates of restructuring empirically focusing on accusative case licensing §4.1, the absence of an underlying subject in the infinitive phrase §4.2, the dependence of the infinitive's temporal reference upon the matrix verb's tense information §4.3, and the all-or-nothing quality of clitic climbing §4.4. §5 addresses the role of information structure for clitic climbing. Concluding remarks are given in §6.

2 The status of West Slavic clitics

It is common to distinguish Polish and Czech clitics along the lines of Zwicky's (1977) simple/special-clitic dichotomy. The second position clitics in Czech are

---

[2]Clitics will be highlighted in italics for ease of reference. I adopt the integer-index-convention from Hana (2007), Rosen (2014), and Kolaković et al. (2022) to indicate the structural hierarchy between the verbal heads as well as the subcategorization relations between a verbal predicate and its dependents.

considered special clitics, whereas Polish clitics being distributed rather freely are simple clitics. Given this distinction, cliticization in these two closely-related languages is an ideal test field for theorizing about clitic phenomena and the micro-typology of Slavic cliticization. However, the often-made statement that Polish clitics are typologically peculiar in comparison to clitics in other West and South Slavic languages, turns out to be controversial on closer inspection. On the one hand, Rappaport (1988), Dziwirek (1998), Kupść (2000), Borsley & Rivero (1994), Franks (2009, 2010), and Franks & King (2000) treat Polish as a language without second position clitics – hence, not possessing special clitics. On the other hand, Rothstein (1993: 725), Urbańczyk (1976: 62), Veselovská (1995: §4.8 Footnote 23), and Dimitrova-Vulchanova (1999: 85) consider Polish to be a second position clitic language essentially. Spencer (1991: 390) regards Polish as a special clitic language. From a different angle, Czech exhibits positional deviations from second position cliticization with third, fourth, and fifth position placement being attested (see Hana 2007: 103–112, Junghanns 2021: 177–178). Therefore, Czech and Polish clitics will be treated alike throughout the paper. I restrict myself to the set of short pronominals adapting a traditional terminology from Slavic studies here.[3] The respective sets for Czech and Polish are provided in Table 1 (see Fried 1994, Avgustinova & Oliva 1997, Rosen 2001, Junghanns 2002b, Petkevič 2009 for Czech, Kupść 2000 for Polish).[4] I follow the spirit of Dotlačil (2007) and – most recently – Adam (2019) in refraining from hypothesizing about the exact syntactic status of the short pronouns, e.g. whether they are syntactic phrases or heads, or whether they are weak rather than clitic pronouns.[5]

---

[3]The notion rests upon the formal distinction of "short" (e.g. Cz. *mu*, *ho*) vs. "long" pronouns (e.g. Cz. *jemu*, *jeho*). Note that terminology differs between authors. The short pronouns are referred to as *konstantní příklonky* 'constant clitics' in Czech linguistics (see Trávníček 1959, Rosen 2001, Hana 2007), Avgustinova & Oliva (1997) propose the term *pure clitics*, and Junghanns (2002b) coins *lexikalische Klitika* 'lexical clitics'.

[4]Short dative reflexive *se* occurs in colloquial Polish, but remains unconsidered in most analyses (e.g. Spencer 1991, Kupść 2000). Rubadeau (1996: 137) claims that "Polish [...] does not have a clitic form of the dative reflexive". On the other hand, Urbańczyk (1976: 58) discusses *se* in his outline of Polish dialects. Franks & King (2000: 150) list *se* among the Polish clitics, but note that it "is used only in the spoken language" (cf. also Rothstein 1993: 702). Aguado & Dogil (1989) explicitly take *se* into consideration.

[5]An anonymous reviewer pointed out that Czech short pronouns are true second position clitics, whereas Polish short pronouns are weak pronouns. This point of view is reminiscent of Cardinaletti & Starke's (1999) tripartite typology of pronouns. Since there is no general consensus on this matter, it appears that the typology of clitics in Slavic still needs further investigation. For an alternative view, see Jung & Migdalski (2022), who who propose an extension of Cardinaletti & Starke's (1999) approach to a four-way classification.

Irenäus Kulik

Table 1: Unambiguous short pronominals in Czech and Polish

English	Czech	Polish
you$_{ACC}$	*tě*	*cię*
him$_{ACC}$	*ho*	*go*
me$_{DAT}$	*mi*	*mi*
you$_{DAT}$	*ti*	*ci*
him$_{DAT}$	*mu*	*mu*
refl$_{ACC/GEN}$	*se*	*się*
refl$_{DAT}$	*si*	*se*

3 Clitic climbing in West Slavic

As in the Italian example (2), CC occurs from embedded infinitives in Czech and Polish.[6] Several scholars point out that infinitive-hood is a necessary, but not a sufficient condition for CC (cf. Junghanns 2002a: 69 on Czech, Kupść 2000: 58 on Polish, Golden 2003: 221–222 on Slovene). The infinitival domain must not be introduced by a SUBORDINATOR.[7] Note that the clitic, which is subject to climbing, is not necessarily an argument, e.g. REFL of a reflexive tantum or in impersonal constructions. Therefore, I adopt the term DEPENDENT from dependency grammar as a general notion for clitics licensed by a verbal head. CC occurs in a variety of syntactic constructions, i.e. raising, subject and object control, and the ACCUSATIVE WITH INFINITIVE (ACI, from Latin *accusativus cum infinitivo*) (see Junghanns 2002a, Golden 2008, Kupść 2000).[8] Note that clitic climbing is ungrammatical in object control constructions in Romance, but not in Slavic (cf. Golden 2008: 315). Note also that standard Polish and its vernacular do not possess the ACI construction (see Przepiórkowski & Rosen 2005: 33, Kupść 2000: 96). The lack of ACI is a general property of Polish syntax, it is not a particular feature of the Polish clitic system. The ACI is attested in several diatopic varieties

---

[6]An anonymous reviewer pointed out that CC is possible from a subset of morphologically finite *da*-clauses in Serbian, i.e. from subjunctive-like *da*-clauses (see Progovac 1993, 1996 for the relevant distinction of indicative und subjunctive *da*-clauses). It is not necessary to rely on the indicative-subjunctive distinction to account for CC in West Slavic, which is best captured by the conditions of the embedded verb's infinitive-hood and the absence of a subordinator.

[7]The term SUBORDINATOR is meant to broadly cover elements introducing subordinate clauses of different kinds, i.e. (i) complementizers introducing argument clauses, (ii) subordinate conjunctions introducing adjunct clauses, and (iii) relative pronouns and adverbs introducing relative clauses.

[8]The accusative with infinitive is known as exceptional case marking in generative grammar.

of Polish (see Urbańczyk 1976: 56). It has been observed that CC is not obligatory and clitics may be realized in situ as in Italian (2a). In the same way, both a- and b-examples are grammatical in Czech (3) and Polish (4). The question arises then, why CC does come into being and what are the conditions for the local vs. non-local realization of the clitics.

(3) a. Asi ho_2 chtěla$_1$ usušit$_2$ pomalu.
perhaps him.ACC want.PST.SG.F dry.INF slowly
'Perhaps she wanted to dry it slowly.'

   b. Asi chtěla$_1$ usušit$_2$ ho_2 pomalu.
perhaps want.PST.SG.F dry.INF him.ACC slowly
'Perhaps she wanted to dry it slowly.' (Czech; Junghanns 2002a: 82)

(4) a. Jan go_2 chciał$_1$ obudzić$_2$ o szóstej.
Jan.NOM him.ACC want.PST.SG.M wake-up at six
'Jan wanted to wake him up at six o'clock.'

   b. Jan chciał$_1$ obudzić$_2$ go_2 o szóstej.
Jan.NOM want.PST.SG.M wake-up him.ACC at six
'Jan wanted to wake him up at six o'clock.' (Polish; Kupść 2000: 60)

It has been proposed to account for CC in Slavic in terms of a restructuring approach by Rezac (2005) for Czech and Aljović (2004) for Bosnian-Croatian-Montenegrin-Serbian (BCMS) (cf. Rizzi 1982 on restructuring in Italian and Wurmbrand 2001 for a general analysis of restructuring properties on the basis of German and Japanese). According to such an approach the optionality of CC is only an alleged one. While clitics must remain in situ in true bi-clausal structures, they are forced to climb under restructuring, which is underlyingly mono-clausal due to the structural deficiency of the embedded so-called RESTRUCTURING INFINITIVE (RI). Being bare VPs, RIs lack the *v*P- and TP-shell.[9] Several correlates have been put forward to support the restructuring analysis: (a) RIs are unable to license accusative case, (b) RIs do not have an underlying subject (= PRO), (c) RIs do not constitute a binding domain for principle B, (d) either all clitics climb as a consequence of the infinitive's structural deficiency or none, (e) RIs are temporally dependent upon the matrix verb's tense. Criteria (a)–(d) are taken from Rezac (2005), criterion (e) is taken from Todorović (2012).

In what follows I will test the hypothesis that CC is dependent upon restructuring by assessing the above-mentioned correlates empirically towards corpus

---

[9]Note that Wurmbrand (2001 and subsequent work) proposes a multi-way distinction of restructuring, which is not limited to binary parametrization. I cannot take these proposals into consideration here due to space limitations and must leave them for future discussion.

Irenäus Kulik

data from the *Český národní korpus* 'Czech national corpus' (ČNK) and the *Narodowy Korpus Języka Polskiego* 'National corpus of Polish' (NKJP) respectively. In particular, the Czech data are drawn from the subcorpus SYN version 8 (see Křen et al. 2019, Hnátková et al. 2014).[10] For Polish, I searched the full NKJP corpus through the Poliqarp search engine (see Przepiórkowski et al. 2012).[11]

4 Clitic climbing and correlates of restructuring

4.1 Case licensing

Due to the lack of vP/TP, RIs are unable to license accusative case. Clitics climb in order to receive case in the matrix phrase then.[12] Lenertová (2004) and Dotlačil (2004) recognize independently for Czech that CC into passivized matrix domains contradicts this argument. It is generally known from Burzio's generalization that passivized verbs are unable to license accusative case (cf. Burzio 1986). Data like (5) and (6) contradict the case-based argument, as the case of the clitic *ho* 'him.ACC' cannot be licensed by the passivized matrix verb, but only by the embedded infinitive. Lenertová's (2004) and Dotlačil's (2004) arguments are corroborated by examples (7) and (8) for Czech and Polish respectively.

(5) '(Přivezl   puk za   švýcarskou branku,) ale tam *ho₃*
    bring.PST.SG.M puck behind Swiss   goal   but there him.ACC
    byl₁   donucen₂   předat₃ Lubinovi.
    be.PST.SG.M forced.PASS.SG.M give.INF Lubin.DAT
    '(He brought the puck behind the Swiss goal,) but there he was forced to
    give it to Lubina.'   (Czech; Lenertová 2004: 159)

(6) Pavel   *ho₃*   byl₁   nucen₂   zničit₃.
    Pavel.NOM him.ACC be.PST.SG.M force.PASS.SG.M destroy.INF
    'Pavel was forced to destroy it.'   (Czech; Dotlačil 2004: 88)

(7) a. […], kdo by *ho₃*   byl₁   oprávněn₂   zbavit₃   z
           who COND him.ACC be.PST.SG.M entitle.PASS.SG.M relieve.INF of
        odpovědnosti za osud Ruska.
        responsibility for fate Russia.GEN.SG
        '…, who would have been entitled to relieve him of his responsibility
        for the fate of Russia.'

---

[10]https://www.korpus.cz/

[11]http://www.nkjp.pl/

[12]It is irrelevant for the purpose of the present study, how case licensing is technically implemented.

   b.   […], kteří *ho*$_3$ byli$_1$ připraveni$_2$ zatknout$_3$.

         who him.ACC be.PST.PL.MAN prepare.PASS.PL.MAN arrest.INF

        '…, who were prepared to arrest him.'   (Czech; ČNK)

(8) a.   […] że już   nigdy nie będę$_1$ *cię*$_3$ zmuszona$_2$

          that already never NEG be.FUT.1SG you.ACC force.PASS.SG.F

          oglądać$_3$.

          look.INF

          '… that I will never be forced to look at you, again.'

    b.   bo   z   powodu drżenia twoich rąk   będę$_1$ *cię*$_3$

         because from reason tremor your   hands be.FUT.1SG you.ACC

         zmuszony$_2$ wrzucić$_3$ do KF

         force.PASS.SG.M throw.INF to KF

         'because of your hands' tremor I will be forced to throw you to the

         KF [= kill file]'   (Polish; NKJP)

4.2 Missing subjects

The lack of *v*P yields RIs without having an underlying subject (PRO). Rezac (2005: 114) states that RIs "will not constitute a binding domain of their own, and coreference between a pronominal argument of the infinitive and any argument of the upstairs verb should be blocked." He provides the minimal pair in (9a)–(9b).[13] In (9a) the embedded clitic *ji* 'her.ACC' is co-referential with both the matrix subject *Anna* (index *a*) and a distinct discourse referent beyond the sentence-level (index *b*). In (9b) the clitic has climbed due to restructuring. As a consequence, there is no clause boundary between the matrix and subordinate domain, thus, co-reference between *Anna* and *ji* is excluded. Rezac (2005) accounts for (9b) by a violation of binding principle B, according to which "[a] pronominal is free [i.e. unbound] in its governing category [i.e. clause]" (Chomsky 1981: 188). However, principle B is inconclusive. It determines semantic co-reference by syntactic non-co-membership, which does not reveal anything about clause boundaries here. Co-reference between *Anna* and *ji* is still excluded by principle B in presence of a clause boundary, for both subject and clitic are co-members of the matrix domain, cf. (9c). Note, furthermore, that the matrix object *mu* 'him.DAT' still co-refers with the kisser of the embedded kissing-event despite CC of *ji*.

---

[13]An anonymous reviewer pointed out that *políbit ji nashledanou* is an uncommon calque on the basis of English 'to kiss someone goodbye'. This does not have an impact on CC, however.

Irenäus Kulik

(9) a. [Anna$_a$ *mu$_c$* dovolila [PRO$_c$ políbit *ji$_{a/b}$*
         Anna.NOM him.DAT allow.PST.SG.F kiss.INF her.ACC
         nashledanou]].
         good-bye
         'Ana permitted him to kiss her good-bye.'

     b. [Anna$_a$ *mu* *ji*$_{*a/b}$ dovolila políbit nashledanou].
         Anna.NOM him.DAT her.ACC allow.PST.SG.F kiss.INF good-bye
         'Ana permitted him to kiss her good-bye.'

     c. [Anna$_a$ *mu$_c$* *ji*$_{*a/b}$ dovolila [(PRO$_c$) políbit
         Anna.NOM him.DAT her.ACC allow.PST.SG.F kiss.INF
         nashledanou]].
         good-bye
         'Ana permitted him to kiss her good-bye.'   (Czech; Rezac 2005: 114)

Rezac (2005: 114–115) further states that neither matrix argument binds subject-oriented anaphora *svým* 'one's.POSS.PL.DAT' in (10b) in contrast to (10a). As the clitic *je* 'them.ACC' has climbed, restructuring must have occurred and PRO is missing thence. However, co-reference between the matrix subject and the anaphorical possessive pronoun should be still expected in a restructuring context. In fact, Dotlačil (2007) and Skoumalová (2005) judge (10b) grammatical with both interpretations, such that embedded *svým* is bound by either matrix argument (*Pavel, Janovi*) despite CC. These judgements are corroborated by the corpus data in (11) and (12). First, the matrix subject *stavitel* 'constructor' binds the possessive anaphor *své* 'one.POSS' after CC in (11) as expected. Second, and even more intriguing, example (12) shows that the matrix object clitic *mu* 'him.DAT' binds the embedded possessive anaphor *své* in spite of the climbed embedded clitic *ho* 'him.ACC'. While the binding relations in (11) are expected under standard assumptions in any mono-clausal domain, the binding facts in (12) are best analyzed by assuming an underlying subject in the embedded infinitive (i.e. PRO under standard generative assumptions).[14]

(10) a. Pavel$_a$ přikázal$_1$ Janovi$_b$ dát$_2$ *je$_2$* svým$_{a/b}$
         Pavel.NOM order.PST.SG.M Jan.DAT give.INF them POSS
         přátelům.
         friends.DAT
         'Pavel$_a$ ordered Jan$_b$ to give them to his$_{a/b}$ friends.'

---

[14]I do not adopt Hornstein's (1999) proposal in abandoning the raising-control distinction, which has originally been the main motivation for the assumption of PRO (cf. Przepiórkowski & Rosen 2005 for a similar account in HPSG and Culicover & Jackendoff 2001, Landau 2003 for a critique).

188

b. * Pavel_a   *je_2*   Janovi_b přikázal_1   dát_2   svým_{a/b}
Pavel.NOM them Jan.DAT order.PST.SG.M give.INF POSS
přátelům.
friends.DAT
Intended: 'Pavel_a ordered Jan_b to give them to his_{a/b} friends.'

(Czech; Rezac 2005: 114–115)

(11) Stavitel_a   *mu_2*   nechtěl_1   vnucovat_2 své_a mínění.
constructor.NOM him.DAT NEG.want.PST.SG.M impose.INF POSS opinion
[...]
'The constructor didn't want to impose his opinion on him ...'

(Czech; ČNK)

(12) (A právě [_NP "ten   myš"]_a se sourozenci_b zalíbil
and exactly   this.SG.M mouse REFL sibling.SG.DAT please.PST.SG.M
natolik, že mě   požádal,) abych *mu_{1/b}*   *ho_{2/a}*
so-much that me.ACC ask.PST.SG.M so-that him.DAT him.ACC
dovolil_1   použít_2 v jedné své_b písničce.
allow.PST.SG.M use.INF in one   POSS song
'(And [my] sibling liked exactly "this he-mouse" so much that he asked
me,) if I would allow him to use it in one of his songs.'   (Czech; ČNK)

Another argument that challenges the predicted binding correlations of restructuring has been put forth by Golden (2008: 316) for Slovene. She observed that certain object control constructions are semantically ambiguous, although the embedded clitic has climbed (13). That is, an object may be interpreted as either being subcategorized for by the matrix verb, whereby the object controls the embedded PRO-subject, or by the embedded infinitive and no control occurs. The ambiguity remains in case of CC, although PRO should be absent due to restructuring, such that the control reading should not be available. Correspondingly, the ambiguity of Czech (14) and Polish (15) calls for an analogue of a PRO-analysis for the infinitive.[15]

(13) a. Janez   *ji_{1/2}*   *jih_2*   je   dovolil_1   kupiti_2.
Janez.NOM her.DAT them.ACC AUX.3SG. allow.PST.SG.M buy.INF
(i) 'Janez allowed her to buy them.'
(ii) 'Janez allowed (someone) to buy them/it for her/them.'

(Slovene; Golden 2008: 316)

---

[15]An anonymous reviewer pointed out that a PRO-less analysis is available following work by Gennaro Chierchia (see Chierchia 1984). As the consequences of this approach are not clear to me at this moment, I will leave it for future research.

Irenäus Kulik

b. Jaz *sem* *ji*$_{1/2}$ *ga*$_2$ dovolil$_1$ poslati$_2$ po pošti.
   I.NOM AUX.1SG. her.DAT him.ACC allow.PST.SG.M send.INF by mail
   (i) 'I allowed her to send it by mail.'
   (ii) 'I allowed (somebody) to send it to her by mail.'

(Slovene; Golden 2008: 312)

(14) […], strýc *mu*$_{1/2}$ *ho*$_2$ nedovolí$_1$ přečíst$_2$.
     uncle him.DAT him.ACC NEG.allow.PRS.3SG read.INF
     (i) 'The uncle doesn't allow him to read it.'
     (ii) 'The uncle doesn't allow (someone) to read it to him.' (Czech; ČNK)

(15) […] każą$_1$ *mu*$_{1/2}$ *go*$_2$ rozebrać$_2$.
     order.PRS.3PL him.DAT him.ACC deconstruct.INF
     (i) '… they order him to deconstruct it.'
     (ii) '… they order (someone) to deconstruct it for him.' (Polish; NKJP)

4.3 Temporal reference

It has been argued that the RI's temporal reference is dependent upon the one presupposed by the matrix verb. RIs are ungrammatical with a temporal adverb, which refers to a time frame deviating from the matrix verb's one. Wurmbrand's (2001) German example (16a) provides a grammatical utterance without restructuring. The main verb encodes the past tense (morphosyntactically encoded by the analytical perfect form), but the embedded infinitive refers to the future by the time adverb *morgen* 'tomorrow'. On the other hand, the presence of the time adverb is ungrammatical in a restructuring context like (16b). Example (17) from Aljović (2004) suggests that the same holds for Slavic. The presence of the time adverb *sutra* 'tomorrow' in the embedded clause is grammatical in the BCMS example (17a), as long as the pronominal clitic *ga* 'him.ACC' is in situ. When restructuring occurs and the clitic climbs, then the realization of the time adverb yields the utterance ungrammatical (17b).

(16) a. Hans hat beschlossen (morgen) zu verreisen.
        Hans have.PRS.3SG decide.PTCP tomorrow to travel.INF
        'John decided to go on a trip (tomorrow).'
     b. Hans hat versucht (*morgen) zu verreisen.
        Hans have.PRS.3SG try.PTCP tomorrow to travel.INF
        'John tried to go on a trip.'

(German; Wurmbrand 2001: 73)

(17) a. On želi₁   da *ga₂*   (sutra)   Jovanu
          he want.PRS.3SG that him.ACC tomorrow Jovan.DAT
          predstavi₂.
          introduce.PRS.3SG
          'He wants to introduce him to John tomorrow.'

    b. *On *ga₂*   želi₁   da (sutra)   Jovanu
          he him.ACC want.PRS.3SG that tomorrow Jovan.DAT
          predstavi₂.
          introduce.PRS.3SG
          Intended: 'He wants to introduce him to John tomorrow.'

                                                                     (BCMS; Aljović 2004: 193)

Lenertová (2004) notes that the aforementioned argument does not hold for Czech, where CC co-occurs with the embedded infinitive's independent temporal reference. The clitic *ho* 'him.ACC' in (18) has climbed to the matrix domain headed by the past tense verb *rozhodl* 'decide.PST.SG.M'. However, the realization of the temporal adverb *příště* 'next time' or adverbial PP *na moment* 'for a moment' within the infinitive's domain is grammatical. Lenertová's observation is corroborated for Czech (19) and Polish (20) by corpus data. Note that climbing is grammatical irrespective of whether the time adverb(ial)s (ADV) intervene between matrix verb and embedded infinitive or not. This fits Junghanns's (2002a: 66) observation that the cascade of verbs, which constitutes an environment for CC, does not form a verb cluster (Germ. *Verb[al]komplex*) in Czech, i.e. they do not need to be contiguous (cf. also Golden 2008: 313).

(18) Místo   toho *se₁* *ho₂*   rozhodl₁   [_ADV na moment] /
      instead-of this REFL him.ACC decide.PST.SG.M   on moment
      [_ADV příště]   ignorovat₂.
          next-time ignore.INF
      'Instead, he decided to ignore him for a moment / next time.'

                                                        (Czech; Lenertová 2004: 157)

(19) a. Přitom   *ho₂*   chtěla₁   odstartovat₂ [_ADV příští sobotu]
          but-in-fact him.ACC want.PST.SG.F launch.INF   next Saturday
          při příležitosti oslav   700 let   od udělení   městských
          at occasion   celebration 700 years from awarding city
          práv Sokolovu.
          rights Sokolov
          'But in fact, [the town's administration] wanted to launch it on the

occasion of the 700th anniversary of Sokolov receiving its town
charter.'

b. Lidé,   kteří *se*₂ chtěli₁   [_ADV zítra   večer]
people who REFL want.PST.PL.MAN   tomorrow evening
bavit₂   při filmu Borat mají   smůlu.
entertain.INF at film   Borat have.PRS.3PL bad-luck
'Those people, who wanted to enjoy the Borat movie tomorrow
evening, have bad luck.'   (Czech; ČNK)

(20) a. Ja   *się*₂ postanowił-em   nie podrapać₂ [_ADV jutro
I.NOM REFL decide.PST.SG.M.M.1SG. NEG scratch.INF   tomorrow
o 12.15] […].
at 12.15
'I decided not to scratch myself tomorrow at 12:15 … .'

b. ja   *mu*₂ zdecydowałem₁   *się*₁ odpowiadać₂ [o
I.NOM him.DAT decide.PST.SG.M.M.1SG. REFL reply.INF   [at
ile   na jakieś posty będzie   warto] [_ADV po   24
how-much on some posts be.FUT.3SG worth]   after 24
godzinach]
hours.
'I decided to respond to him [as far as some posts will be worth it]
after 24 hours.'   (Polish; NKJP)

4.4 All or nothing

CC has been deemed an "all-or-nothing phenomenon" (Rezac 2005: 111), whereby
either all embedded clitics climb or none (see Aljović 2004: 194 for a similar po-
sition regarding BCMS). Due to the RI's structural deficiency, the clitics escape
the infinitival domain to satisfy their formal requirements in the matrix phrase,
where they are placed in the respective clitic cluster. DIACLISIS of co-dependents
poses a problem for such an approach then.[16] In the empirically attested Ser-
bian example (21), both the pronominal clitic *mi* 'me.DAT' and the reflexive clitic
se 'REFL' are subcategorized for by the embedded verb *vrti* 'spin.PRS.3SG' of the
da-clause.[17] However, it is only the pronominal clitic that climbs, the reflexive

---

[16]I adopt the term DIACLISIS for split-clitic-constructions from Kolaković et al. (2022: 34), who
took the notion from Janse's (1998) discussion of clitics in Cappadocian Greek.

[17]An anonymous reviewer pointed out that the availability of diaclisis in BCMS was already
recognized by Sandra Stjepanović (see Stjepanović 1998, 1999, 2004). The data remained con-
troversial, as the positive judgement of diaclisis has not been generally accepted (cf. Aljović
2004: 192 Footnote 3, Franks & King 2000: 335).

remains in situ (see Kolaković et al. 2022: 307). As restructuring is supposed to affect all embedded clitics equally, the approach is unable to predict differences in the distribution of co-dependent clitics. This is also true for diaclisis in Czech (22) and Polish (23).[18] Again, both reflexive and pronominal clitics are co-dependents of the embedded infinitives. Only the reflexive clitic occupies the clausal second position, whereas the pronominal clitic appear further to the right and does not build a cluster with the reflexive. Uwe Junghanns (p.c.) pointed out that it is impossible to determine a priori, whether the pronominal clitics in (22)–(23) have in fact climbed or whether they are still positioned within the infinitival phrase (see also Junghanns 2002a: 67–68).[19] While this behavior would appear unsurprising for Polish given the rather peculiar status its clitic system is assigned, the occurrence of the same pattern in Czech is unexpected. The Czech data are problematic for Bošković's (2001) PF-filtering approach to clitic clustering, according to which clitics in a second position clitic language are placed according to two parameters: first, initial positioning in an intonation phrase (ιP), second, being suffixed to a prop. Bošković accounts for diaclisis in Polish by assuming that Polish clitics do not possess the second position requirement of being ιP-initial.[20] Note that the diaclitic distribution is independent of the argument status of the

---

[18]An anonymous reviewer suggested that an all-or-nothing analysis might be available for Czech, if one assumes a verb-adjacent placement pattern with both pronouns leaning on the matrix verb. Such an approach is debatable. It remains unclear, why Czech special clitics do not build a cluster (cf. also Footnote 20). Bulgarian and Macedonian clitics are verb-adjacent and obey clitic clustering (but do not have clitic climbing, cf. Franks & King 2000: 241). In the scenario suggested, the reflexive needs to procliticize, while the pronominal encliticizes on the same host. Thus, the prosodic orientation appears to be rather arbitrary, contra Toman (1996), who argues that, while short pronouns encliticize by default in Czech, procliticization is a function of the phonological environment, i.e. the lack of a prop due to a prosodic break (see also Junghanns 2021: 178–179).

[19]The contrast is schematised in (i) and (ii).

    (i) [$_\alpha$ V $_\alpha$ CL$_\beta$ [$_\beta$ V $_\beta$]] (CC/non-local)

    (ii) [$_\alpha$ V$_\alpha$ [$_\beta$ CL$_\beta$ V $_\beta$]] (in situ/local)

Proclisis to the following infinitive would be indicative of clitic in situ placement. Proclisis is available in Czech (see Toman 1996) and Polish (see Kraska-Szlenk 1995: 62–64) for the set of clitics relevant here.

[20]The Czech facts cannot be captured, since the approach predicts that split clitics are placed in distinct ιPs. However, there is only one ιP for the relevant clause in Czech (22), cf. (i)–(ii). I thank Martina Berrocal (p.c.) for the judgement (# marks a pause).

    (i) [$_{\iota P}$ Lehce *si* (*#) uměla *ho* představit]

    (ii) [$_{\iota P}$ kdekdo *se* (*#) začal *mi* smáti]

clitics. The reflexives in the a-examples are true arguments of the embedded infinitives, whereas they are not in the b-examples. Both Cz. *smát(i) se* 'laugh' and Pol. *bać się* 'be afraid' are reflexiva tantum, i.e. the appearance of the reflexive is lexically specified. This finding supplements Lenertová's (2004: 138–139) observation that Czech clitics do not need to cluster together, as she found that conditional and reflexive/pronominal clitics may occur non-contiguously.

(21) [...] i   počelo$_1$ *mi$_2$*   *je*   [da *se$_2$* vrti$_2$ u glavi].
      and start.PST.SG.N me.DAT AUX.3SG that REFL spin.PRS.3SG in head
      '... and I started to feel dizzy.'   (Serbian; Kolaković et al. 2022: 307)

(22) a. Lehce *si$_2$* uměla$_1$ *ho$_2$*   představit$_2$, [...]
          easily REFL be-able.PST.SG.F him.ACC imagine.INF
          'She could easily imagine him, ...'

      b. [...] kdekdo   *se$_2$* začal$_1$ *mi$_2$*   smáti$_2$.
          almost-everybody REFL start.PST.SG.M me.DAT laugh.INF
          '... almost everybody started to laugh at me.'   (Czech; ČNK)

(23) a. My   *się$_2$* musimy$_1$ *go$_2$*   nauczyć$_2$.
          we.NOM REFL must.PRS.1SG him.ACC teach.INF
          'We have to learn it.'

      b. Już   *się$_2$* zaczęli$_1$ *go$_2$*   bać$_2$, [...]
          already REFL start.PST.PL.MAP him.ACC fear.INF
          'They already started to be afraid of him, ...'   (Polish; NKJP)

5 Clitic climbing and information structure

The previous sections showed that a purely syntactic account in terms of restructuring cannot cope with CC in Czech and Polish. The corpus data provided in §4 contradict the predictions of the approach. Therefore, I agree with Dotlačil (2004: 87) in that CC does not occur because of restructuring and that both should be regarded as independent phenomena.

If restructuring is not responsible for CC, then the question arises, what is. Which alternatives are available, if one refrains from accounting for CC by syntax proper? Proposals in terms of phonology, morphology, and syntax-prosody interaction have been put forward, but were also criticized. For instance, Franks & King (2000: 287–291, 293–305) critically review purely phonological and purely syntactic accounts of cliticization (incl. CC) and conclude that both types of approach face several problems in accounting for clitic phenomena (cf. also the

discussion in Bošković 2001: 36–80). However, I do not pursue a mixed syntax-PF account like Franks & King's (2000: §11–12) and Franks's (2010) PF-filtering approach. Another mixed account is Halpern's (1995) Prosodic Inversion, which is critically reviewed in detail by Bošković (2001: 11–36). On the other hand, Bošković's (2001) own intonational-phrase-based proposal has been criticized by Lenertová (2004: 150–151) and Golden (2008: passim) (see also Footnote 20). An explanatory account of CC has to shed light on the actuation or causation of CC vs. clitic in situ positioning.[21] I propose to take a candidate into consideration that repeatedly appears in the literature on clitics, but has been mostly neglected: information structure. Stjepanović (2004: 206) considers the possibility that CC is an instance of object shift, which has been reported to rely on information structural notions in Northern Germanic: only objects having background status are shifted.

Junghanns (2002a: 82–83) takes the farthest step towards information structure I am aware of and proposes that information structure is the actual reason for CC, whereas syntax merely restricts which domains clitics can escape. Consequently, a clitic climbs, iff it belongs to the background of the whole sentence, else it remains in situ. The else-case covers utterances in which the clitic is part of a topicalized or focused constituent. The clitic itself does not bear topic or focus, but is an element of a domain specified as either [+Topic] or [+Focus].

Accounting for the ban of CC across CP, Dotlačil (2004: 93, 98; 2007: 89) suggests that clitics cannot escape CPs, because they cannot bear the discourse functions of topic or focus. He observes that topicalized or focused constituents are able to escape CPs in Czech. If clitics are hosted in such a topicalized or focused domain, they can cross a CP as a part of the respective constituent. However, no CC occurs, as the clitics remain in their licensing domain.

I adopt Junghanns's (2002a) proposal and paraphrase it tentatively with the notation in (24), which reads as follows: for every *x*, iff *x* has property CL (= is a clitic) and *x* is element of the information structural background, then *x* is realized in domain α and co-indexed with a gap *e* in domain β. Implication (24b) specifies the else-case with the clitic in situ realization. Note that the asymmetry between α and β is purely syntactic and refers to dominance or hierarchic order, but does not tell us anything about linear order.

(24) a. $\forall x(\text{CL}(x) \wedge x \in [\textsc{Background}] \rightarrow [_\alpha x_j [_\beta e_j]])$
      b. $\forall x(\text{CL}(x) \wedge x \in \{[\textsc{Topic}], [\textsc{Focus}]\} \rightarrow [_\alpha ... [_\beta x]])$

---

[21]I refer to the traditional notion of explanation based on causality, not to Chomsky's (1965: 25–26) concept of explanatory adequacy. Instead, I allude to the actuation problem coined by Weinreich et al. (1968) towards the background of historical linguistics and adopt it for the field of synchronic grammar research.

Irenäus Kulik

6 Conclusion

The paper empirically tested the hypothesis that clitic climbing in Czech and Polish is contingent upon a mono-clausal restructuring environment. In particular, I reviewed the predicted correlates of the proposal that clitics escape defective infinitival complements, which are bare VPs. Utilizing data from both the Czech National Corpus and the National Corpus of Polish, it has been shown that (i) clitics climb to passivized domains, i.e. not for accusative case licensing, (ii) binding phenomena and ambiguities in climbing constructions call for an underlying subject (PRO) analysis of the infinitival domain, (iii) embedded infinitives possess temporal reference independent of the finite matrix verb, (iv) co-dependent clitics do not behave uniformly with respect to climbing and end up non-contiguously. In sum, the respective predictions of the restructuring approach have been falsified. Clitic climbing thus cannot be regarded as being restricted to mono-clausal structures, but occurs in what is considered an underlyingly bi-clausal structure. This state of affairs yields the approach uneligible for clitic climbing in Czech and Polish. More generally, as syntax proper does not provide us with an explanatory account for the very existence of clitic climbing, alternatives have to be taken into consideration seriously. Following Junghanns (2002a), I referred to information structure, whereby clitics climb, if they are elements of the background of the entire sentence, but remain in situ, if they are elements of a topicalized or focused constituent. Admittedly, the present study neither addresses how to account for diaclisis in terms of information structure nor how topic and focus domains are determined in order to capture clitic in situ positioning. This has to be dealt with in future research. What is more, the resemblance between Czech and Polish clitic distributions suggests that the typological peculiarity of Polish is not well-grounded. I propose to revisit and refine the micro-typology of Slavic cliticization on a sound empirical basis.

Abbreviations

1	first person	N	neuter
2	second person	NEG	negation
3	third person	NKJP	National Corpus of Polish
ACC	accusative	NOM	nominative
ACI	accusative with infinitive	PASS	passive
ADV	adverb(ial)	MAN	masculine-animate
AUX	auxiliary	MAP	masculine-personal
CC	clitic climbing	PL	plural
ČNK	Czech National Corpus	POSS	possessive
DAT	dative	PRS	present tense
F	feminine	PST	past tense
FUT	future tense	PTCP	participle
GEN	genitive	REFL	reflexive
INF	infinitive	RI	restructuring infinitive
M	masculine	SG	singular

Acknowledgments

References

Adam, Nina. 2019. Clitics at the interfaces of grammar: Defining the "first position" in Czech. *Typology of Morphosyntactic Parameters* 2(1). 11–27. https://tmp.sc/application/files/9816/6793/2318/Adam-2019-2-1.pdf.

Irenäus Kulik

Aguado, Miquel & Grzegorz Dogil. 1989. Clitics in lexical phonology: Alleged counterevidence? *Linguistische Berichte* 120. 99–116.

Aljović, Nadira. 2004. Cliticization domains: Clitic climbing in Romance and in Serbo-Croatian. In Olivier Crouzet, Hamida Demirdache & Sophie Wauquier-Gravelines (eds.), *Actes des / Proceedings of JEL'2004. [DOMAIN_{[E]}S]: Journées d'études linguistiques 2004, 5–6–7 Mai 2004, Nantes, France*, 191–197. Nantes: Université de Nantes.

Avgustinova, Tania & Karel Oliva. 1997. On the nature of the Wackernagel position in Czech. In Uwe Junghanns & Gerhild Zybatow (eds.), *Formale Slavistik*, 25–47. Frankfurt am Main: Vervuert.

Borsley, Robert D. & María Luisa Rivero. 1994. Clitic auxiliaries and incorporation in Polish. *Natural Language & Linguistic Theory* 12(3). 373–422. DOI: 10.1007/BF01118135.

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier. DOI: 10.1163/9780585474250.

Burzio, Luigi. 1986. *Italian syntax: A government-binding approach* (Studies in Natural Language and Linguistic Theory 1). Dordrecht: D. Reidel Publishing Company. DOI: 10.1007/978-94-009-4522-7.

Cardinaletti, Anna & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.), *Clitics in the languages of Europe* (Empirical Approaches to Language Typology 20), 145–233. Berlin/New York: Mouton de Gruyter. DOI: 10.1515/9783110804010.145.

Chierchia, Gennaro. 1984. *Topics in the syntax and semantics of infinitives and gerunds*. PhD thesis. Amherst, MA: University of Massachusetts. (Doctoral dissertation).

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures*. Dordrecht: Foris.

Culicover, Peter W. & Ray Jackendoff. 2001. Control is not movement. *Linguistic Inquiry* 32(3). 493–512. DOI: 10.1162/002438901750372531.

Dimitrova-Vulchanova, Mila. 1999. Clitics in the Slavic languages. In Henk van Riemsdijk (ed.), *Clitics in the languages of Europe* (Empirical Approaches to Language Typology 20), 83–122. Berlin/New York, NY: Mouton de Gruyter. DOI: 10.1515/9783110804010.83.

Dotlačil, Jakub. 2004. *The syntax of infinitives in Czech*. Tromsø. (MA thesis).

Dotlačil, Jakub. 2007. Why clitics cannot climb out of CP: A discourse approach. In Ulyana Savchenko, Magdalena Goledzinowska & Richard Compton (eds.),

Formal approaches to Slavic linguistics: The Toronto meeting 2006, 76–93. Ann Arbor, MI: Michigan Slavic Publications.

Dziwirek, Katarzyna. 1998. Reduced construction in universal grammar: Evidence from the Polish object control construction. *Natural Language & Linguistic Theory* 16(1). 53–99. DOI: 10.1023/A:1005952030641.

Franks, Steven. 2009. Clitics in Slavic. In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.), *Die slavischen Sprachen: Ein internationales Handbuch zu ihrer Struktur, ihrer Geschichte and ihrer Erforschung* (Handbücher zur Sprach- und Kommunikationswissenschaft 32), 725–738. Berlin/New York, NY: Walter de Gruyter. DOI: 10.1515/9783110214475.1.9.725.

Franks, Steven. 2010. Clitics in Slavic. *Glossos* 10. 1–157. https://slaviccenters.duke.edu/sites/slaviccenters.duke.edu/files/media_items_files/10franks.original.pdf.

Franks, Steven & Tracy H. King. 2000. *A handbook of Slavic clitics*. New York, NY/Oxford: Oxford University Press.

Fried, Mirjam. 1994. Second-position clitics in Czech: Syntactic or phonological? *Lingua* 94(2). 155–175. DOI: 10.1016/0024-3841(94)90024-8.

Golden, Marija. 2003. Clitic placement and clitic climbing in Slovenian. *Sprachtypologie und Universalienforschung [STUF – Language Typology and Universals]* 56(3). 208–233. DOI: 10.1524/stuf.2003.56.3.208.

Golden, Marija. 2008. 2nd position clitic climbing and restructuring. In Gerhild Zybatow, Luka Szucsich, Uwe Junghanns & Roland Meyer (eds.), *Formal description of Slavic languages*, 308–317. Frankfurt am Main: Peter Lang.

Halpern, Aaron. 1995. *On the placement and morphology of clitics*. Stanford, CA: CSLI Publications.

Hana, Jiri. 2007. *Czech clitics in higher order grammar*. Columbus, OH: The Ohio State University. (Doctoral dissertation).

Hnátková, Milena, Michal Křen, Pavel Procházka & Hana Skoumalová. 2014. The SYN-series corpora of written Czech. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, 160–164. Reykjavík, Iceland: European Language Resources Association (ELRA). http://www.lrec-conf.org/proceedings/lrec2014/pdf/294_Paper.pdf.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30(1). 69–96. DOI: 10.1162/002438999553968.

Janse, Mark. 1998. Cappadocian clitics and the syntax-morphology interface. In Geoffrey C. Horrocks Brian D. Joseph & Irene Philippaki-Warburton (eds.), *Themes in Greek linguistics II* (Amsterdam Studies in the Theory and History

of Linguistic Science 159), 257–281. Amsterdam/Philadelphia, PA: John Benjamins. DOI: 10.1075/cilt.159.16jan.

Jung, Hakyung & Krzysztof Migdalski. 2022. Toward a four-way pronoun hierarchy: a view from Slavic. *Journal of Slavic Linguistics* 30(FASL 29 extra issue). 373–422. https://ojs.ung.si/index.php/JSL/article/view/202.

Junghanns, Uwe. 2002a. Clitic climbing im Tschechischen. In Uwe Junghanns (ed.), *Untersuchungen zur Syntax und Informationsstruktur slavischer Deklarativsätze* (Linguistische Arbeitsberichte 80), 57–90. Leipzig: Institut für Linguistik der Universität Leipzig.

Junghanns, Uwe. 2002b. Klitische Elemente im Tschechischen: Eine kritische Bestandsaufnahme. In Thomas Daiber (ed.), *Linguistische Beiträge zur Slavistik: IX. JungslavistInnen-Treffen, Halle a. d. Saale / Lutherstadt Wittenberg, 27. September – 1. Oktober 2000*, 117–150. Munich: Otto Sagner.

Junghanns, Uwe. 2021. Grammatische Regel oder Möglichkeit? In Bernhard Brehmer, Anja Gattner & Tatiana Perevozchikova (eds.), *Von A wie Aspekt bis Z wie zdvořilost: Ein Kaleidoskop der Slavistik für Tilman Berger zum 65. Geburtstag*, 171–184. Tübingen: Tübingen Library Publishing.

Kolaković, Zrinka, Edyta Jurkiewicz-Rohrbacher, Björn Hansen, Dušica Filipović Đurđević & Nataša Fritz. 2022. *Clitics in the wild: Empirical studies on the microvariation of the pronominal, reflexive and verbal clitics in Bosnian, Croatian and Serbian* (Open Slavic Linguistics 7). Berlin: Language Science Press. DOI: 10.5281/zenodo.5792972.

Kraska-Szlenk, Iwona. 1995. *The phonology of stress in Polish*. Urbana, IL: University of Illinois at Urbana-Champaign. (Doctoral dissertation).

Křen, Michal, Václav Cvrček, Tomáš Čapka, Anna Čermáková, Milena Hnátková, Lucie Chlumská, Tomáš Jelínek, Dominika Kováříková, Vladimír Petkevič, Pavel Procházka, Hana Skoumalová, Michal Škrabal, Petr Truneček, Pavel Vondřička & Adrian Zasina. 2019. Corpus SYN, version 8 from 12.12.2019, Prague: Ústav Českého národního korpusu FF UK. http://www.korpus.cz.

Kupść, Anna. 2000. *An HPSG grammar of Polish clitics*. Warsaw/Paris: Institute of Computer Science, Polish Academy of Sciences & Université Paris 7. (Doctoral dissertation).

Landau, Idan. 2003. Movement out of control. *Linguistic Inquiry* 34(3). 471–498. https://www.jstor.org/stable/4179243.

Lenertová, Denisa. 2004. Czech pronominal clitics. *Journal of Slavic Linguistics* 12(1–2). 135–171. http://www.jstor.org/stable/24599870.

Petkevič, Vladimír. 2009. Postavení příklonek v české klauzi v korpusech současné češtiny. In Dana Hlaváčková, Aleš Horák, Klára Osolsobě & Pavel

Rychlý (eds.), *After half a century of Slavonic natural language processing*, 185–196. Brno: Tribun EU.

Progovac, Ljiljana. 1993. Locality and subjunctive-like complements in Serbo-Croatian. *Journal of Slavic Linguistics* 1(1). 116–144. https://www.jstor.org/stable/24599042.

Progovac, Ljiljana. 1996. Clitics in Serbian/Croatian: Comp as the second position. In Aaron L. Halpern & Arnold M. Zwicky (eds.), *Approaching second: Second position clitics and related phenomena*, 411–428. Stanford, CA: CSLI Publications.

Przepiórkowski, Adam, Mirosław Bańko, Rafał L. Górski & Barbara Lewandowska-Tomaszczyk (eds.). 2012. *Narodowy korpus języka polskiego*. Warsaw: Wydawnictwo Naukowe PWN.

Przepiórkowski, Adam & Alexandr Rosen. 2005. Czech and Polish raising/control with or without structure sharing. *Research in Language* 3. 33–66.

Rappaport, Gilbert. 1988. On the relationship between prosodic and syntactic properties of the pronouns in the Slavonic languages. In Alexander M. Schenker (ed.), *American contributions to the 10th international congress of Slavists*, 301–327. Columbus, OH: Slavica.

Rezac, Milan. 2005. The syntax of clitic climbing in Czech. In Lorie Heggie & Francisco Ordóñez (eds.), *Clitic and affix combinations: Theoretical perspectives* (Linguistik Aktuell/Linguistics Today 74), 103–140. Amsterdam: John Benjamins. DOI: 10.1075/la.74.06rez.

Rizzi, Luigi. 1982. A restructuring rule. In Luigi Rizzi (ed.), *Issues in Italian syntax*, 1–48. Dordrecht: Foris Publications.

Rosen, Alexandr. 2001. *A constraint-based approach to dependency syntax applied to some issues of Czech word order*. Prague: Univerzita Karlova. (Doctoral dissertation).

Rosen, Alexandr. 2014. Haplology of reflexive clitics in Czech. In Elżbieta Kaczmarska & Motoki Nomachi (eds.), *Slavic and German in contact: Studies from areal and contrastive linguistics* (Slavic Eurasian Studies 26), 97–116. Sapporo: Slavic-Eurasian Research Center. http://src-h.slav.hokudai.ac.jp/coe21/publish/no26_ses/26-05.pdf.

Rothstein, Robert A. 1993. Polish. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 686–758. London/New York, NY: Routledge.

Rubadeau, Patrice Marie. 1996. *A descriptive study of clitics in four Slavic languages: Serbo-Croatian, Bulgarian, Polish, and Czech*. Ann Arbor, MI: University of Michigan. (Doctoral dissertation).

Skoumalová, Hana. 2005. *Review of: Clitic and affix combinations: Theoretical perspectives*. In LINGUIST List 16.3131, Sat 29 Oct 2005. https://linguistlist.org/issues/16/16-3131.html.

Irenäus Kulik

Spencer, Andrew. 1991. *Morphological theory. An introduction to word structure in generative grammar.* Oxford/Cambridge, MA: Basil Blackwell.

Spencer, Andrew & Ana R. Luís. 2012. *Clitics: An introduction.* Cambridge: University Press.

Stjepanović, Sandra. 1998. On the placement of Serbo-Croatian clitics: Evidence from clitic climbing and VP ellipsis. In Željko Bošković, Steven Franks & William Snyder (eds.), *Annual workshop on formal approaches to Slavic linguistics: The Connecticut meeting 1997*, 267–286. Ann Arbor, MI: Michigan Slavic Publications.

Stjepanović, Sandra. 1999. *What do second position cliticization, scrambling and multiple wh-fronting have in common?* Storrs, CT: University of Connecticut. (Doctoral dissertation).

Stjepanović, Sandra. 2004. Clitic climbing and restructuring with "finite clause" and infinitive complements. *Journal of Slavic linguistics* 12(1–2). 173–212. https://www.jstor.org/stable/24599871.

Todorović, Nataša. 2012. *The indicative and subjunctive da-complements in Serbian: A syntactic-semantic approach.* Chicago, IL: University of Illinois at Chicago. (PhD thesis).

Toman, Jindřich. 1996. A note on clitics and prosody. In Aaron L. Halpern & Arnold M. Zwicky (eds.), *Approaching second: Second position clitics and related phenomena*, 505–510. Stanford, CA: CSLI Publications.

Trávníček, František. 1959. K postavení stálých příklonek po přestávce uvnitř věty. *Naše Řeč* 42(3-4). 65–76. http://nase-rec.ujc.cas.cz/archiv.php?art=4673.

Urbańczyk, Stanisław. 1976. *Zarys dialektologii polskiej.* Warsaw: Państwowe Wydawnictwo Naukowe.

Veselovská, Ludmila. 1995. *Phrasal movement and X^0-morphology: Word order parallels in Czech and English nominal and verbal projections.* Olomouc: Univerzita Palackého. (PhD thesis).

Weinreich, Uriel, William Labov & Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehman & Yakov Malkiel (eds.), *Directions for historical linguistics: A symposium*, 95–195. Austin, TX/London: University of Texas Press.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure* (Studies in Generative Grammar 55). Berlin/New York, NY: Mouton de Gruyter. DOI: 10.1515/9783110908329.

Zwicky, Arnold M. 1977. *On clitics.* Bloomington, IN: Indiana University Linguistics Club.

Chapter 8

Responding to negative biased questions in Russian

Ⓓ Ljudmila Geist[a] & Ⓓ Sophie Repp[b]

[a]University of Stuttgart [b]University of Cologne

The paper investigates polar responses to biased questions with outer vs. inner negation and the particle *razve* 'really' in Russian. We present experimental evidence from two acceptability judgment studies and show that the two question types have slightly different answer patterns. We argue that the meaning previously suggested for the particles *da/net* 'yes/no' must be revised. We propose an analysis of our results which combines a proposal for outer vs. inner negation in terms of the illocutionary operator FALSUM vs. propositional negation (Repp 2006, 2009), and a proposal for response particles in terms of propositional anaphors that realize certain polarity features (Roelofsen & Farkas 2015). We argue that the set of polarity features hitherto assumed should be extended to features that are sensitive to the type of antecedent that polar responses react to: assertion or question.

Keywords: question, question bias, negation, response particle, propositional anaphor, acceptability judgments

1 Introduction

Response particles like *yes* and *no* have been assumed to fulfil two functions: they may affirm or reject the truth of a previous utterance (truth-based function), or they may signal the polarity of the response (polarity-based function). The difference becomes relevant in responses to assertions or questions with a negation. For instance, in reaction to the assertion *Nina didn't sneeze*, a particle like *yes* in principle may signal that the assertion is true, i.e. signal agreement with *Nina didn't sneeze*, but it may also signal that the response is positive, i.e. that *Nina sneezed*. Languages differ with respect to which of these functions the

individual response particles preferably realize – or in how far these functions are combined. There has been much research on cross-linguistic as well as inter-individual variation on this issue in recent years, and earlier assumptions that there might be a division into TRUTH-BASED LANGUAGES and POLARITY-BASED LANGUAGES (Pope 1976, Jones 1999) have been called into question (e.g., Krifka 2013, Goodhue & Wagner 2018, González-Fuente et al. 2015, Kramer & Rawlins 2011, Holmberg 2013, 2015, Meijer et al. 2015, Roelofsen & Farkas 2015, Li et al. 2016, Claus et al. 2017, Farkas & Roelofsen 2019, Repp et al. 2019, Loos et al. 2020).

Response particles are generally thought to be anaphoric devices. They have been analysed as propositional anaphors (Krifka 2013, Roelofsen & Farkas 2015, Farkas & Roelofsen 2019), and as remnants of an elliptic clause (Kramer & Rawlins 2011, Holmberg 2013, 2015). As propositional anaphors they refer to a salient proposition in the previous utterance. While assertions normally are assumed to introduce one proposition (unless they contain a negation), questions are usually assumed to introduce a set of two propositions (e.g., Hamblin 1973). For instance, the polar question *Nina čichnula?* 'Did Nina sneeze?' introduces the positive proposition p, *Nina sneezed*, and the negative proposition $\bar{p}$, *Nina did not sneeze*. In principle, response particles may take up either proposition as antecedent but since anaphors are sensitive to the salience of potential antecedents, and since it has been argued that the particular form of a question may influence the salience of the two propositions, the issue arises which proposition a particle picks up.

Formal aspects potentially influencing the salience of p or $\bar{p}$ include for instance the presence vs. absence of a negative marker (e.g., Roelofsen & van Gool 2010, Roelofsen & Farkas 2015), the form and position of the negative marker, and the presence of certain particles. These formal means mark certain contextual and speaker-related biases, which may correspond to p or $\bar{p}$ (e.g., Ladd 1981, Büring & Gunlogson 2000, Romero & Han 2004, Repp 2009, Sudo 2013, Seeliger 2015, 2019, Gyuris 2017, Seeliger & Repp 2018, Arnhold et al. 2021, Repp & Geist to appear). To illustrate, a question like *Didn't Nina sneeze?* may be used to double-check the truth of *p (Nina sneezed)* because the speaker had assumed that p is true – this might make p salient. The same question may also be used to double-check the truth of *$\bar{p}$ (Nina didn't sneeze)* because this is what the evidence suggests – this might make $\bar{p}$ salient. Most accounts of question bias assume different analyses for the negation in these two question uses (or meanings): as OUTER NEGATION and INNER NEGATION, respectively, so that a question with outer negation (ON-QUESTION) double-checks a positive proposition, and a question with inner negation (IN-QUESTION) checks a negative proposition. Hence, it is to be expected

that *yes* and *no* as well as their correlates in other languages pick up different propositions when answering ON- vs. IN-questions.

In this paper we investigate the meaning and use of the response particles *da/net* 'yes/no' in Russian in responses to biased ON/IN-questions in Russian. We present quantitative evidence from two acceptability judgment experiments. The goal of our investigation is to improve our understanding of bias in questions on the one hand, and of the meaning and use of response particles, on the other hand. In Russian, polar questions typically have a declarative syntax, and are distinguished from declaratives by prosody. To indicate question bias, interrogative particles may be used. The two readings of polar questions as ON- vs. IN-questions are attested, albeit not necessarily by this terminology (e.g., Baranov & Kobozeva 1983, Brown & Franks 1995, Brown 1999, Kobozeva 2004: 307, Meyer 2004, Šatunovskij 2005). As for the meaning and use of response particles, Russian has been argued to combine truth-based and polarity-based strategies (González-Fuente et al. 2015, Esipova 2021). Most previous investigations on this issue focus on lexical, prosodic and (co-speech) gestural answering strategies in responses to positive and negative antecedents without considering a potential difference between ON/IN-question readings. However, work by Restan (1972), Meyer (2004) and, most recently, the experimental work by Pančenko (2021) on *da/net* in responses to negative questions suggests that the ON/IN-difference plays a role for the acceptability of the Russian response particles.

The paper is structured as follows. §2 discusses the notion of question bias in relation to ON/IN-readings both in general and for Russian. §3 discusses the analysis of response particles in one of the anaphora accounts (Roelofsen & Farkas 2015, Farkas & Roelofsen 2019). §4 presents the two acceptability studies. §5 discusses the results and provides a theoretical evaluation.

2 Polar question bias and negation

2.1 Background

As mentioned above, negative polar questions may express certain contextual and speaker-related biases. Two dimensions have proven helpful in the analysis of these biases (Sudo 2013, Gärtner & Gyuris 2017): (i) EPISTEMIC BIAS (roughly: prior speaker belief or speaker knowledge) and (ii) EVIDENTIAL BIAS (current situational evidence, including propositions implied by the addressee).[1] For instance, in the context description in (1) we learn about a belief of the person

---

[1]Epistemic bias has also been associated with the speaker's desires or expectations (Sudo 2013). We are not considering these meaning aspects here.

Ljudmila Geist & Sophie Repp

asking the question, Sarah. Sarah believes that the proposition p, *Ms Miller has already booked the tickets*, is true. This belief implies that the departure time for the flights under discussion cannot be changed. Tom's suggestion to take an earlier flight (= the evidence) therefore is incompatible with Sarah's belief: the evidence suggests that $\bar{p}$ is true. To resolve this conflict between the evidential and the epistemic bias, Sarah asks a negative polar question.

(1) Sarah and Tom are preparing a business trip to Milan. Ms Miller, their secretary, is helping them. Just before they go home, Sarah and Tom are talking about the business trip. Sarah assumes that Ms Miller has organized everything and the departure time of the flights is fixed.

Tom: Maybe we should take an earlier flight.

Sarah: Hasn't Ms Miller booked the tickets?

As mentioned above, a question like Sarah's may double-check the epistemic bias or the evidential bias.[2] Ladd (1981) argued that the presence of a positive polarity item (PPI) vs. a negative polarity item (NPI) disambiguates the two readings. We are showing this for the PPI *already* and the weak NPI *yet* in (2a)–(2b), since we used the Russian counterparts of these elements in our experiments. (2a) contains *already*, (2b) contains *yet*. Both questions are negative but in (2a) the negation does not seem to anti-license the PPI, which is why it is called OUTER NEGATION. The negation licensing the NPI in (2b) is INNER NEGATION (Romero & Han 2004). The idea behind this terminology is that outer negation is "too far out" to anti-license the PPI, whereas inner negation is close enough to license the NPI (Ladd 1981). Table 1 summarises the main characteristics of ON/IN-questions.

(2) a. Hasn't Ms Miller *already* booked the tickets?

   b. Hasn't Ms Miller booked the tickets *yet*?

The difference between the two negations has been analysed in various ways, for instance in terms of scope relations between the negation and an epistemic conversational operator (Romero & Han 2004), as illocutionary vs. propositional

---

[2] For English, this ambiguity only is obligatorily present with so-called preposed negation, i.e. with the negation marker cliticized to the auxiliary like in (1). Questions with non-preposed negation, i.e. *Has Ms Miller not booked the tickets?*, do not necessarily have the implicature that the speaker had a previous belief: they can be asked in neutral contexts (Romero & Han 2004). We are not considering the difference between preposed and non-preposed negation here as we did not manipulate the position of the Russian negation-plus-verb complex in our experimental materials.

Table 1: Characteristics of ON/IN-questions

Form	Polarity item	Epistemic bias	Evidential bias	"Function"	Negation
Hasn't Ms Miller already booked the tickets?	PPI	p	$\bar{p}$ or none	double-checks p	outer
Hasn't Ms Miller booked the tickets yet?	NPI	p	$\bar{p}$	double-checks $\bar{p}$	inner

negation (Repp 2006, 2009, 2013; also Romero 2015), or in terms of scope relations between speech act operators (Krifka 2015); see Romero (2020) for a review. We are following here the analysis proposed by Repp (2006, 2009, 2013).

Repp assumes that outer negation corresponds to the illocutionary (or common ground managing) operator FALSUM. FALSUM expresses that the speaker is sure that the proposition in its scope should not be added to the common ground. Being an illocutionary operator, FALSUM always scopes over a (positive) proposition (unless there are several negation markers), but it scopes under the question operator so that a question with FALSUM asks whether or not the speaker is sure that a given proposition should not be added to the common ground. Thus, in this analysis a biased question is not a set of two propositions but a set of two semantic-pragmatic objects including an illocutionary operator, see (3a) for the proposed logical form (LF) of ON-questions and their meaning. For inner negation, Repp builds on Romero & Han (2004), who assume that preposed negation obligatorily introduces a conversational epistemic operator VERUM (based on Höhle's 1988, 1992 VERUM focus). VERUM expresses that the speaker is sure that the proposition in its scope should be added to the common ground.[3] Repp

---

[3]Romero & Han (2004) propose a VERUM analysis for both ON- and IN-questions. They assume that in ON-questions, VERUM, which itself is in the scope of negation, scopes over a positive proposition: [Q [¬ VERUM p]]. In IN-questions, VERUM scopes over a negative proposition: [Q [VERUM $\bar{p}$]]. Repp (2006, 2009, 2013) departs from this proposal inter alia because an analysis in terms of VERUM in some contexts produces meanings that are "too weak". For instance, for rejections like *She DIDn't buy the tickets*, Romero & Han (2004) also assume a VERUM analysis. However, [¬ VERUM [*she bought the tickets*]] means that the speaker is not sure that the proposition *she bought the tickets* should be added to the common ground, contrary to the intuition of what this rejection expresses, namely that the speaker is sure that this proposition should not be added to the common ground. Also see Romero (2015) for an analysis of negative

assumes that VERUM, like FALSUM, is an illocutionary operator and takes scope over a proposition. In IN-questions, VERUM scopes over a negative proposition because the negation in these questions is propositional negation, see (3b) for the corresponding LF. A question with VERUM asks whether or not the speaker is sure that a given negative proposition should be added to the common ground. Note that the occurrence of PPIs in ON-questions and of NPIs in IN-questions is predicted by this account because only in the latter is there propositional negation, which by hypothesis is required to license NPIs.

(3)   a.   ON-question: $[Q \, [\text{FALSUM} \, p]] = \{\text{FALSUM} \, p, \neg \, \text{FALSUM} \, p\}$

      b.   IN-question: $[Q \, [\text{VERUM} \, \bar{p}]] = \{\text{VERUM} \, \bar{p}, \neg \, \text{VERUM} \, \bar{p}\}$

Repp's account predicts that in responses to ON- vs. IN-questions, different propositions are made available for anaphoric uptake: p and $\bar{p}$, respectively. Evidence that this might indeed be the case comes from acceptability rating studies in German. Claus et al. (2016) and Repp et al. (2022) show that ON-questions are answered as if they were positive questions. This is expected if the negation in ON-questions is not propositional. Responses to IN-questions do not show this pattern. In our study, we will test whether the predictions of Repp's account for ON- vs. IN-questions can be confirmed for Russian.

2.2 Question bias and negation in Russian

As already mentioned, Russian polar questions by default have the form of declarative sentences: subject-verb-object order without subject-auxiliary inversion. Questionhood is marked by intonation: whereas in (out-of-the-blue) declaratives the default nuclear accent is on the object of the clause, in (out-of-the-blue) interrogatives it is on the verb (Bryzgunova 1975, Ladd 1996). The accent in interrogatives is described as a steep rise $L + H^*$ with peak delay into the postnuclear syllable, which may be followed by a secondary L^* target (Meyer & Mleinek 2006; cf. Bryzgunova 1980).

Russian has interrogative particles that indicate different question biases: *razve*, *neuželi*, *li*, *ved'*, *že*, among others (e.g., Švedova et al. 2005: 387f.). Here we discuss the particle *razve* 'really', which we used in our experiments. *Razve* is used in situations where there is an evidential bias for the proposition denoted by the clause that is used as question, and an epistemic bias for the complement of this proposition (Repp & Geist to appear). For instance in (4), A's utterance implies that Ivan is married (evidential bias for *p*). The occurrence of *razve* in B's

_____

    polar questions that uses both VERUM and FALSUM.

questions (*p*?) indicates that B originally had the belief that Ivan is not married (epistemic bias for *p̄*). The use of *razve* in B's question indicates moderate surprise or doubt concerning the evidence in view of B's original belief (Apresjan 1980, Rathmayr 1985, Baranov 1986, Kirschbaum 2001, Mat'ko 2014), and signals that B wishes to double-check the evidential bias *p* (*he is married*).

(4) A: Ivan ezdil v otpusk so   svoej   ženoj.
     Ivan went in holiday with his.own wife
     'Ivan was on holiday together with his wife.'

   B: A   razve on ženat?
     but PART he married
     'But is he really married?'   (Zaliznjak 2020: 5)

Razve can also occur in negative questions. Negation in Russian is expressed by the preverbal particle *ne*. Repp & Geist (to appear) present experimental evidence which indicates that negative questions (*p̄*?) with *razve* are more acceptable when they occur in biased contexts, i.e. in contexts where there is evidence for *p̄* and the speaker had a previous belief for *p*, than when they occur in neutral contexts. Negative questions without *razve* display the opposite pattern.

As already mentioned, there are descriptions of ON- and IN-question readings in the literature (Restan 1972, Baranov & Kobozeva 1983, Brown & Franks 1995, Brown 1999, Meyer 2004, Kobozeva 2004, Šatunovskij 2005, Pančenko 2021, Repp & Geist to appear). Whether or not the position of the negation-verb complex (clause-initial or not) contributes to the different readings is controversial (Brown & Franks 1995, Meyer 2004). Repp & Geist (to appear) discuss data from the Russian National Corpus (ruscorpora.ru; Rachilina 2008) with the negation-verb complex in non-initial position which show that both ON- and IN-readings are available in questions with *razve* (see (9) and (10) further below). Repp & Geist (to appear) assume that Russian *eščë*, the approximate counterpart of the English NPI *yet*, indicates the inner negation reading, and Russian *uže*, the approximate counterpart of the English PPI *already*, indicates the outer negation reading.[4]

As just indicated, *eščë* and *uže* cannot be fully identified with *yet* and *already*: *eščë* and *uže* have many different uses (Boguslavskij 1996). The polarity-sensitive

---

[4]There are other diagnostics in Russian to distinguish the two readings. For instance, Pančenko (2021) provides experimental evidence showing that ON is marked by the combination of the particle *li* with *ne* (*ne...li*). Meyer (2004), following Restan (1972), argues that certain modal particles and sentence adverbs, for instance *že* '≈ but', *ved'* '≈ but', *konečno* 'of course' and *stalo byt'* 'apparently', may only occur in IN-questions and not in ON-questions. See Brown & Franks (1995) and Meyer (2004) for other morphosyntactic cues. The role of intonation is uncertain (Meyer 2004, Pančenko 2021).

uses that we are interested in here are attested in combination with a verb in perfective aspect. In this context, *eščë* patterns with the English NPI *yet* and needs licensing by negation, (5), whereas *uže* patterns with *already* and is excluded under sentence negation, (6).

(5) John has left {already / *yet}. *positive clause*
     Ivan uechal {uže / *eščë}.

(6) John has not left {*already / yet}. *negative clause*
     Ivan ne uechal {*uže / eščë}.

The polarity sensitivity of *eščë* and *uže* furthermore shows up in combination with other NPIs and PPIs. *Eščë* may co-occur with strong NPIs like the negative pronoun *nikuda* 'nowhere', (7), but cannot co-occur with PPIs like the intensifier *gorazdo* 'considerably', (8) (cf. van der Wouden 1997 for intensifiers as PPIs). For *uže* it is the other way round.

(7) Ivan {*eščë nikuda$_{NPI}$ / ^{OK}uže   gorazdo$_{PPI}$ bystree} uechal.
     Ivan yet nowhere   already considerably faster   left
     Intended: 'Ivan hasn't left anywhere yet.' /
     'Ivan has left already considerably faster.'

(8) Ivan {^{OK}eščë nikuda$_{NPI}$ / *uže   gorazdo$_{PPI}$ bystree} ne uechal.
     Ivan   yet nowhere   already considerably faster   not left
     'Ivan hasn't left anywhere yet.' /
     Intended: 'Ivan has left already considerably faster.'

Turning now to the occurrence of *eščë* and *uže* in negative questions in the Russian National Corpus, as discussed by Repp & Geist (to appear), consider (9) and (10). In (9) speaker B has an epistemic bias for the positive proposition p (*A has already told me the main thing*). However, A's utterance provides evidence for $\bar{p}$. To resolve the conflict, B asks a question containing the NPI *eščë*, double-checking $\bar{p}$, the evidential bias.

(9) A: Sejčas ja tebe skažu glavnoe.
         now   I you tell main.thing
         'Now I am telling you the main thing.'
     B: Razve eščë ne skazal?
         PART yet not said
         'Haven't you told it to me yet?'   *IN-question*
                     [A. I. Spasovskiy, "Bolšaja kniga peremen / Volga" 2010]

The assumption that a negative *razve*-question containing *eščë* is indeed an IN-question is supported by the observation that the strong NPI *ni razu* 'not once' can occur in such a question:

(10) Razve eščë ni razu ne skazal?
 PART yet NEG once not said
 'Haven't you ever told me?'

Example (11) shows that *uže* can occur in a *razve*-question, indicating that *razve*-questions can be ON-questions. The *razve*-question in (11) conveys the same biases as the *razve*-question in (9): an epistemic bias for *p* (*You have dragged me out of the past already*), and an evidential bias for *p̄*. To resolve the conflict, the speaker asks the question. Here it is the epistemic bias that is checked, as is indicated by the presence of the PPI *uže*. The question is an ON-question.

(11) A: Čestnoe slovo, ne znaju, kak vytaščit' tebja iz prošlogo.
 honest word not know how drag you out.of past
 'Frankly, I don't know how to drag you out of the past.'
 B: Razve ty uže ne vytaščila menja iz prošlogo?
 PART you already not dragged me out.of past
 'Haven't you dragged me out of the past already?' *ON-question*
 [Alexander Bogdan, Gennadi Praškewič. "Čelovek Č" 2001]

As is shown in (12), the outer negation in the *razve*-question in (11) anti-licenses the strong NPI *ni razu*, which supports the assumption that the question in (11) indeed is an ON-question.

(12) * Razve ty uže ni razu ne vytaščila menja iz prošlogo?
 PART you already NEG once not dragged me out.of past
 Intended: 'Haven't you dragged me out of the past once already?'

We conclude that a negative *razve*-question *p̄*? comes with an epistemic bias for *p* and an evidential bias for *p̄*. The question may double-check, and – by hypothesis – make salient, different propositions. Which proposition is double-checked and made salient may be disambiguated by polarity-sensitive items like *eščë* and *uže*.

3 Response particles

There are various analyses of response particles, which fall into two major types: anaphora and ellipsis analyses. We already mentioned in §1 that response particles have been analysed as propositional anaphors, i.e. they take up a salient

Ljudmila Geist & Sophie Repp

proposition in the discourse context (Krifka 2013, Roelofsen & Farkas 2015, Farkas & Roelofsen 2019). Ellipsis accounts treat response particles as remnants of elided response clauses (Kramer & Rawlins 2011, Holmberg 2013, 2015). All these accounts aim at explaining the gradual differences in the acceptability and use of response particles that have been observed in recent years. For reasons of space, we only discuss one of the anaphora accounts here, namely Roelofsen & Farkas's feature model (Roelofsen & Farkas 2015, Farkas & Roelofsen 2019).

3.1 The feature model

Roelofsen & Farkas (2015) assume that response particles like English *yes* and *no* realize two types of semantic presuppositional features, which are formal instantiations of the two functions that response particles were argued to have in earlier literature: to indicate the polarity of the response or the truth of the antecedent (e.g., Pope 1976, Jones 1999). Accordingly, the first type of feature are ABSOLUTE POLARITY features, which presuppose that the polarity of the response is positive (feature [+]) or negative (feature [−]). The second type are RELATIVE POLARITY features, which presuppose that the response has the same or the opposite polarity of the antecedent (the features [AGREE] and [REVERSE]).

In the feature model, language-specific FEATURE-PARTICLE MAPPINGS indicate which particle may realize which feature. For instance, English maps [+] and [AGREE] onto *yes*, and [−] and [REVERSE] onto *no*. Some languages map feature combinations onto a dedicated particle, like German does for [+, REVERSE], which maps onto *doch*. The feature-particle mapping for English in comparison to German as suggested by Roelofsen & Farkas (2015) is given in Table 2.

Table 2: The feature-particle mapping for English and German

English:	
[+] and [AGREE] → *yes*	[−] and [REVERSE] → *no*
German:	
[+] and [AGREE] → *ja*	[−] and [REVERSE] → *nein*
	[+, REVERSE] → *doch*

In responses to positive assertions and questions, the absolute and relative polarity of response particles coincide, but in responses to negative questions and assertions these two functions come apart. This is illustrated in (13), where the feature combination of the whole response is given in square brackets and the

feature realized by the respective particle is marked by a frame. In (13a)–(13b) the absolute polarity feature to be realized is [−] because the response clause contains negation, and the relative polarity feature to be realized is [AGREE] because the polarity of the response is the same as the polarity of the antecedent. In (13a) *no* realizes [−], and in (13b) *yes* realizes [AGREE].

(13)   Antecedent:
       Ms Miller hasn't booked the tickets. / Hasn't Ms Miller booked the
       tickets?

       Response:

       a.   No, she hasn't.   [$\boxed{-}$, AGREE]
       b.   Yes, she hasn't.   [−, $\boxed{\text{AGREE}}$]
       c.   No, she has.   [+, $\boxed{\text{REVERSE}}$]
       d.   Yes, she has.   [$\boxed{+}$, REVERSE]

The pattern shown in (13) reflects the feature-particle mapping for English but it does not represent the actual preference patterns for *yes* and *no* in English in the various discourse contexts. In other words, although both particles may in principle realize both types of features, there are clear differences in (graded) acceptability and use (Roelofsen & Farkas 2015, Repp et al. 2019). To account for such observations, Farkas & Roelofsen (2019) model the realization of features in a stochastic optimality-theoretic (OT) framework. In this model, different constraint weightings are used to explain language-specific answering patterns and gradual preference patterns. Table 3 lists the constraints.

Table 3: OT constraints in the feature model (Farkas & Roelofsen 2019)

MAXIMIZE MARKED:	Maximize the realization of marked polarity features or feature combinations.
EXPRESSIVENESS:	Maximize the expression of feature content.
MAXIMIZE RELATIVE:	Maximize the realization of relative polarity features.
MAXIMIZE ABSOLUTE:	Maximize the realization of absolute polarity features.

The constraint MAXIMIZE MARKED is a typical OT markedness constraint and thus is thought to be generally operative in response systems. It favours the realization of marked features or feature combinations. The features [−] and

[REVERSE] are thought to be marked: negation [−] is assumed to be hard to process, and disagreeing in discourse [REVERSE] is a dispreferred discourse move. The feature combination [+, REVERSE] also is considered to be marked. In a language where the constraint MAXIMIZE MARKED has a particularly high weight, marked features have a particularly high realization need and a particle that realizes a marked feature (combination) will be preferred over other particles.

The constraint EXPRESSIVENESS is relevant if there is a preference in a language for particles expressing more rather than less features. For instance, for German, EXPRESSIVENESS is assumed to have a high weight, which explains why the particle *doch*, which realizes [+, REVERSE], is more accepted in [+, REVERSE] responses than particles realizing only one of the features [+] and [REVERSE]. EXPRESSIVENESS can be viewed as an instance of the general principle MAXIMIZE PRESUPPOSITION! (Heim 1991): the polarity features are presuppositional.

The constraints MAXIMIZE RELATIVE and MAXIMIZE ABSOLUTE, by which relative and absolute polarity features, respectively, have a high realization need, are response-specific constraints, and arguably cannot be linked to more general principles. However, given that languages do display different general tendencies to express truth vs. polarity (see §1), it seems warranted to assume these constraints.

To see how these constraints can be used to explain gradual preferences for response particles, consider how Repp et al. (2019) explain findings from an acceptability judgment experiment testing *yes* and *no* responses to negative assertions in English. Repp et al. suggest that the relative weight of two of the above constraints is relevant to account for the data (the other constraints have low weights), see (14), where ≻ stands for 'has greater weight than'.

(14)   REALIZE ABSOLUTE FEATURES ≻ REALIZE MARKED FEATURES

The acceptability patterns found by Repp et al. are shown in (15).[5] As before, a frame indicates the feature that is realized. In addition, marked features are highlighted in grey. (15) shows that in agreeing responses, (15a), *no* was much more acceptable (≫) than *yes*. In these responses, *no* realizes absolute, marked [−], and *yes* realizes relative, unmarked [AGREE]. In rejecting responses, (15b), *yes* was more acceptable (>) than *no* but the difference was not so extreme. In rejecting responses, *yes* realizes absolute, unmarked [+], and *no* realizes relative, marked [REVERSE]. Thus, in both agreeing and rejecting responses, the particle realizing the absolute feature was more acceptable than the particle realizing the relative feature. However, only in agreeing responses the particle realizing the

---

[5]We are glossing over the inter-individual differences found by Repp et al. (2019).

marked feature was more acceptable than the particle realizing the unmarked feature. This pattern can be explained with the weighting indicated in (14): realizing absolute features has more weight in English than realizing marked features.

(15)   Antecedent:
       Ms Miller hasn't booked the tickets.

       Response:
       a.   No, she hasn't. [$\boxed{-}$, AGREE] ≫ *Yes, she hasn't. [$-$, $\boxed{\text{AGREE}}$]
       b.   Yes, she has. [$\boxed{+}$, $\boxed{\text{REVERSE}}$] > No, she has. [+, $\boxed{\text{REVERSE}}$]

3.2 Russian response particles in the feature model

Russian has two response particles: *da* and *net*. In two recent feature model analyses (Esipova 2021, González-Fuente et al. 2015), which do not distinguish between ON- and IN-questions, Russian has been proposed to differ from English in its feature-particle mapping. Like English *no*, Russian *net* may realize the absolute feature [−] or the relative feature [REVERSE]. Unlike English *yes*, however, Russian *da* may only realize the relative feature [AGREE]. Thus, the proposed feature-particle mapping is the one given in (16), and the corresponding acceptability pattern is illustrated in (17) from Esipova (2021).[6]

(16)   Russian: [AGREE] → *da*   [−] and [REVERSE] → *net*

(17)   Antecedent:
       Nina ne sdala   ekzamen {?,.}
       Nina not passed exam
       '{Did Nina not pass the exam?, Nina did not pass the exam.}'

       Response:
       a.   Net, ne sdala.
            no   not passed
            'No, she didn't.'   [$\boxed{-}$, AGREE]

       b.   Da, ne sdala.
            yes not passed
            'Yes, she didn't.'   [−, $\boxed{\text{AGREE}}$]

---

[6]Esipova (2021) assumes the same pattern for questions and assertions as antecedents. However, she does not specify the bias profile or the ON/IN-readings of the questions. The non-preposed position of the negation in the English translation given by Esipova might be taken to hint at a 'bias-free' reading, which like the IN-negation reading arguably makes $\bar{p}$ salient, but Esipova is not explicit on this issue.

    c.   Net, sdala.
           no   passed
           'No, she did.'   [+, REVERSE]

    d. * Da, sdala.
           yes passed
           Intended: 'Yes she did.'   [+ , REVERSE]

<div align="right">Esipova (2021: 3f.)</div>

Meyer (2004) (following Restan 1972, Brown & Franks 1995) distinguishes between "purely informative" negative questions (questions without a bias) as antecedents, and questions with a negative implicature (the speaker expects a negative answer). For the former type of question, Meyer suggests that only the responses given in (17a) and (17d) are acceptable. Thus, the pattern is clearly different from the one given by Esipova (2021) in (17). According to Meyer, *da* and *net* undoubtedly indicate absolute polarity as responses to such questions, i.e. [+] and [−]. However, Repp & Geist (2022) report experimental evidence on responses to unbiased questions in rich discourse contexts which does not confirm Meyer's claims: *da* was clearly degraded in responses to such questions whereas *net* was rated as acceptable – both independently of the polarity of the response. For questions with a negative implicature – which is a category that does not fit our description of biases – Meyer (2004) proposes the same pattern as the one given by Esipova in (17a)–(17d). He also highlights that the pattern would be the same with assertions as antecedents, thus corroborating Esipova's suggestion. However, since the question type is not specified by Esipova, a comparison is difficult. Overall, this empirical picture leaves open many questions and needs careful empirical investigation, especially in rich discourse contexts so that the exact question meaning can be controlled. For our investigation, we will work with the hypothesis that *da* can only realize [AGREE] (Esipova 2021, González-Fuente et al. 2015).

    For sake of completeness, it should be noted here that in addition to particles, Russian uses lexico-syntactic response strategies. For instance, González-Fuente et al. (2015) identify the echoic answering strategy, where the speaker may repeat the verb without a particle, for instance to mark a rejection like (17d). We restrict our investigation to the response particles *da* and *net*.

4 Acceptability judgment experiments

In this section we are presenting the acceptability judgment experiments that we conducted to explore the feature-particle mapping for Russian *da* and *net* as sum-

marized in (16), for responses to biased ON/IN-questions, where the two types of negation are signalled by the polarity-sensitive items *uže* 'already' and *eščë* 'yet'. Specifically, we explored the predictions that can be made on the basis of Repp's (2006, 2009, 2013) analysis of such questions in English and German. Recall that according to this analysis, ON-questions vs. IN-questions make different propositions available for anaphoric uptake, which predicts that the type of negation will influence the felicity of *da/net* for expressing that p or $\bar{p}$ is true. We hypothesized that in responses to ON-questions, which check the epistemic bias for p and according to Repp have the LF [Q [FALSUM p]], the positive proposition p is taken up by *da/net*. In responses to IN-questions, which check the evidential bias $\bar{p}$ and have the LF [Q [VERUM $\bar{p}$]], it is the negative proposition $\bar{p}$ which is taken up by *da/net*.

Table 4 summarizes our specific predictions. For responses expressing that p (= the epistemic bias) is true, we predict that after ON-questions only *da* is felicitous because only *da* can realize one of the features that potentially can be realized in such discourses ([AGREE] and [+]): *da* realizes [AGREE], which presupposes that antecedent polarity and response polarity are the same. After IN-questions, we predict that only *net* is felicitous: it realizes [REVERSE], which presupposes that antecedent polarity and response polarity are the opposite. For responses expressing that $\bar{p}$ (the evidential bias) is true, we predict that after ON-questions, only *net* is felicitous: *net* indicates the negative polarity of the response, and it indicates that the polarities of antecedent and response are the opposite. After IN-questions, *net* should be felicitous because it expresses negative response polarity, and *da* should be felicitous because it signals that antecedent and response polarity are the same. However, *net* should be preferred over *da* by MAXIMIZE MARKED FEATURES because *net* realizes a marked feature whereas *da* does not.

We note here that although ON/IN-questions by their structure are assumed to introduce only one propositional discourse referent, the context might make additional propositions available. ON-questions double-check the epistemic bias for p for a reason: there is evidence for $\bar{p}$ in the context. Therefore, it might be the case that $\bar{p}$ is salient to some extent. Similarly, IN-questions double-check the evidential bias $\bar{p}$ for a reason: the speaker believed p to be true. So p might be perceived to be salient to some extent. This interplay is not reflected in the LF of the questions and raises the interesting issue of the discourse status of the "unchecked" biases. We will come back to this issue in §5.

Ljudmila Geist & Sophie Repp

Table 4: Predictions for feature realization preferences in responses to Russian ON/IN-questions

State of affairs = polarity of response	Antecedent	
	ON-question *Hasn't ... already...?* [Q [FALSUM p]]	IN-question *Hasn't ... yet...?* [Q [VERUM $\bar{p}$]]
p	[+, AGREE] → *da*	[+, REVERSE] → *net*
$\bar{p}$	[− , REVERSE] → *net* [−, REVERSE] → *net*	[− , AGREE] → *net* [−, AGREE] → *da* *net > da*

4.1 Method

In our acceptability judgment experiments, we presented participants with question-answer dialogues embedded in contexts which make clear what the contextual evidence, the speaker's previous beliefs, and the actual state of affairs (SoA) are. Experiment 1 tested responses to ON-questions, and Experiment 2 tested responses to IN-questions. We describe the two experiments together because of the great overlap in materials and method.

The materials of our study were based on those used in the experiments reported in Claus et al. (2017) (also see Meijer et al. 2015). Claus et al. investigated responses to assertions in German, so we translated and localized the materials, and we adapted the contexts to license the question biases. The experimental items were descriptions of short scenarios including a question-answer dialogue between two interlocutors, Dima and Katja. The question was an ON-question (Experiment 1) or an IN-question (Experiment 2), and the answer consisted of a response particle (*da*, *net*) and an answer clause.

Both experiments had a 2 × 2 design with the factors STATE OF AFFAIRS (SoA) and PARTICLE. (18) is a sample item. Each item started with a description of a situation, which informed the reader about the general setting, including information on whether or not a certain SoA obtained or not (= factor SoA). In (18) the SoA concerned whether Marina Petrovna had booked tickets for a flight or not. For mnemonic reasons, we are using the strings DONE and NOT DONE to indicate whether the relevant SoA obtains (p is true), or not ($\bar{p}$ is true). The SoA was what the question-answer dialogue was about. The description of the situation further contained information about the knowledge states and assumptions of the inter-

locutors and the existing contextual evidence (epistemic and evidential bias). The person asking the question, Katja, always believed that p is true (epistemic bias for p), and the contextual evidence always suggested that $\bar{p}$ might be true. Thus, there was a conflict between the epistemic and the evidential bias, which produces doubt or surprise in Katja. To dispel her doubt, Katja asks a question. In Experiment 1 the question contained the PPI *uže* 'already' and thus by hypothesis was an ON-question checking the epistemic bias. In Experiment 2 the question contained the NPI *eščë* 'yet' and thus by hypothesis was an IN-question double-checking the evidential bias. Dima's response consisted of a response particle (factor PARTICLE: *da*, *net*) and a response clause (where the subject was elided), which – depending on the question – contained *uže* or *eščë*. The response clause was always truthful: it reflected the actual state of affairs.

(18) SAMPLE ITEM

Dima i Katja gotovjatsja k komandirovke v Milan. Im pomogaet ich sekretar'– Marina Petrovna Mironova. *'Dima and Katja are preparing a business trip to Milan. Marina Petrovna Mironova, their secretary is helping them'.*

SOA DONE: Segodnja utrom Dima razgovarival s Mariej Petrovnoj i uznal, čto ona uže zabronirovala aviabilety. *'Dima talked to Marina Petrovna this morning and learned that she has already booked the tickets'.*

SOA NOT DONE: Segodnja utrom Dima razgovarival s Mariej Petrovnoj i uznal, čto ona budet bronirivat' aviabilety na sledujuščej nedele. *'This morning Dima talked to Marina Petrovna and learned that she would book the tickets next week'.*

Nezadolgo do okončanija rabočego dnja Dima i Katja obsuždajut predstojaščuju komandirovku. Katja uverena v tom, čto Marina Petrovna uže vsë organizovala i vremja vyleta uže izvestno. Poėtomu ona udivljaetsja, kogda Dima predlagaet letet' bolee rannim rejsom. *'Just before they go home, Dima and Katja are talking about the business trip. Katja assumes that Marina Petrovna has organized everything and that the departure time is fixed. So she is a little surprised when Dima suggests taking an earlier flight'.*

Katja:

ON-question, Experiment 1

Ljudmila Geist & Sophie Repp

> Razve Marina Petrovna uže   ne zabronirovala aviabilety?
> PART   Marina Petrovna already not booked   flight.tickets
> 'Hasn't Ms Miller already booked the tickets?'
>
> *IN-question, Experiment 2*
>
> Razve Marina Petrovna eščë ne zabronirovala aviabilety?
> PART   Marina Petrovna yet   not booked   flight.tickets
> 'Hasn't Ms Miller booked the tickets yet?'
>
> Dima:
>
> Net/ Da, uže   zabronirovala.
> no   yes already booked
> 'No/Yes, she has already booked the tickets.'   *Experiments 1, 2*
>
> Net/ Da, eščë ne   zabronirovala.
> no   yes yet   not booked
> 'No/Yes, she has not booked the tickets yet.'   *Experiments 1, 2*

Each experiment contained 24 lexicalizations in the four conditions just described. In addition to the experimental items, there were 24 lexicalizations which were very similar to the scenarios in the experimental items except that the question was positive and there was no bias. Otherwise they had the same 2 × 2 design. The fillers served mainly as control items and we will not discuss them here. The 48 lexicalizations were distributed over four lists in a Latin square design so that each list contained 24 experimental and 24 filler items. In addition, there were two practice items on each list.

The task of the participants was to judge the naturalness of the answer as a response to the question in view of the information described in the scenario. The judgment was given on a seven-point-scale with one scale end labelled *očen' estestvenno* 'very natural' and the other scale end *očen' stranno* 'very strange'. For the statistical analysis, these end points were transformed to the numbers 7 and 1, respectively, with the other scale points sitting in between. In addition to giving the acceptability judgment, participants verified a statement about the context, which was to ensure that they read the scenarios carefully. The verification statement was shown to the participants after they had read the test item and given the acceptability judgment.

The experiments were run as a web experiment on SoSci Survey (soscisurvey.de; Leiner 2021). For Experiment 1, 36 participants (28 female, 8 male; mean age: 35.3; age range: 29–54) with Russian as their native language were recruited via Prolific (www.prolific.co). For Experiment 2, 39 participants (30 female, 8 male, 1 unspecified; mean age: 37.5; age range: 20–56) were recruited.

Before taking part in the experiment, they gave informed consent. Due to the recruiting strategy via Prolific, we had not originally planned to conduct cross-experimental comparisons because we did not expect the same participants to take part in both experiments, which were conducted two weeks apart. As it turned out, 29 participants took part in both experiments. We decided to pool the data for these participants from both experiments for the statistical analysis because this allowed a direct comparison between the two question types. We discarded the data of the other participants.

To tackle the problem which recruiting participants via prolific brings about – the danger that most of the participants might be heritage speakers with potentially low levels of proficiency in Russian – we collected sociodemographic data of our participants. Of the 29 participants that took part in both experiments, 18 were born in Russia, 3 in Estonia, 3 in Latvia, 3 in the Ukraine, 1 in Moldavia, and 1 in Mongolia.[7] Almost all had also spent the longest part of their lives in these countries, except for two people born in Russia, who had spent most time in the Ukraine and in the UK, respectively, and one person from the Ukraine and one from Moldavia, who both had spent most time in the UK. We take these numbers to indicate that our participants are proficient Russian speakers, although we note that the age of one of the people having spent most time in the UK indicates a pre-adult move to the UK. We note that 26 participants reported to speak English on a daily basis, for one this was the case for French, and for one for Ukrainian. There were several other languages that were used less frequently.

4.2 Results

All 29 participants reached at least 80 percent correctness for the verification task so no participant was excluded on that criterion. The data from three participants were excluded from the analysis because they had not chosen the expected side of the naturalness scale in more than ten percent of the filler items, where the judgment for the use of *da* or *net* is unequivocal. This left 1248 data points for analysis. The analysis was conducted by fitting a cumulative link mixed model for ordinal data (R package ordinal, Christensen 2019). QUESTION TYPE (= Experiment), SOA

---

[7]We assigned participants that had indicated the Soviet Union as birth place to the respective post-Soviet countries. Russian is a widespread native language in all the above-mentioned countries, except Mongolia. None of participants born in Estonia, and Moldavia indicated that they speak Estonian or Moldavian. One person from Latvia speaks Latvian regularly but only several times per month. The person from Mongolia, and the Latvian person just mentioned were excluded from the statistics for poor performance on the control items (see §4.2) along with one other person.

Ljudmila Geist & Sophie Repp

and PARTICLE were fixed factors. They were sum-coded. Initially, participant and lexicalization were random factors. However, since the random effects of lexicalization produced models that were a singular fit, the final model only contained random intercepts and slopes for the experimental factors and their interaction per participant and not per lexicalization.

Figure 1 shows the results in terms of proportions of rating levels broken down for the experimental conditions including the median ratings per condition. Table 5 shows the model estimates. There were main effects of QUESTION TYPE (experiment) and of PARTICLE, and an interaction of PARTICLE and SOA.
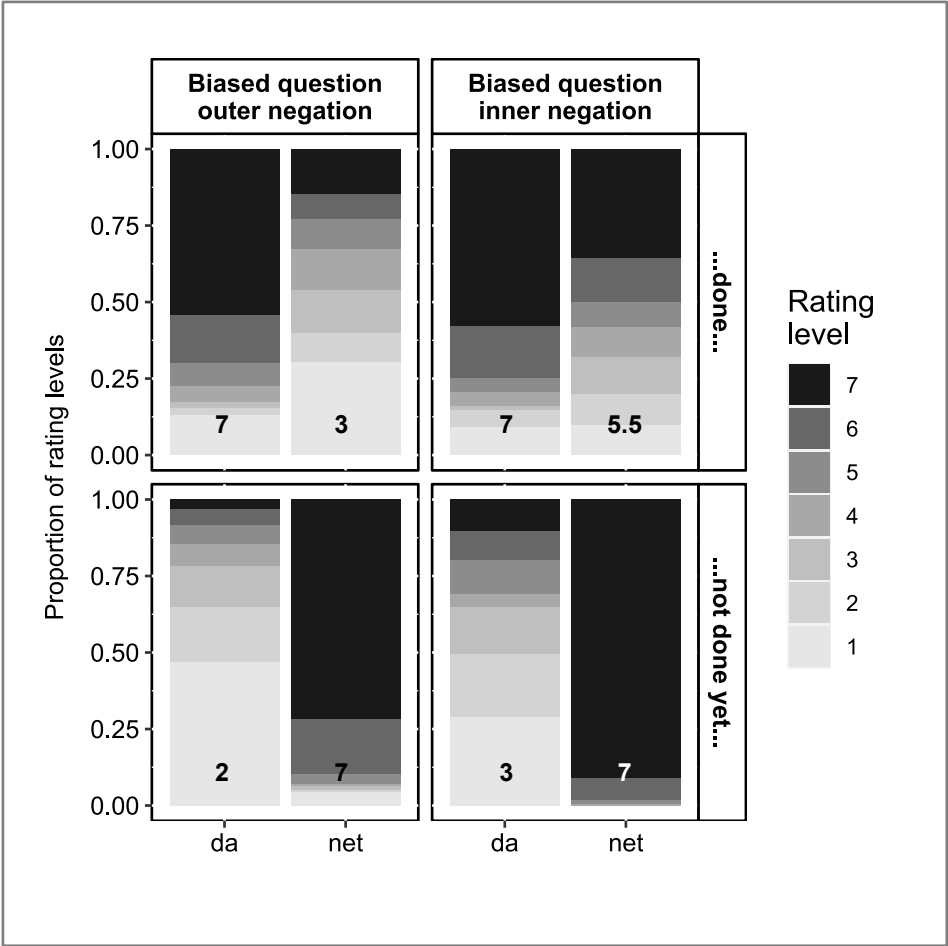

Figure 1: Proportions of rating levels for responses to ON/IN-questions. Numbers on the bars are the medians per condition

Table 5: Model estimates for the pooled data of both experiments

	Estimate	SE	z-value	p-value
QUESTION TYPE	0.62	0.25	2.52	0.012*
SOA	−0.02	0.16	−0.11	0.912
PARTICLE	1.04	0.29	3.53	< 0.001***
QUESTION TYPE × SOA	0.02	0.15	0.10	0.921
QUESTION TYPE × PARTICLE	0.28	0.15	1.84	0.065
PARTICLE × SOA	2.58	0.40	6.53	< 0.001***
QUESTION TYPE × SOA × PARTICLE	−0.16	0.17	−0.92	0.357

Overall, the particles were judged to be more natural after IN-questions, and *net* was more natural than *da*. We resolved the interaction PARTICLE × SOA by subsetting the data for each SoA. In the DONE context, *da* received higher ratings than *net* ($b = −1.43, SE = 0.51, z = −2.81, p = 0.005$). In the NOT DONE contexts, *net* received higher ratings than *da* ($b = 3.95, SE = 0.46, z = 8.51, p < 0.001$). Since QUESTION TYPE did not interact reliably with the other two factors, we take the effect of question type to be present in both SoAs and for both particles. Looking at the medians, however, the effect becomes particularly visible for *net* in the DONE contexts: After IN-questions *net* has a median in the scale part towards naturalness (median = 5.5) whereas after ON-questions *net* has a median that is in the scale part towards unnaturalness (median = 3). For *da* in NOT DONE contexts we observe only differences in the scale part toward unnaturalness: *da* is judged to be more unnatural after ON-questions (median = 2) than after IN-questions (median = 3).

Since previous research has found considerable inter-individual variation in the acceptability of response particles in various languages (Claus et al. 2017, Repp et al. 2019), we investigated this issue for our data. Figure 2 and Figure 3 show the variation for ON-questions and for IN-questions respectively. The figures indicate that the variation is fairly similar. In DONE contexts, the majority of participants judge *da* as natural (median 6 or 7), and as more natural than *net*. There are a few participants, however, who judge *net* more natural than *da*, and some who find neither particle natural after ON-questions (median below 6). In NOT DONE contexts, almost all participants find *net* natural whereas for *da* naturalness ratings vary considerably.

To better assess the difference between the two question types, we plotted the inter-individual variation in a way that allows us to directly compare par-

Ljudmila Geist & Sophie Repp
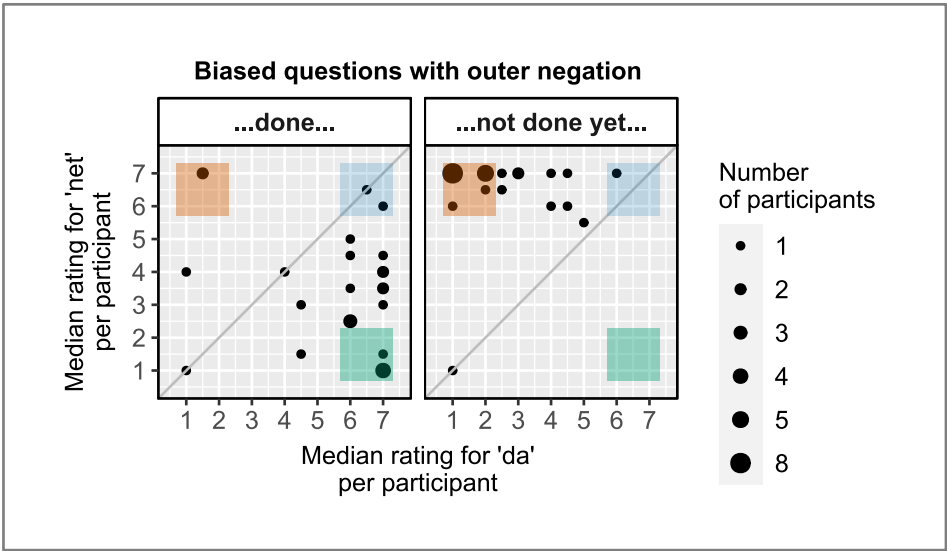
Figure 2: Inter-individual variation in responses to ON-questions. Dot size represents the number of participants with the same combination of median rating for *da* and median rating for *net* for the respective SoA. Dots in the orange box represent participants for whom *net* had a median of at least 6 and *da* had a median of maximum 2, i.e. for whom the difference between the particles was very pronounced. Dots in the green box represent participants for whom *da* had a median of at least 6 and *net* had a median of maximum 2. Dots in the blue box represent participants for whom both *da* and *net* had a median of at least 6.

ticipants' medians across question types, see Figure 4. Figure 4 has two facets which indicate differences between the question types: For *da* in the NOT DONE context, many dots are quite far away from the (perfect correlation) diagonal in both directions, which suggests that the speakers' judgments for the two question types differ in scale direction. For *net* in the DONE context, the dots are above the diagonal, which indicates generally higher ratings after IN-questions. Hence, we assume that there is a real difference for many speakers between the two question types here.

5 Discussion and Conclusion

Table 6 summarizes the results of our experiments in comparison to our predictions. Confirmed predictions are marked with ✓. Unpredicted results are marked with ✗. The table shows that many of our expectations were confirmed. Especially for ON-questions, our hypotheses seem to be on the right track: what is

224
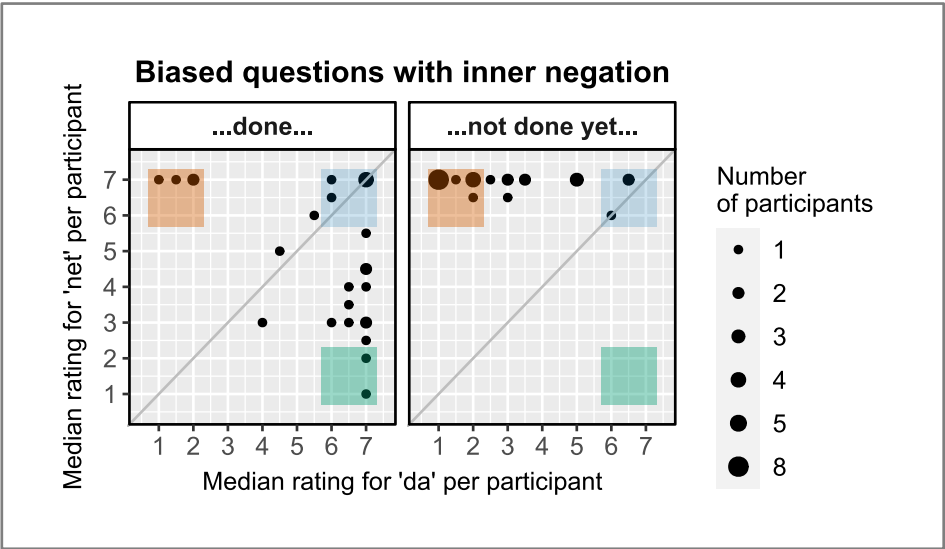
Figure 3: Inter-individual variation in responses to IN-questions. For the coding system, see caption of Figure 2

checked by an ON-question is a positive proposition p, and p is the proposition that serves as the antecedent for *da* and *net*. Accounts assuming an LF where ON-questions contain only a positive proposition can explain these findings. For IN-questions, we obtained several unexpected results, especially concerning *da*. We will discuss these in detail in what follows.

Table 6: Results and predictions

SoA	ON-question *Hasn't ... already...?* [Q [FALSUM p]]		IN-question *Hasn't ... yet...?* [Q [VERUM $\bar{p}$]]	
	Antecedent			
p (DONE)	[+, AGREE] → *da*	✓	[+, REVERSE] → *net*	✓ ✗ *da > net*
$\bar{p}$ (NOT DONE)	[[−], REVERSE] → *net* [−, REVERSE] → *net*	✓	[[−], AGREE] → *net* [−, AGREE] → *da* *net > da*	✓ ✗ ?? *da* ✓

The high acceptability of *da* in responses to IN-questions in DONE contexts

Ljudmila Geist & Sophie Repp
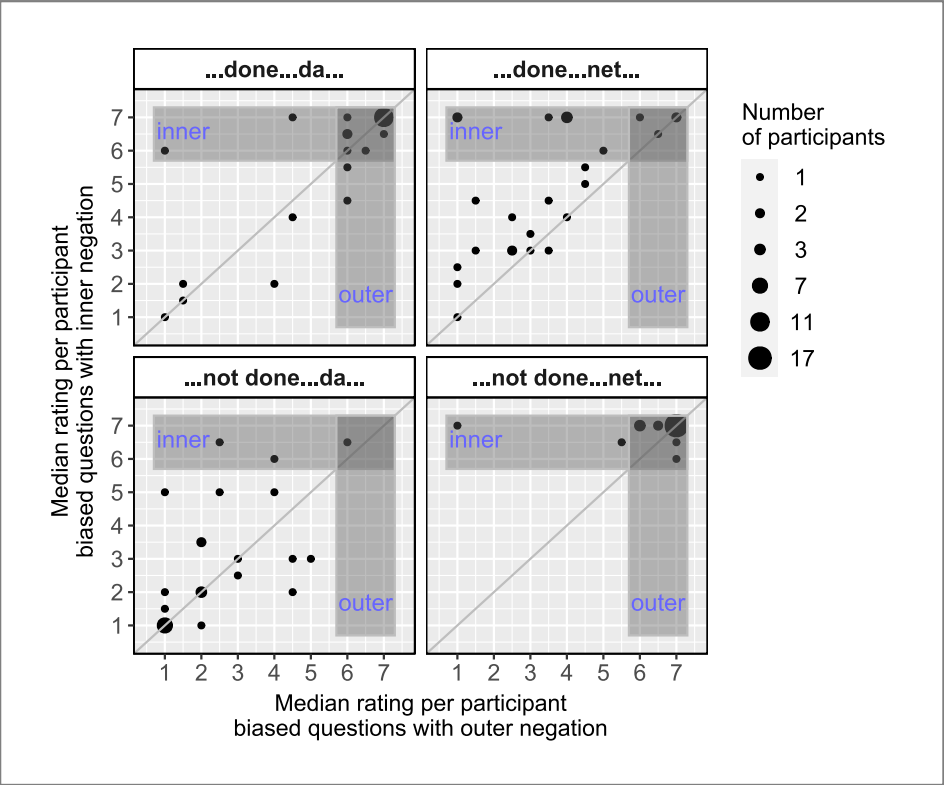
Figure 4: Median ratings per participant for ON- vs. IN-questions. Dot size represents the number of participants with the same combination of median rating for ON-questions and for IN-questions. Dots on the diagonal line represent participants that had the same ratings for both question types. Dots in the grey bars represent ratings of 6 or 7 for IN-questions (horizontal bar) or ON-questions (vertical bar) or both (overlap of bars).

(median = 7) is completely unexpected. Recall that *da* by hypothesis only realizes [AGREE], and an IN-question by hypothesis only makes the negative proposition $\bar{p}$ available. Since the response is supposed to express that p is true, the presupposition of [AGREE] is not met. We conclude from this finding that either *razve*-questions with *eščë* do not have the LF proposed for IN-questions by Romero & Han (2004) and Repp (2006, 2009), or the hypothesis for *da* that we developed on the basis of Esipova (2021) and González-Fuente et al. (2015) is wrong. A third avenue for explaining the result is re-investigating the salience of the various propositions and the role of the particle *razve*. We will discuss these three options for the DONE contexts and also consider the repercussions for the other contexts.

Regarding the potential conclusion that IN-questions do not have the assumed LF, there is a finding in our experiments that in our view speaks against it: *net* is fairly acceptable after IN-questions in DONE contexts (median = 5.5), in contrast to ON-questions (median = 3). Indeed, the median for *net* is on the acceptable scale end for IN-questions, which is not the case for ON-questions. This finding suggests that an IN-question does make $\bar{p}$ available, which can serve as the antecedent that is required for the presupposition of [REVERSE] in a DONE context: [REVERSE] is the feature that is realized by *net*.[8]

Regarding a different feature-particle mapping for *da*, we will consider two options: one makes the mapping more general, the other makes it more specific. Starting with the more general one, we could assume that instead of [AGREE] → *da*, the mapping is [+], [AGREE] → *da*, i.e. *da* may realize [AGREE] as well as [+], just like English *yes*. This could explain the high ratings in the DONE context in IN-questions in the following way. If in Russian the constraint REALIZE ABSOLUTE FEATURES has a considerably higher weight than MAXIMIZE MARKED FEATURES and than REALIZE RELATIVE FEATURES, the observed preference for *da* over *net* in DONE contexts is explained: *da* realizes absolute, unmarked [+], *net* realizes relative, marked [REVERSE]. This assumption could also explain the low ratings for *da* after IN-questions in NOT DONE contexts (median = 3), where *da* realizes relative, unmarked [AGREE], whereas *net* (median = 7) realizes absolute, marked [−]. However, there also is a problem. Recall from §3.2 that Esipova (2021) claims that *da* cannot be used in [+, REVERSE] contexts after negative assertions, see (17d) above. This claim is fully confirmed by experimental findings in Repp & Geist (2022). So assuming that *da* can realize [+] seems to be on the wrong track

---

[8]Note that the high acceptability of *net* in a DONE context does not parallel Meyer's (2004) empirical claims about unbiased questions: in Meyer's example, *net* is unacceptable in this context.

Ljudmila Geist & Sophie Repp

because of substantial empirical differences between IN-questions and negative assertions as antecedents. We will return to this issue further below.

The more specific feature-particle mapping that is a promising candidate to explain our findings is: [+, AGREE] → *da*. Here, we would have to assume that the presupposition of [AGREE] is fulfilled in IN-questions by the presence of the (less salient) epistemic bias *p*, which – recall our discussion in §2.1 – is an integral part of biased ON/IN-questions although this is not reflected in the LF of IN-questions. If *da* realizes [+, AGREE], a high weighting of EXPRESSIVENESS will ensure the preference of *da* over *net* because *da* realizes more features than *net* does. This more specific feature-particle mapping would also be able to explain why *da* is quite unacceptable (median = 3) as a response to IN-questions in NOT DONE contexts: *da* cannot express [AGREE] if the response clause is a negative proposition. However, the more specific feature-particle mapping also faces the problem that there is a difference with previous findings for assertions. Recall from §3.2 that Esipova (2021) claims that *da* is acceptable in NOT DONE contexts if the antecedent is a negative assertion, see (17b) above – the answer with the features [−, AGREE]. Repp & Geist (2022) present experimental evidence supporting this claim, at least to some extent.

Regarding the salience assumptions, we could also take a more drastic step and assume that the epistemic bias *p* is made very salient by the interrogative particle *razve*, so that *p* is more salient than the evidential bias *p̄*, which is part of the LF of IN-questions. On this assumption, we would not have to alter the feature-particle mapping [AGREE] for *da* because after IN-questions in DONE contexts *da* just picks up the more salient proposition *p* and therefore is more acceptable than *net* (median 7 vs. 5.5). After IN-questions in NOT DONE contexts, *da* is expected to be unacceptable because signalling the same polarity of epistemic bias and response does not express the intended meaning *p̄*. To test the relative salience of the biases in *razve*-questions, follow-up studies with other interrogative particles are needed. Note, however, that the sketched salience account essentially assumes the same salience differences between *p* and *p̄* in IN- and ON-questions, so that subtle differences between the question types – for instance in responses with *net* – cannot be explained.

An anonymous reviewer suggests that by using *da* the speaker indicates agreement with the interlocutor's epistemic bias independently of salience considerations. This proposal could indeed explain the patterns for ON- and IN-questions for *da*, because for *da* the difference does not seem to matter (a lot). It would also be compatible with the observation that *da* can be used to signal agreement with a negative assertion (Esipova 2021, Repp & Geist 2022), because asserting *p̄* plausibly presupposes having a bias for *p̄*. Finally, this proposal would also be com-

patible with the observation in Repp & Geist (2022) that *da* is clearly degraded in responses to unbiased negative questions, independently of the response polarity (see §3.2). However, intuitively, *da* seems to be the appropriate answer to a positive question with *razve*, like B in (4) in §2.2, if the response polarity is positive:

(19) B: A   razve on ženat?
          but PART he married
          'But is he really married?'

      A: Da, on ženat.
          yes he married
          'Yes, he is married.'

As laid out in §2.2, the epistemic bias of B in this example is $\bar{p}$. A does not agree with this bias, but with the evidential bias. The evidential bias is the bias that arguably is made salient by the question.

In the final part of this discussion, we will sketch a way to reconcile the observed differences between questions and assertions as antecedents. We think that these differences can only be explained on the assumption that *da* is ambiguous, and that the ambiguity must involve a presupposition regarding the type of antecedent. At present we cannot decide between the mappings that we discussed to account for our results for ON/IN-questions, [+, AGREE] → *da*, or [+], [AGREE] → *da*. The former has the advantage that it is more parsimonious in the overall setup because there will be less ambiguity, but the choice is an empirical question that must be addressed in future research.

Our new proposal is that *da* also can realize a feature that we will call [ACCEPT]. (20) gives the presupposition of [ACCEPT] in abbreviated form. It contains an illocutionary component: the conversational table (Farkas & Bruce 2010).

(20) [ACCEPT] presupposes the existence of a single proposition on the conversational table, which has the same polarity as the response clause.

(20) shows that [ACCEPT] is sensitive to how many propositions there are on the table. We have no space to discuss this here but we assume that questions place a set of propositions on the table, which might be more or less salient, and it is up to the addressee to decide which proposition enters the common ground (if any). Assertions place only one proposition on the table. Roelofsen & Farkas (2015) emphasize that for any anaphor, including response particles, there must be a *unique salient* antecedent in the context. The presupposition in (20) is stricter than that: it allows only one proposition on the table at all, irrespective of

the non-salience of potential other propositions. Assuming that a constraint like MAXIMIZE PRESUPPOSITION! (Heim 1991) is generally operative, [ACCEPT] will be the feature that is relevant in responses to assertions. In responses to questions there will be a presupposition failure for [ACCEPT], so that (one of) the other feature-particle mapping(s) for *da* applies (depending on the answers regarding the future research questions above, $[+, \text{AGREE}] \rightarrow da$ or $[+], [\text{AGREE}] \rightarrow da$).

We are not the first to suggest that questions and assertions receive different responses. Holmberg (2015) has made suggestions along these lines for English. Similarly, Repp et al. (2022) propose for German that *nein* 'no' is used to express a counterpart of [ACCEPT] in responses to assertions, namely [REJECT]. The observed differences require much more quantitative empirical research, also because there is substantial inter-individual variation, as we could also verify for Russian.

Overall, our investigation has shown that the answer patterns for Russian *da/net* differ depending on whether the antecedent is an IN-question or an ON-question. We have also discussed some differences with assertions, which, however, were not the focus of the present study. On the basis of our findings, we assume that *da* and *net* are sensitive to the interpretation of the negation in biased questions with *razve*, as it is indicated by the polarity-sensitive items *uže* and *eščë*. The account of inner vs. outer negation in terms of propositional negation vs. the illocutionary operator FALSUM goes some way to explaining the answer patterns for these questions. However, we also saw that we might have to make additional assumptions concerning the salience of a bias that is not double-checked. This is an issue that needs further attention in future research as it poses interesting empirical and theoretical challenges. Specifically, we need to find out more about potential differences in salience between epistemic bias and evidential bias. After all, the evidential bias for $\bar{p}$ does not seem to play a role for responses to ON-questions. Furthermore, we need a model that integrates the biases in a more explicit way, which explains how they become part of the discourse representation.

Acknowledgments

References

Apresjan, Jurij Derenikovič. 1980. Tipy informacii dlja poverchnostno-semantičeskogo komponenta modeli 'Smysl – Text'. *Wiener Slawistischer Almanach* 1. 1–119.

Arnhold, Anja, Bettina Braun & Maribel Romero. 2021. Aren't prosody and syntax marking bias in questions? *Language and Speech* 64(1). 141–180. DOI: 10.1177/0023830920914315.

Baranov, Anatolij Nikolaevič. 1986. 'Predpoloženie' vs 'fakt': *Neuželi* vs *razve. Zeitschrift für Slavistik* 31(1). 119–131.

Baranov, Anatolij Nikolaevič & Irina Michajlovna Kobozeva. 1983. Semantika obščich voprosov v russkom jazyke (kategorija ustanovki). *Izvestija AN SSSR. Ser. lit. i jaz* 42(3). 263–274.

Boguslavskij, Igor' Mihajlovič. 1996. *Sfera dejstvija leksicheskich edinic.* Moskva: Škola "jazyki russkoj kul'tury".

Brown, Sue. 1999. *The syntax of negation in Russian: A minimalist approach.* Stanford, CA: CSLI Publications.

Brown, Sue & Steven Franks. 1995. Asymmetries in the scope of Russian negation. *Journal of Slavic Linguistics* 3(2). 239–287. https://www.jstor.org/stable/24598950.

Bryzgunova, Elena Andreevna. 1975. The declarative–interrogative opposition in Russian. *Slavic and East European Journal* 19(2). 155–161. DOI: 10.2307/306768.

Bryzgunova, Elena Andreevna. 1980. Intonacija. In Natalija Jul'evna Švedova et al. (eds.), *Russkaja grammatika*, 96–122. Moskva: Akademija nauk.

Büring, Daniel & Christine Gunlogson. 2000. Aren't positive and negative polar questions the same? Manuscript. UCLA and UCSC. https://semanticsarchive.net/Archive/mYwOGNhO/polar_questions.pdf.

Christensen, Rune Haubo Bojesen. 2019. *Package ordinal. Regression models for ordinal data.* Version 2012.09-11. CRAN Repository. https://github.com/runehaubo/ordinal.

Claus, Berry, Felix Frühauf, A. Marlijn Meijer & Manfred Krifka. 2016. Particle responses to negative polar questions with high vs low negation. Poster presented at the *Annual Meeting Priority Program XPrag.de*, June 2016.

Ljudmila Geist & Sophie Repp

Claus, Berry, A. Marlijn Meijer, Sophie Repp & Manfred Krifka. 2017. Puzzling response particles: An experimental study on the German answering system. *Semantics and Pragmatics* 10. 1–51. DOI: 10.3765/sp.10.19.

Esipova, Maria. 2021. Polar responses in Russian across modalities and across interfaces. *Journal of Slavic Linguistics* 29(FASL 28 extra issue). 1–11. https://ojs.ung.si/index.php/JSL/article/view/151.

Farkas, Donka & Floris Roelofsen. 2019. Polarity particles revisited. *Semantics and Pragmatics* 12(15). 1–16. DOI: 10.3765/sp.12.15.

Farkas, Donka F. & Kim B. Bruce. 2010. On reacting to assertions and polar questions. *Journal of Semantics* 27(1). 81–118. DOI: 10.1093/jos/ffp010.

González-Fuente, Santiago, Susagna Tubau, M. Teresa Espinal & Pilar Prieto. 2015. Is there a universal answering strategy for rejecting negative propositions? typological evidence on the use of prosody and gesture. *Frontiers in Psychology* 6. 293–316. DOI: 10.3389/fpsyg.2015.00899.

Gärtner, Hans-Martin & Beáta Gyuris. 2017. On delimiting the space of bias profiles for polar interrogatives. *Linguistische Berichte* 251. 293–316.

Goodhue, Daniel & Michael Wagner. 2018. Intonation, *yes* and *no. Glossa: A Journal of General Linguistics* 3(1). DOI: 10.5334/gjgl.210.

Gyuris, Beáta. 2017. New perspectives on bias in polar questions: A study of Hungarian *-e. International Review of Pragmatics* 9(1). 1–50. DOI: 10.1163/18773109-00000003.

Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 10(1). 41–53. https://www.jstor.org/stable/25000703.

Heim, Irene. 1991. Artikel und definitheit [Articles and definiteness]. In Armin v. Stechow & Dieter Wunderlich (eds.), *Semantik/semantics*, 487–535. Berlin: de Gruyter.

Höhle, Tilmann. 1988. VERUM-Fokus. *Sprache und Pragmatik* 5. 2–7.

Höhle, Tilmann. 1992. Über verum Fokus in Deutschen. *Linguistische Berichte Sonderheft* 4. 112–141.

Holmberg, Anders. 2013. The syntax of answers to polar questions in English and Swedish. *Lingua* 128. 31–50. DOI: 10.1016/j.lingua.2012.10.018.

Holmberg, Anders. 2015. *The syntax of* yes *and* no. Oxford: Oxford University Press.

Jones, Bob Morris. 1999. *The Welsh answering system.* Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110800593.

Kirschbaum, Ernst-Georg. 2001. *Grammatik der russischen Sprache.* Berlin: Cornelsen Volk und Wissen Verlag.

Kobozeva, Irina Michajlovna. 2004. *Lingvističeskaja semantika.* Moskva: URSS.

Kramer, Ruth & Kyle Rawlins. 2011. Polarity particles: an ellipsis account. In Suzi Lima, Kevin Mullin & Brian Smith (eds.), *NELS 39: Proceedings of the 39th Annual Meeting of the North East Linguistic Society*, vol. 1, 479–491. Amherst, MA: GLSA.

Krifka, Manfred. 2013. Response particles as propositional anaphors. In Todd Snider (ed.), *SALT 23: Proceedings from the 23rd Conference on Semantics and Linguistic Theory*, 1–18. Ithaca, NY: CLC Publications. DOI: 10.3765/salt.v23i0.2676.

Krifka, Manfred. 2015. Bias in commitment space semantics: Declarative questions, negated questions, and question tags. In Sarah D'Antonio, Mary Moroney & Carol Rose Little (eds.), *Semantics and Linguistic Theory (SALT) 25*, 328–345. Washington, D.C.: LSA Open Journal Systems. DOI: 10.3765/salt.v25i0.3078.

Ladd, D. Robert. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In Roberta A. Hendrick, Carrie S. Masek & Mary Frances Miller (eds.), *Proceedings of the Chicago Linguistic Society 17*, 164–171. Chicago, IL: CLS.

Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511808814.

Leiner, Dominik J. 2021. *Sosci survey (version 3.2.31) [computer software]*. Available at https://www.soscisurvey.de.

Li, Feifei, Santiago González-Fuente, Pilar Prieto & M.Teresa Espinal. 2016. Is Mandarin Chinese a truthbased language? Rejecting responses to negative assertions and questions. *Frontiers in Psychology* 7. DOI: 10.3389/fpsyg.2016.01967.

Loos, Cornelia, Markus Steinbach & Sophie Repp. 2020. Affirming and rejecting assertions in German Sign Language (DGS). In Michael Franke, Nikola Kompa, Mingya Liu, Jutta L. Mueller & Juliane Schwab (eds.), *Proceedings of Sinn und Bedeutung 24*, vol. 2, 1–19. Osnabrück University / Humboldt-Universität zu Berlin. DOI: 10.18148/sub/2020.v24i2.883.

Mat'ko, Irina Dmitrievna. 2014. K voprosu o pragmatičeskich ėkvivalentach: Perevod russkich voprositel'nych častic neuželi / razve na francuzskij jazyk. *Vesnik Hrodzenskaha Dzjaržaŭnaha Universitėta Imja Janki Kupaly* 2(179). 40–50.

Meijer, Marlijn A., Berry Claus, Sophie Repp & Manfred Krifka. 2015. Particle responses to negated assertions: preference patterns for German *ja* and *nein*. In Thomas Brochhagen, Floris Roelofsen & Nadine Theiler (eds.), *Proceedings of the 20th Amsterdam Colloquium*, 286–295. Amsterdam: Institute for Language, Logic & Computation. Universität Amsterdam.

Meyer, Roland. 2004. Prosody, mood, and focus. A study of so-called 'intonationally marked' yes-no questions in Russian. In M. Rivero, O. Arnaudova & D. Stojanović (eds.), *Formal Approaches to Slavic Linguistics 12. The Ottawa Meeting*, 333–352.

Meyer, Roland & Ina Mleinek. 2006. How prosody signals force and focus – A study of pitch accents in Russian yes–no questions. *Journal of Pragmatics* 38(10). 1615–1635. DOI: 10.1016/j.pragma.2005.05.011.

Pančenko, Julija Dmitrievna. 2021. *Otvety na obščie voprosy s otricaniem v russkom jazyke (ėksperimental'noe issledovanie raznych vozrastnych grupp)*. Moskovskij gosudarstvennyj universitet im. M.V. Lomonosova. Bachelor thesis.

Pope, Emily. 1976. *Questions and answers in English*. The Hague: Mouton.

Rachilina, Ekaterina V. 2008. Nacional'nyj korpus russkogo jazyka: začem on slavistam? *Die Welt der Slaven* 1(LIII). 1–10.

Rathmayr, Renate. 1985. *Die russischen Partikeln als Pragmalexeme*. München: Otto Sagner. DOI: 10.3726/b12537.

Repp, Sophie. 2006. ¬(A&B). Gapping, negation and speech act operators. *Research on Language and Computation* 4(4). 397–423. DOI: 10.1007/s11168-006-9022-y.

Repp, Sophie. 2009. *Negation in gapping*. Oxford: Oxford University Press.

Repp, Sophie. 2013. Common ground management: Modal particles, illocutionary negation and VERUM. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives: explorations in use-conditional meaning*, 231–274. Leiden: Emerald.

Repp, Sophie, Berry Claus & Felix Frühauf. 2022. Responses to biased questions in German. Manuscript University of Cologne.

Repp, Sophie & Ljudmila Geist. 2022. Polar responses to negative antecedents in Russian: assertions and biased questions. Manuscript University of Cologne, University of Stuttgart.

Repp, Sophie & Ljudmila Geist. to appear. Negative polar questions in Russian: question bias and question concern. In Anton Benz, Manfred Krifka, Tue Trinh & Kazuko Yatsushiro (eds.), *Perspectives on biased questions*. Berlin: Language Science Press.

Repp, Sophie, A. Marlijn Meijer & Nathalie Scherf. 2019. Responding to negative assertions in Germanic: On *yes* and *no* in English, Dutch and Swedish. In M. Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung 23*, vol. 2, 267–285. Bellaterra (Cerdanyola del Vallès): Universitat Autònoma de Barcelona. DOI: 10.18148/sub/2019.v23i2.611.

Restan, Per. 1972. *Sintaksis voprositel'nogo predloženija.* Oslo: Universitetsforlaget.

Roelofsen, Floris & Donka F. Farkas. 2015. Polarity particle responses as a window onto the interpretation of questions and assertions. *Language* 91(2). 359–414. https://www.jstor.org/stable/24672234.

Roelofsen, Floris & Sam van Gool. 2010. Disjunctive questions, intonation, and highlighting. In Maria Aloni, Harald Bastiaanse, Tiktu de Jager & Katrin Schulz (eds.), *Proceedings of the 17th Amsterdam Colloquium*, 384–394. Berlin: Springer. DOI: 10.1007/978-3-642-14287-1_39.

Romero, Maribel. 2015. High negation in subjunctive conditionals and polar questions. In Eva Csipak & Hedde Zeijlstra (eds.), *Proceedings of Sinn und Bedeutung 19*, 519–536. DOI: 10.18148/sub/2015.v19i0.247.

Romero, Maribel. 2020. Form and function of negative, tag, and rhetorical questions. In Viviane Déprez & M. Teresa Espinal (eds.), *The Oxford handbook of negation*, 235–254. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780198830528.013.8.

Romero, Maribel & Chung-Hye Han. 2004. On negative yes/no questions. *Linguistics and Philosophy* 27(5). 609–658. DOI: 10.1023/B:LING.0000033850.15705.94.

Šatunovskij, Il'ja Borisovič. 2005. Osnovnye kognitivno-kommunikativnye tipy obščich voprosov v russkom jazyke. In I. M. Kobozeva, A. S. Narin'jani & V. P. Selegej (eds.), *Dialog: Komp'juternaja lingvistika i intellektual'nye technologii*, 502–506. Moskva: Nauka.

Seeliger, Heiko. 2015. 'Surely that's not a negative declarative question?' Polar discourses in Swedish, German and English. In Eva Csipak & Hedde Zeijlstra (eds.), *Proceedings of Sinn und Bedeutung 19*, 591–609. DOI: 10.18148/sub/2015.v19i0.251.

Seeliger, Heiko. 2019. *Swedish and German rejecting questions.* Berlin: Humboldt-Universität zu Berlin. (Doctoral dissertation). DOI: 10.18452/19686.

Seeliger, Heiko & Sophie Repp. 2018. Biased declarative questions in Swedish and German: Negation meets modal particles (*väl*, and *doch wohl*). In Christine Dimroth & Stefan Sudhoff (eds.), *The grammatical realization of polarity contrast. Theoretical, empirical and typological approaches*, 129–172. Amsterdam: Benjamins. DOI: 10.1075/la.249.05see.

Sudo, Yasutada. 2013. Biased polar questions in English and Japanese. In Daniel Gutzman & Hans Martin Gärtner (eds.), *Beyond expressives: explorations in use-conditional meaning*, 1–58. Leiden: Brill. DOI: 10.1163/9789004183988_009.

Švedova, Natalija Jur'evna et al. 2005. *Russkaja grammatika.* Vol II. Reprint 1980. Moskva: RAN.

Ljudmila Geist & Sophie Repp

van der Wouden, Ton. 1997. *Negative contexts: collocation, polarity and multiple negation.* London/New York: Routledge. DOI: 10.4324/9780203437636.

Zaliznjak, Anna. 2020. Russkoje *razve*: ot predloga k voprositel'noj častice. *Izvestija Rossijskoj akademii nauk. Serija literatury i jazyka* 79(4). 5–11.

Chapter 9

Phi-congruence and case agreement in close apposition in Russian

Ora Matushansky

SFL (CNRS/Université Paris-8/UPL)/ILS (Utrecht University)

I demonstrate that case-marking on the proper name in close apposition in Russian depends on two factors: the semantic sort of the proper name (where object-denoting proper names differ from place-denoting proper names, i.e., toponyms) and within the latter category on the lexical-semantic class of the toponym: major landmarks, such as cities and countries, special landmarks (rivers, streets, etc.) and the rest. While animate proper names necessarily agree in case with their sortals and inanimate ones obligatorily appear in the nominative case, case agreement with toponyms is conditioned by phi-congruence: cities and countries require number congruence, special landmarks need gender congruence and for the residue only phi-congruent adjectival toponyms may agree in case. I suggest that the phi-congruence condition should be analyzed as semantic agreement and hypothesize that toponyms differ from object-denoting proper names in that the former may have interpretable phi-features despite being inanimate.

Keywords: case, semantic agreement, close apposition, Russian, phi-features

1 Introduction

Proper names in Russian are divided into three categories in function of their case-marking in close apposition: those that must agree in case with the sortal (animates), those that can agree (toponyms) and those that cannot agree (the rest). Within the second category the possibility of case agreement is conditioned by PHI-CONGRUENCE: the values of certain phi-features of the toponym must match those of the sortal. The question arises how to model these facts and what they tell us about the nature of agreement.

Ora Matushansky

I begin with the presentation of the phenomenon of close apposition in general (§2). In §3 I will describe what is known about the empirical landscape of case-agreement in close apposition in Russian, focusing on toponyms and the impact of phi-congruence. §4 will deal with and dismiss several possible analyses of these facts. §5 is dedicated to a sketch of a proposal, linking case agreement to semantic agreement. §6 concludes.

2 Close apposition

Appositions can be defined as a single constituent containing more than one NP yet only one referent. The obvious difference between close apposition in (1) and loose apposition in (2) is that in (1) there is no intervening pause and in (2) the proper name or kind name alone refers to the same individual as the NP combining with it:

(1) a. the element engoopium
     b. the material polyacrynilate
     c. the actor John Gielgud   (Jackendoff 1984)

(2) a. This element, engoopium, was invented by Ray Jackendoff.
     b. The prima/Maria Callas, the best Carmen ever, outsings everyone in this role.

In the type of close apposition exemplified in (1) the first noun (henceforth, THE SORTAL) is the syntactic head (Jackendoff 1984, Lasersohn 1986, McCawley 1996, 1998; contra Haugen 1953, Burton-Roberts 1975, Noailly 1991, Keizer 2005), as shown by the fact that agreement is determined by the phi-features of the sortal rather than by those of the proper name (or the second noun), as in the Russian example (3), and that the case assigned to the NP as a whole must surface on the sortal (and may, on the proper name), as in (4): [1,2]

---

[1] We set aside here several other types of close apposition, such as *Francis Bacon the philosopher* (restrictive, picking out one of the possible name bearers), *Karl Marx the Jew* (singling out a particular guise or aspect of an individual) or other marginal instances where NP$_2$ is headed by a common noun and contains an overt determiner, since those of them that can be reliably translated into Russian are all animate and exhibit obligatory case agreement.

[2] The first generalization seems to be contradicted by animate proper names, where the sortal may be masculine while the proper name (and the referent) is feminine, as in *doktor Liza*, which triggers feminine agreement. This contradiction is only apparent, as human-denoting nouns in Russian can agree semantically (Corbett 1979, see also Footnote 19).

(3) Krejser "Avrora" ne {streljal / *streljala}.
     cruiser.M Aurora.F NEG fired.M.SG fired.F.SG
     'The cruiser Aurora was not firing.'   agreement

(4) na {ulice   / *ulica}   {Jakimanke   /
     in street.F.SG.LOC street.F.SG.NOM Yakimanka.F.SG.LOC
     Jakimanka}
     Yakimanka.F.SG.NOM
     'on the Yakimanka street'   case

The fact that the sortal NP may contain a complement ((5), after McCawley 1998: 473) shows that the proper name, which is clearly not a semantic argument of the sortal anyway, must be treated as a modifier (Figure 1).

(5) byvšij prezident SŠA   i   gollivudskaja kinozvezda Ronal'd
     former president USA.GEN and Hollywood.ADJ movie star Ronald
     Rejgan
     Reagan
     'the former president of the US and Hollywood star Ronald Reagan'
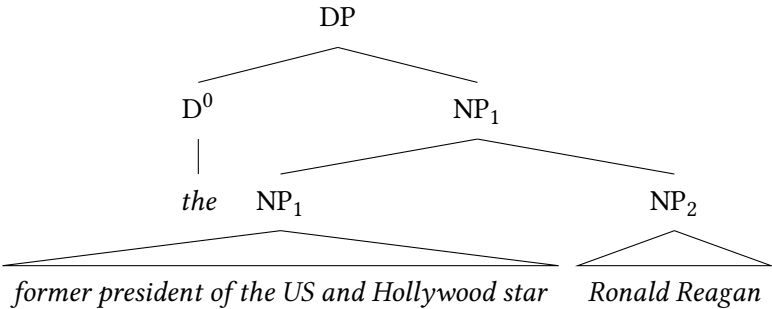
Figure 1: Close apposition: internal structure

Case-marking in close apposition (4) not only offers insight into its internal structure, but also suggests that agreement can occur between two noun phrases, as the choice of case can be conditioned by phi-congruence of the two nouns: having the same values for the gender and number features on the proper name as on the sortal may be a necessary condition for having an agreeing case on the proper name.

Ora Matushansky

3 Case-marking in close apposition with proper names

As (4) shows, close apposition permits two options for the proper name: the proper name can bear either the same case as the sortal or the default nominative case. The availability of either option depends on the lexical-semantic class of the proper name. Three broad groups can be established:

(6) animate referents: obligatory case agreement

    a. o   russkom   poėte   {Bloke   /
       about Russian.M.SG.LOC poet.M.SG.LOC Blok.M.SG.LOC
       *Blok}
       Blok.M.SG.NOM
       'about the Russian poet Blok'   [+animate]

    b. o   russkom   poėte   {Cvetaevoj   /
       about Russian.M.SG.LOC poet.M.SG.LOC Tsvetaeva.F.SG.LOC
       *Cvetaeva}
       Tsvetaeva.F.SG.NOM
       'about the Russian poet Tsvetaeva'

(7) non-toponymic proper names: forbidden case agreement

    a. s   familiej   {Blok   / *Blokom}
       with surname.INS Blok.NOM   Blok.INS
       'with the surname Blok'

    b. o   krejsere   {"Moskva"   / *"Moskve"}
       about cruiser.M.SG.LOC Moscow.F.SG.NOM   Moscow.F.SG.LOC
       'about the cruiser Moscow'

(8) toponyms: case agreement restricted by phi-feature congruence

    a. na ulice   {Jakimanka   / Jakimanke}
       in street.F.SG.LOC Yakimanka.F.SG.NOM Yakimanka.F.SG.LOC
       'on the Yakimanka street'   ✓ phi-congruent

    b. na ulice   {Balčug   / *Balčuge}
       in street.F.SG.LOC Balčug.M.SG.NOM   Balčug.M.SG.LOC
       'on the Balčug street'   ✗ phi-congruent

Close apposition is also possible with kind names, as in (1b)–(1c) and (9). The eight native speakers I asked split fifty-fifty as to which variant they accept and

no one has accepted both, so kind names seem to pattern either as city/country names or as non-toponymic names.[3]

(9) uroven' gormona   {%kortizol   / %kortizola}
    level   hormone.GEN   cortisol.NOM   cortisol.GEN
    'the level of the hormone cortisol'

While with animate referents (6) non-agreeing case on the proper name is disallowed in close apposition, with non-toponymic proper names (7) nominative is required on the proper name. Finally, for the third category, which only contains toponyms, both options are possible and, as (8) shows, the availability of the agreeing option is conditioned by their phi-features.

The focus of this paper is on case agreement for toponyms, which has been shown to depend on phi-congruence, i.e., on whether the sortal and the proper name match in phi-features. In addition to prescriptivist works like Rozental et al. (1998), two corpus studies, Graudina et al. (1976) and Logvinova (2018), show that within the broad category of toponyms different lexical-semantic classes can be distinguished in function of whether they require matching only in number or also in gender. After having examined the empirical picture provided by these works and discussed which deviations from these patterns are possible and why,[4] I will argue (§4) that the first hypotheses that come to mind cannot account for them and then advance an approach based on the assumption that phi-congruence enables semantic agreement (§5).

---

[3]Month names allow only the odd reversed construction in (i) with obligatory agreement. Sortals are not used with days of the week or event names (like *WWII*), perhaps because these proper names refer unambiguously and so a sortal is pragmatically infelicitous. Finally, holiday names allow the appositive oblique, as in (ii), as do the names of galaxies, constellations, and certain others, see Logvinova (2018, in press).

   (i) v   marte   mesjace
       in March.LOC month.LOC
       'in the month of March'

   (ii) prazdnik Pasxi
       holiday   Easter.GEN
       'the Holiday of Easter'

[4]Most of the generalizations below come from Graudina et al. (1976) and Rozental et al. (1998) and are verified by Logvinova (2018, in press). Deviations from and extensions of the patterns described there have been cross-checked in the National Russian Language Corpus (RNC 2003—2022), on Google, and with some native speakers.

Ora Matushansky

3.1 Number congruence and optional case agreement

The most permissive category of toponyms are proper names introduced by the sortals *gorod* 'city, town' (M), *stolica* 'capital' (F) and *strana* 'country' (F) (although not the coextensional *gosudarstvo* 'state' (N)), where agreeing and non-agreeing cases can be in free variation with no obvious difference in interpretation. Yet the phi-feature specification of the proper name is relevant for case-agreement, as can be seen from morphologically plural proper names. While both masculine and feminine city and country names generally allow case agreement (10), plural ones, as in (11), do not (Graudina et al. 1976: 141, Rozental et al. 1998: 281, confirmed by Logvinova 2018: 25-28, in press; the same is true for Ukrainian Gorpinič 1987):[5]

(10) a. v gorode   {Moskva   / Moskve}
        in city.M.SG.LOC Moscow.F.SG.NOM Moscow.F.SG.LOC
        'in the city of Moscow'   singular sortal, feminine PN

     b. v gorode   {Tallinn   / Tallinne}
        in city.M.SG.LOC Tallinn.M.SG.NOM Tallinn.M.SG.LOC
        'in the city of Tallinn'   singular sortal, masculine PN

     c. o   strane   {Francija   / Francii}
        about country.F.SG.LOC France.F.SG.NOM France.F.SG.LOC
        'about the great country France'   feminine sortal, feminine PN

     d. o   strane   {Kitaj   / Kitaje}
        about country.F.SG.LOC China.M.SG.NOM China.M.SG.LOC
        'about the great country China'   feminine sortal, masculine PN

(11) a. v gorode   {Gagry   / *Gagrax}
        in city.M.SG.LOC Gagra.PL.NOM   Gagra.PL.LOC
        'in the city of Gagra'   singular sortal, plural PN

     b. v gorode   {Velikie Luki   / *Velikix Lukax}
        in city.M.SG.LOC Velikie Luki.PL.NOM Velikie Luki.PL.LOC
        'in the city of Velikie Luki'   singular sortal, complex plural PN

A caveat should be introduced here. The non-agreeing pattern is an innovation in the history of Russian and is anecdotally taken to have arisen as a response to the

---

[5]For the sake of simplicity I avoid neuter toponyms, as these tend to behave as indeclinables, appearing in the nominative even without a sortal (Graudina et al. 1976: 138–140). The neuter sortals *selo* 'village' and *gosudarstvo* 'state' avoid case agreement even with phi-congruent toponyms, though the former allows it with phi-congruent adjectival proper nouns (there exist no adjectival state names).

logistical challenges of WWI, when the use of the nominative form after a sortal could distinguish one location from another. Prior to that time the preference was for case agreement between the sortal and the proper name, and this pattern is still attested even for number-incongruent proper names. While (12a) can be taken to result from artificial archaization, (12b), forming a near-minimal pair with (11), is taken from a recent article about paragliding, indicating that the language change is still in progress.

(12) a. V gorode   Fivax   pravili car' Laj i   carica
       in city.M.SG.LOC Thebes.PL.LOC ruled   king Laius and queen
       Iokasta.[6]
       Jocasta
       'King Laius and Queen Jocasta ruled in the city of Thebes.'

   b. V majskie prazdniki on paril   nad   pljažami v kurortnom
      in May   holidays   he soared above beaches in resort.M.SG.LOC
      gorode   Gagrax.[7]
      town.M.SG.LOC Gagra.PL.LOC
      'During May holidays he soared above the beaches in the resort
      town of Gagra.'

Examples (12) were not ungrammatical for some of the native speakers I consulted, including those who, when asked earlier about (11), had rejected the agreeing variant.

There also exists a more restrictive group of speakers, who reject agreeing case on a country or city name that is not gender-congruent with the sortal (see also Rozental et al. 1998: 281). This is in fact the pattern described by Rozental et al. (1998) for toponymic sortals other than the masculine *gorod* 'city, town' and the feminine *strana* 'country'. In most of the current usage, however, as shown by the statistical data in Logvinova (2018: 43), case agreement with the masculine sortal *gorod* 'city, town' is not affected by gender. Interestingly, however, there is one context where the gender factor seems active for this category:

(13) a. A vy   znaete, čto v Rossii est' dva goroda {Pavlovska
       and you.PL know.2PL that in Russia is   two city.GEN Pavlovsk.GEN
       / *Pavlovsk}?
         Pavlovsk.NOM
       'Are you aware that there are two cities named Pavlovsk in Russia?'

---

[6]https://lit.wikireading.ru/hbGcTPBY34

[7]http://www.paraplanerism.ru/kolomenskoe.php

Ora Matushansky

  b. A vy znaete, čto v Štatax est' dva goroda {Moskva
     and you.PL know.2PL that in States is two city.GEN Moscow.NOM
     / *Moskvy}?
      Moscow.GEN
     'Are you aware that there are two cities named Moscow in the
     States?'

The fact that in exactly the same environment case agreement is grammatical for a masculine toponym and ungrammatical, for a feminine one, strongly suggests that case agreement is not correlated with a major difference in meaning.

3.2 Gender congruence as a condition on case agreement

For the sortals *derevnja* 'village', *selo* 'village', *posëlok* 'village', *reka* 'river', *xutor* 'farm' and *ulica* 'street' (the exact list varies from source to source, and Logvinova 2018 claims that in contemporary Russian *reka* 'river' and *gora* 'mountain' reflect this tendency), the toponym in apposition does not agree in case unless congruent with the sortal both in number and in gender (Rozental et al. 1998: 281, Graudina et al. 1976: 140):

(14) a. na ulice   {Jakimanka   / Jakimanke}
    in street.F.SG.LOC Yakimanka.F.SG.NOM Yakimanka.F.SG.LOC
    'on the Yakimanka street'   ✓ phi-congruent
  b. na ulice   {Balčug   / *Balčuge}
    in street.F.SG.LOC Balčug.M.SG.NOM Balčug.M.SG.LOC
    'on the Balčug street'   ✗ phi-congruent
  c. na ulice   {Čistye Prudy   / *Čistyx Prudax}
    in street.F.SG.LOC Čistye Prudy.M.PL.NOM Čistye Prudy.M.PL.LOC
    'on the Čistye Prudy street'   ✗ phi-congruent

The lack of agreement in (14c), containing a masculine plural proper name with a feminine singular sortal, could be due to number incongruence, gender incongruence or both (the lack of familiar plural street names precludes the construction of a gender-congruent example). Moreover, the toponym in (14c) is also internally complex, which, oddly enough, introduces an additional factor to be discussed in §3.5. Since gender is not syntactically active in the plural in Russian, the question arises if number congruence in this category of toponyms should be analyzed as a separate factor, which it is for the toponyms discussed in the previous section, or as merely reflecting the syntactic inactivity of gender in the plural.

3.3 Case agreement with number-congruent pluralia tantum toponyms

While Graudina et al. (1976), Rozental et al. (1998) and other prescriptive sources agree that morphologically plural toponyms disallow agreement, such is not the case when the sortal itself is plural, as with archipelagos (15a) or mountain chains (15b):

(15) a. Kak žit'   na rajskix   ostrovax Mal'divax   za suščie
          how live.INF on Paradise.ADJ islands.LOC Maldives.LOC for real
          groši?[8]
          pennies
          'How to live in the island paradise of the Maldives for peanuts?'
      b. gorami   Al'pami[9]
          mountains.PL.INS Alps.PL.INS
          'with the Alps'

Confirming this observation, Logvinova (2018) also points out that case agreement is possible when a plural sortal is followed by a conjunction of singular toponyms:

(16) v gorodax   Balakove   i Saratove
      in city.M.PL.LOC Balakov.M.LOC and Saratov.M.LOC
      'in the cities of Balakov and Saratov'   (Logvinova 2018)

As the proper name here is a conjunction of two singular toponyms and is therefore plural only by virtue of its semantics, it cannot be argued that number congruence as a precondition for case agreement is ensured by the proper name agreeing with the sortal.

3.4 Case agreement with phi-congruent adjectival proper names only

For the remaining categories of toponyms case agreement in close apposition is possible only with morphologically adjectival toponyms on the condition of both gender and number congruence with their sortals:

(17) a. do stancii   {Bologoe   / *Bologogo}
          until station.F.SG.GEN Bologoe.N.SG.NOM Bologoe.N.SG.GEN
          'until the station Bologoe'   ✗ phi-congruent, ✓ adjective

---

[8]https://arissston.livejournal.com/140512.html
[9]https://limon.kg/ru/news:67260

Ora Matushansky

    b.   do   stancii   {Moskva   / *Moskvy}
        until station.F.SG.GEN Moscow.F.SG.NOM   Moscow.F.SG.GEN
        'until the station Moscow'   ✓ phi-congruent, ✗ adjective

    c.   do   stancii   {Tixoreckaja   / Tixoreckoj}
        until station.F.SG.GEN Tixoreckaja.F.SG.NOM Tixoreckaja.F.SG.GEN
        'until the station Tixoreckaja'   ✓ phi-congruent, ✓ adjective

An incomplete list of such sortals includes ports, lakes, bays, volcanoes, hills (especially the Far Eastern *sopka*), mountains, planets, and railway stations. Prescriptive grammars may insist that case agreement is impossible with such proper names or include in this list islands, republics, etc., but this is because adjectival toponyms are rarely considered. Thus, toponyms preceded by the sortals *aul* 'a village in the Caucasus and Central Asia' and *kišlak* 'a village in Central Asia' are often claimed to never agree for case, but this is because the names of such villages are extremely unlikely to be morphologically adjectival: when an adjectival toponym is used, case agreement becomes possible:

(18) v kišlake   / aule   {Severnom   /
      in kishlak.M.SG.LOC aul.M.SG.LOC Northern.M.SG.LOC
      Severnyj}
      Northern.M.SG.NOM
      'in the kishlak / aul Severnyj'

The observation (Graudina et al. 1976: 143, confirmed by Logvinova) that foreign toponyms do not agree in case when combining with such sortals as *štat* 'state', *respublika* 'republic', etc., is explained by the non-existence of morphologically adjectival foreign proper names.

    As far as I could ascertain, adjectival toponyms can always agree in case with their sortal if they are phi-congruent. In this they differ from proper names of other entities, which do not allow this option:

(19) a.   na minonosce   {"Blestjaščij"   / *"Blestjaščem"}
        on torpedo boat.M.SG.LOC Shining.M.SG.NOM   Shining.M.SG.LOC
        'on the torpedo boat The Shining'

    b.   o   romane   {"Nepobedimyj"   / *"Nepobedimom"}
        about novel.M.SG.LOC Invincible.M.SG.NOM   Invincible.M.SG.LOC
        'about the novel The Invincible'

The contrast between adjectival and nominal toponyms strongly suggests that the latter do not contain an implicit sortal (which would have made them nominal).

246

3.5 Complex toponyms

One more important characterization of close apposition in Russian is that complex toponyms appear to be more restrictive than simplex toponyms. As noted in Graudina et al. (1976: 142), syntactically complex city and country names differ from syntactically simple ones in that the former agree in case only on the condition of gender congruence, just like street names, (20):[10]

(20) a.   v gorode   {Belaja Cerkov'   / *Beloj Cerkvi}
            in city.M.SG.LOC   White Church.F.SG.NOM   White Church.F.SG.LOC
            'in the city of Belaya Cerkov (lit. White Church)'   ✗ phi-congruent

      b.   v gorode   {Petropavlovsk-Kamčatskij   /
            in city.M.SG.LOC   Petropavlovsk-Kamčatka.ADJ.M.SG.NOM
            Petropavlovske-Kamčatskom}
            Petropavlovsk-Kamčatka.ADJ.M.SG.LOC
            'in the city of Petropavlovsk-Kamchatskij (lit. Petropavlovsk of
            Kamchatka)'

As before, Internet searches locate some instances of case agreement for (20a) that probably reflects an earlier stage of the linguistic change in progress, whereas the native speakers that I consulted conform to the generalization in Graudina et al. (1976): only phi-congruent complex city names can agree in case, exhibiting the more restricted pattern associated with street names (§3.3). Logvinova (2018) supports this generalization showing that complex masculine city names (the word *gorod* 'city, town' is masculine) are less likely to agree than simplex masculine city names of comparable frequency.[11]

A similar effect is reported for internally complex street names, such as *Novaja Zarja* 'the New Dawn'. While street names are generally asserted to require gender congruence (as in (14) in §3.2), some prescriptivists claim that complex feminine street names behave like masculine street names and disallow case-agreement (recall that the sortal *ulica* 'street' is feminine), resulting in the pattern in (21a).[12] Others only draw a distinction between feminine street names (which

---

[10]Gorpinič (1987) asserts that in Ukrainian complex toponyms in close apposition do not agree in case, but a quick informal check has shown that such is not the case for at least some native speakers.

[11]Graudina et al. (1976: 149) also claims that while agreeing adjectival modifiers have this effect, PP modifiers do not. Logvinova (2018) does not examine such cases and I have not been able to verify this claim or disprove it.

[12]E.g., https://newslab.ru/article/465957

Ora Matushansky

agree in case) and masculine ones (which do not).[13] Importantly, complex adjectival street names allow case agreement (21b).

(21) a. na ulicu   {Novaja Zarja   / *Novuju Zarju}
        on street.ACC New   Dawn.NOM New   Dawn.ACC
        'on(to) the street New Dawn'

    b. na ulicu   {Malaja Bronnaja   / Maluju Bronnuju}
        on street.ACC Small   Hauberk.ADJ.NOM Small   Hauberk.ADJ.ACC
        'on(to) the Lesser Hauberk street'

    c. na Maluju Bronnuju   ulicu
        on Small   Hauberk.ADJ.ACC street.ACC
        'on(to) the Lesser Hauberk street'

Even though adjectives do not modify adjectives and *malaja* 'small' in (21b) is originally a restrictive modifier (the Small Hauberk street, as opposed to the bigger one), it seems unlikely that (21b) contains a null head noun, or it would behave the same as (21a). One more possibility is that (21b) is derived by inversion from (21c), where the sortal forms part of the toponym, yet inversion is generally impossible with toponyms (22)–(23), except in poetry:

(22) a. na Sennoj   ploščadi
        on hay.ADJ.F.SG.LOC Square.F.LOC
        'on Hay Square'

    b. * na ploščadi   Sennoj
        on Square.F.LOC hay.ADJ.F.SG.LOC

(23) a. na Nevskom   (prospekte)
        on Nevsky.M.SG.LOC avenue.M
        'on the Nevsky (Prospekt)'

    b. * na prospekte   (Nevskij   / Nevskom)
        on avenue.M.LOC Nevsky.M.SG.NOM   Nevsky.M.SG.LOC

The fact that complex adjectival toponyms do not behave as nominal ones provides additional support for the lack of an implicit sortal in adjectival toponyms, which the contrast between adjectival and nominal toponyms has already suggested.

---

[13]E.g., http://new.gramota.ru/spravka/buro/search-answer?s=295848

3.6 Intermediate summary

The behavior of toponyms clearly shows that case agreement depends on phi-congruence and that the strictness of this condition is determined by the lexical-semantic class of the proper name: while animate proper names require case agreement and non-toponymic inanimate ones disallow it, toponyms permit case agreement on variable conditions of phi-congruence: while for cities and countries number congruence is a sufficient condition for case agreement, street names require gender congruence in addition, and other toponyms can agree in case only if they are adjectival, as shown below.

Table 1: Case agreement with proper names

	no case	+adjecti- val	gender	number	no con- gruence
animates	✗	✓	✓	✓	✓
cities, countries, rivers...	✓	✓	✓	✓	✗
streets, villages...	✓	✓	✓	✗	✗
other toponyms	✓	✓	✗	✗	✗
non-toponymic inanimates	✓	✗	✗	✗	✗

For some speakers certain lexical-semantic classes seem to be more restrictive than described by the existing sources and "shifted downwards" in the table, and the same appears to be the case for internally complex toponyms, though the facts are yet far from clear.

Several facts should be accounted for, which excludes some analyses that appear plausible at a first glance:

(24)   • animate sortals require a case-agreeing proper name
   • case agreement is impossible with inanimate non-toponymic proper names
   • without an overt sortal all proper names are appropriately case-marked
   • it is the sortal that determines how the entire NP agrees
   • the same proper noun (e.g., *Moskva* 'Moscow' in (7b) and in (10a)) may behave differently with different sortals
   • it does not seem that agreeing toponyms permit some interpretation or usage that non-agreeing ones do not

- internally complex toponyms may yield different congruence restrictions, though the entire empirical picture is yet unclear
- at a prior stage of the language toponyms did not require phi-congruence for case agreement
- with cardinals, sorted city names require gender congruence (13)

The distinction between toponyms and other proper names suggests that the lexical-semantic class of a proper name is reflected in its syntax in a principled way.

4 Excluded hypotheses

The empirical generalizations established above provide the desiderata for an explanation that exclude several immediately obvious and not-so-obvious hypotheses.

4.1 Semantic type distinction

A question that needs to be addressed by any theory of close apposition is the semantic type of the proper name (or kind name, for that matter). Two options are available: a predicate and an individual.

The standard approach to proper names is to regard them as individual constants: in argument positions the name *Alice* denotes the individual *a*. However, since, as first pointed out in this context by Sloat (1969), proper nouns can also appear in positions where such a denotation is impossible (25), an additional denotation for them is needed, where they denote predicates.

(25) a. *Some/✓ sóme Smith/man stopped by.
      b. Some/sóme Smiths/men stopped by.
      c. Smiths/men must breathe.
      d. The clever Smith/man stopped by.
      e. The Smith/man who is clever stopped by.
      f. A clever Smith/man stopped by.
      g. The Smiths/men stopped by.
      h. The *Smith/✓ man stopped by.
      i. Smith/*man stopped by.   (Sloat 1969)

The predicative approach to proper names (see Matushansky 2008, Gray 2015, and Fara 2015 for recent takes and references) argues that the denotation in (26a)

can and must be derived as a referential definite description built on the basis of the predicative denotation presented in a simplified form in (26b).

(26)   a. $[\![\text{Alice}]\!] = \textsc{a}$
      b. $[\![\text{Alice}]\!] = \lambda x \in D_e \,.\, x$ is called /ælɪs/

Yet for our purposes it is sufficient that the toponym in close apposition can in principle be referential or predicative.[14] Can case agreement be taken as an argument for the simultaneous availability of both options and used to differentiate between the two?

Several reasons can be provided why this approach should not be taken. Firstly, the fact that animate proper names require case agreement, while inanimate non-toponymic proper names disallow it is hard to square with different denotations: we do not expect animacy to interact in this way with the semantic type. Secondly, if case-agreeing toponyms are referential and non-agreeing ones are predicative (or *vice versa*), we expect that there is some context of use that the non-agreeing close apposition in (17a)–(17b) lacks and the agreeing close apposition in (17c) has, which does not seem to be the case. While a more detailed survey might reveal such a difference, no research so far has indicated that there is some meaning or use that (27a) might have while (27b) would lack it, nor is there any obvious interpretational distinction for the agreeing vs. non-agreeing options for one and the same toponym in (17c) or for the gender-distinct toponyms in exactly the same environment in (13).

(27)   a.   v gorode   Moskve
         in city.M.SG.LOC Moscow.F.SG.LOC
         'in the city of Moscow'   singular sortal, feminine PN

      b.   v gorode   Gagry
         in city.M.SG.LOC Gagra.PL.NOM
         'in the city of Gagra'   singular sortal, plural PN

It can be argued that a predicative proper noun, as in (26b), can be combined with the definite article (or the corresponding type-shift, the iota-operator) to give rise

---

[14] It is tempting to appeal to the lack of the article in *the river Rhine* as an argument for treating the toponym as non-referential. However, in the next language over, Dutch, the article is present:

   (i)   de rivier de Rijn
       the river the Rhine
       'the river Rhine'   (Dutch, van Riemsdijk 1998)

to a definite NP with an interpretation that is virtually indistinguishable from (26a), as in (28). While the predicative approach to proper names argues that this is in fact how their referential use is derived, the referential approach may rely on the ambiguity in (26a)–(26b) to derive the two syntactic options: the proper name (26a) and the definite DP (28).

(28) $\iota x \in D_e$. x is called /ælıs/

While at first blush such an analysis could be taken as an argument in favor of the referential approach to proper names, two problems arise as a result. Firstly, in general, if both options are available in principle, how do we know which one we are dealing with in *Alice is here?* Secondly, specifically to the empirical issue at hand, why should one of the two options be unavailable for animate proper names (which require case agreement in close apposition) and the other, for inanimate non-toponymic proper names (which require nominative) and why should gender features, as in (13), be relevant? The same two issues arise for any view that derives the variation in case agreement from a difference in the interpretation, and the theory to be discussed now is no exception.

4.2 Quotation

The semantic distinction between mention and use looks like a plausible explanation for the two different syntactic options. It is an immediately obvious hypothesis that case invariability involves quotation, and even the objection raised at the end of §4.1 might be overcome: maybe quotations are obligatorily inanimate and cannot function as anthroponyms or zoonyms, thus explaining why animate proper names require case agreement.

Two issues remain, however. Firstly, it is still an open question why inanimate non-toponyms disallow case agreement. Secondly, if the interpretation of the proper name is not the same in agreeing vs. non-agreeing cases, some difference in use is expected. There are, however, no cases where a phi-congruent and hence agreeing toponym is possible and another toponym, which does not permit agreement due to phi-incongruence, is excluded. In other words, the fact that a certain toponym cannot agree with a given sortal does not preclude its appearance in any context where an agreeing toponym with the same sortal can appear, which strongly suggests no difference in semantics for case-agreeing and invariant toponyms.

4.3 The sortal as the locus of variation

Although case agreement variation for toponyms is usually described in terms of lexical-semantic classes, it is tempting to hypothesize that it is not the toponyms that are responsible for it, but their sortals, e.g., that some nouns can enter the derivation underspecified for some phi-features. The advantage of this approach is that it can explain why the same proper nouns (e.g., *Moskva* 'Moscow' in (7b) and in (10a)) behave differently by suggesting that it is not in the proper noun but in the sortal where the difference lies.[15] The flip side is the prediction that different sortals applying to the same set of proper names are not expected to behave the same. Testing this prediction is difficult: the same behavior for different sortals can easily be attributed to coincidence. In fact, the feminine *stolica* 'capital', which combines with a subset of the toponyms that the masculine *gorod* 'city, town' can combine with, also requires only number congruence, whereas the difference between the coextensional *strana* 'country' and *gosudarstvo* 'state' (see Footnote 15) can be due to the difference in their gender, so confirming or disproving this prediction is impossible.

Another problem with this hypothesis is that it cannot explain why phi-congruent adjectival toponyms can always agree in case with their sortal, and why non-toponymic proper names never do: if the source of the relevant phi-features is the proper name, adjectival and nominal proper names should not differ, and the same is true for toponyms vs. non-toponyms. One more problem is motivation: these sortals do not exhibit any obvious semantic or syntactic peculiarities in any other contexts (which, however, is also true for the toponyms themselves). Finally, the very mechanism of "agreement as valuation" is ill-suited for dealing with phi-congruence, as we will now see.

4.4 Phi-congruence as valuation

Two mechanisms are provided by the current syntactic theory for comparing the phi-features of two constituents: agreement and semantic matching. As it is generally assumed that gender features of inanimate nouns are not interpretable, the feminine of *Jakimanka* and that of *ulica* 'street' in (14a) cannot be matched by

---

[15]Logvinova (2018) documents a difference in the behavior of the same toponyms with the feminine *strana* 'country' (case agreement conditioned by number congruence) as opposed to the neuter *gosudarstvo* 'state' (no case agreement). While the question is open whether the (sorted) toponyms denote the same entity, the syntax could still be the same, as the observed difference would also follow from the gender of the sortal: there were no neuter country names in the data set. Furthermore, as discussed in Footnote 5, neuter toponyms resist case-marking even without a sortal.

Ora Matushansky

ensuring that their presuppositions match: they do not introduce any.[16] Syntactic agreement remains then the only option.

While number can reasonably be argued to not be inherent to a noun, gender arguably is. It is possible, however, that the gender feature is introduced on a special functional head (e.g., *n*, see Kihm 2005, Lowenstamm 2007, Acquaviva 2009, Percus 2011, and Kramer 2015, among others) and some additional (and independently needed) mechanism ensures that it correlates properly with the semantics of the noun (for animates) and its declension class. How can we then implement the fact that some sortals, e.g., *ulica* 'street', can agree with the toponym?

Suppose that *ulica* 'street' can combine directly with the toponym and the gender-introducing functional head *n* (be it categorizing or not) enters the derivation afterwards:
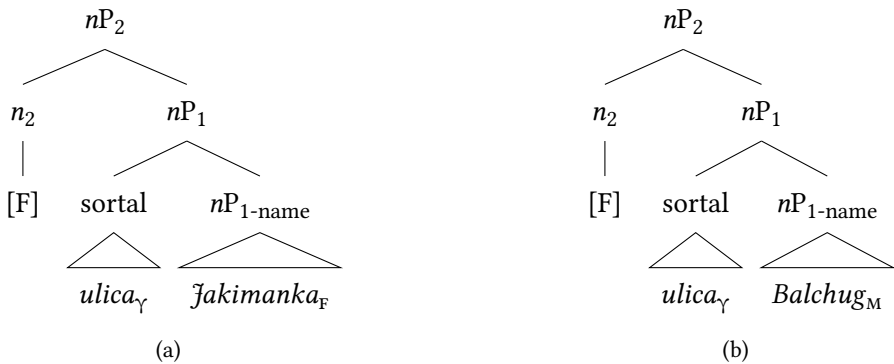
Figure 2: Gender congruence: sortal *n* as a head external to close apposition

Setting aside many technical details, consider Figure 2b, where the gender values of the sortal and of the proper name do not match. The proper name is masculine (a valued feature), so *ulica* 'street' should also be assigned masculine, contrary to its declension class, which assigns it to feminine (and the gender feature of the resulting complex NP (nP_2) should also be feminine). Nouns whose gender does not match its declension class, such as semantically feminine nouns ending in a consonant (29), do not decline in Russian.

(29) a. k   ėtoj   {madam   / *madame   /
          towards this.DAT   madam.DAT$_{INDECL}$ madam.DAT$_{a-DECL}$

---

[16]This assessment will be reexamined in §5.

       *madamu}

       madam.$DAT_{c\text{-}DECL}$

       'towards this madam'

  b.   s   {Karmen   / *Karmenoj   / *Karmenom}

       with Carmen.DAT_{INDECL} Carmen.$DAT_{a\text{-}DECL}$ Carmen.$DAT_{c\text{-}DECL}$

       Ivanovnoj

       Ivanovna.$DAT_{a\text{-}DECL}$

       'with Carmen Ivanovna'

At the nP_1 level the prediction is that *ulica* 'street' would not agree. This is a wrong result, so let us suppose that the feminine feature of n_2 somehow overrides the masculine obtained from $nP_{1\text{-name}}$, both on the sortal and on the proper name. Feminine gender specification contradicts the morphological properties of the toponym, so the structure in Figure 2b would result in a non-agreeing form, as desired.[17]

This approach, however, cannot be extended to toponyms agreeing in case on the condition of number congruence. Firstly, number is generally associated with the presence of plural semantics, i.e., a *-operator or a cardinal (or both, this depends on the adopted approach to cardinals). In the case of number-congruent pluralia tantum toponyms, like in (15), where both the sortal and the toponym bear plural morphology, there seems to be no reasonable way in which one of them could be unvalued.[18] To see this, consider the following structures:

The toponym *Al'py* 'the Alps' in Figure 3 corresponding to (15b) is plural, on both morphological and semantic grounds, so its number feature is valued. Consider first Figure 3a, where the number feature of the sortal is unvalued and so can in principle agree with the valued number feature of the toponym. However, the semantics of Figure 3a is incorrect: if *Al'py* 'the Alps' is referential here, then the higher nP_1 node denotes the set of singular mountains that is the Alps, i.e., the empty set. If *Al'py* is predicative, then the higher nP_1 node denotes a set of mountains each of which either is called (the) Alps or is a plurality called (the) Alps, which is equally incorrect.

Consider now Figure 3b as the structure for (15b), assuming that Num^0 of the sortal is the source of the plural semantics (if it isn't, the same problem arises as

---

[17]The fact that phi-congruent toponyms may still not agree in case requires an additional richer structure, where the sortal is specified for gender and the toponym, not having agreed with it, does not count as part of the same NP for the purposes of case-assignment (or more likely, concord).

[18]I note here that in the singular the feminine noun *gora* 'mountain' allows case agreement on the condition of gender congruence, though to a lesser degree than *strana* 'country' or *reka* 'river' (Logvinova 2018: 22).

Ora Matushansky
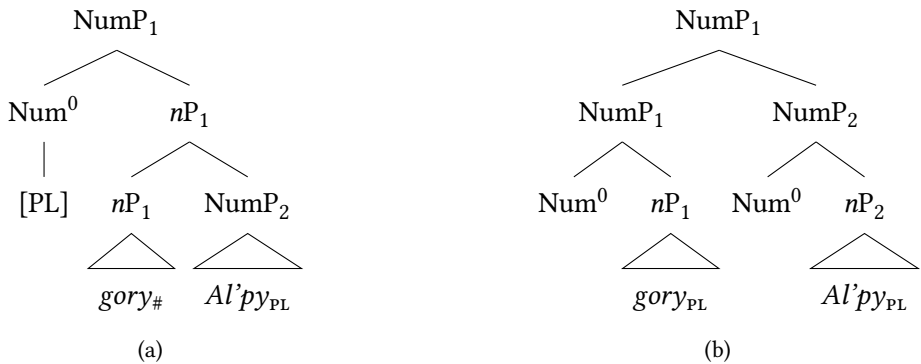
Figure 3: Number congruence: the position of Num in close apposition

in Figure 3a). The semantics is now correct, but the number feature of the sortal cannot be unvalued.

Two more options are available in principle. One (Figure 4a) is to assume that the unvalued number feature is on the toponym, contrary to what has been assumed before (and despite the fact that it is a *plurale tantum*). The second (Figure 4b) is to treat number features as unvalued on both the sortal and the toponym.
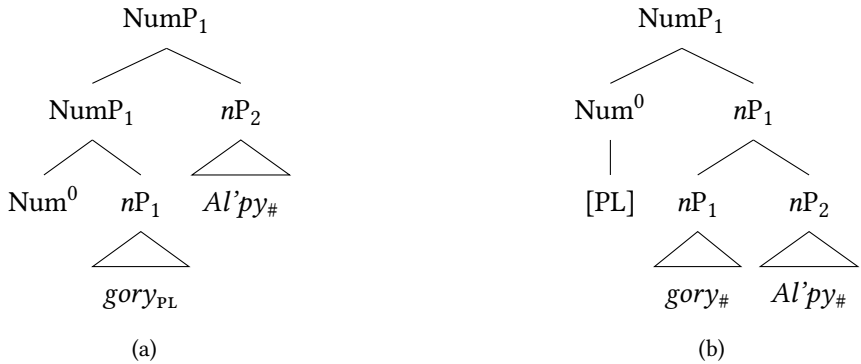
Figure 4: Number congruence: uninterpretable number is on the name

Even setting aside their syntactic plausibility, both options fail with the conjoined singulars in (16) where the toponym cannot be reasonably regarded as having unvalued number: a non-intersective conjunction of two singulars (be it a sum of two individuals or a set-product of two predicates) can under no assumptions be non-plural semantically.

We conclude that case agreement with a phi-congruent plural sortal poses an unsurmountable obstacle to treating the phi-congruence condition in toponymic close apposition as valuation.

4.5 Intermediate summary

We have examined four theories that can be advanced to explain the phenomenon of varying case agreement in close apposition in Russian. Two of them suggest a semantic difference between agreeing proper names (assumed to be referential) and non-agreeing proper names (which are attributed predicative semantics (or maybe indirectly referential semantics) or the semantics of quotation). The other two address the syntactic side of the problem: the locus of the unvalued features that should drive case agreement and the applicability of the theory of agreement as feature valuation to close apposition.

The failure of syntactic theories is due to the fact that phi-congruence is established between interpretable features that can be simultaneously valued on the sortal and on the proper name. On the semantic side one problem is that the immediately obvious potential solutions do not take into consideration the difference between lexical-semantic classes of proper names, and another, that there is no independent evidence for a semantic distinction.

What follows is a sketch of a solution based on two assumptions: (a) that agreement in close-apposition is semantic and as such, based on feature-value matching rather than valuation and (b) that the semantic sort of toponyms is different from that of other proper names, so they can be singled out on semantic grounds.

5 Toponyms as a semantic sort

One of the main facts to be accounted for is the distinction between animate proper names (which obligatorily agree in case), toponyms (which may do so) and inanimate non-toponyms (which cannot do so).

As case agreement is clearly dependent on phi-congruence, it is natural to hypothesize that a proper name counts as part of the same NP as the sortal if it agrees with it in some feature. Case agreement then becomes something of a free-rider in the sense that case-assignment to the proper name forming part of the same NP as the sortal (which is what agreement enables) can be viewed as concord: multiple realizations of the case assigned to the entire NP. Without further elaboration of this hypothesis, I further suggest that different lexical-semantic classes of proper names underlyingly have different semantic phi-feature specifications and attempt to motivate these distinctions by independent factors.

Ora Matushansky

5.1 The role of animacy

Being a subtype of nouns, proper names have valued formal phi-features determined by their semantics and their declension class. Since formal gender (for inanimate nouns) and formal number (for pluralia tantum) can be inherently valued and fail to agree, the only remaining option for agreement in close apposition are semantic phi-features. The first such feature is obviously animacy.

I will not decide here how this feature value is set. Three possibilities can be envisaged: from the sortal, from the denotation of the proper name itself (if it is referential) or from the denotation of the entire appositive noun phrase. What is crucial is that semantically, animacy is a privative feature, so inanimate nouns lack it. This means that a proper name can semantically agree for animacy only with animate sortals, which would explain why only animate proper names agree in close apposition.[19]

The question is now why toponyms do not behave as other inanimate proper names.

5.2 Locative nominals as a lexical-semantic class

There is mounting evidence that the syntax of nouns denoting places is different from that of nouns denoting other entities. Thus Haspelmath (2019) shows that cross-linguistically nouns denoting places are less marked in locative environments than regular object-denoting nouns and Matushansky (2019) argues that crosslinguistic use of toponyms and a few common nouns as locative adverbials with zero or special marking indicates denotation in the special locative domain (variants of which have been independently postulated to account for the semantics of spatial prepositions, see Bierwisch 1988, Wunderlich 1991, Zwarts & Winter 2000, Kracht 2002, Bateman et al. 2010, etc.). Evidence for a special status

---

[19]One might object that animate proper names also have semantic gender, which they need not share with the sortal (Footnote 2). A counterargument to this objection is that a human-denoting NP in Russian may acquire semantic gender that overrides its formal gender (Crockett 1976, Corbett 1979, Rothstein 1980, Nikunlassi 2000, Asarina 2008, Pesetsky 2013, etc.):

(i) U nas byla ocen' xorošaja zubnoj vrač.
with us was.F.SG. very good.F.SG. dental.M.SG doctor.M
'We had a very good dentist.' (Crockett 1976)

In other words, sortals whose gender is different from that of the anthroponym can also agree on the basis of the gender of the referent outside close apposition, so arguably either do not possess underlying semantic gender or can acquire the gender of their referent by an independently motivated mechanism and then presumably agree with the proper name.

of locative place names in Martinican Creole can also be found in Zribi-Hertz & Jean-Louis (2014), 2017, 2018. In Russian itself, support for this view comes from the so-called LOCATIVE-II: the special form of the Russian locative case that certain nouns take when appearing with the prepositions *v* 'in' or *na* 'on' denoting the default locative relations with these nouns (30).[20] Other nouns (including other location nouns) do not show this distinction:

(30)   a.   voda v taz-u
            water in hand-basin-LOC$_{II}$
            'water in the hand-basin'   default locative meaning

       b.   nadpis' na taz-e
            writing on hand-basin-LOC
            'writing on the hand-basin'   non-default locative meaning
                                                        (Plungjan 2002)

The fact that adjectival modification of nouns in locative II is allowed shows that they cannot denote in the loci domain (since loci, be they regions, sets of points, or sets of vectors, do not have the same domain structure as objects and cannot be modified by the same modifiers). Yet locative II provides evidence for a crucial underlying distinction between object nouns and place nouns, and I propose that toponyms can be distinguished from other proper names on precisely these grounds (even though toponyms are never marked with locative II in Russian). Moreover, since locative-II nouns denote not only places, but also objects (i.e., any such noun can enter the derivation with either sort), we expect that the non-agreeing option will be possible in the latter denotation.

The question is how this distinction translates into optional case agreement on the condition of phi-congruence.

5.3 Number features of toponyms

Importantly, Russian toponyms are not syntactically uniform. Their behavior with respect to case agreement separates them into three classes (cf. Table 1):

(31)   •   countries and cities: number congruence is required for case
           agreement
       •   rivers, villages, etc.: number and gender congruence is required

---

[20]The distribution of the "second prepositional case" (locative II) is very complicated, as discussed in Plungjan (2002), Brown (2007) and Itkin (2016) (see Nesset 2004 for its use in temporal expressions).

- others: only agreeing adjectival toponyms agree

I stipulate that, unlike other proper names, toponyms by virtue of their semantic sort cannot be mass. This generates the semantic feature of number, which is denotation-based. For most toponyms this would mean singular, but it is over-ridden by the formal plural with a plurale tantum toponym. It is only when the sortal is plural as well that no conflict arises.

The question is then what to do with gender.

5.4 Semantic agreement and referentiality

The appeal to semantic agreement raises the question of whether case-agreeing proper names are referential since semantic agreement is known to rely on the properties of the denotatum. Importantly, case agreement is known to be facili-tated if the toponym is familiar (Graudina et al. 1976, Rozental et al. 1998, Logvi-nova 2018, in press).[21] While it seems plausible therefore that case agreement in close apposition correlates with the referentiality of the toponym, testing this hypothesis with native speakers does not support this conclusion:

(32) a. obsledovanie domašnix xozjajstv žitelej goroda
examination home economy.GEN residents.GEN city.GEN
Ekaterinburga a takže naxodjaščixsja na territorii
Ekaterinburg.GEN and also located.PL.GEN on territory
Sverdlovskoj oblasti gorodov Pervoural'ska i
Sverdlovsk.ADJ region cities.GEN Pervouralsk.GEN and
Kamensk-Ural'skogo
Kamensk-Ural'sky.GEN
'an examination of the housekeeping of the residents of the city of
Ekaterinburg as well as of the towns of Pervouralsk and
Kamensk-Uralsky, located in the Sverdlovsk region'   (RNC)
b. Krome goroda Pavlovska pod Piterom, est' ešče odin –
besides city.GEN Pavlovsk.GEN under Piter is also one
pod Voronežem.
under Voronezh
'Besides the town of Pavlovsk near St. Petersburg, there is one more
near Voronezh.'

---

[21]Logvinova (in press: p. 56) provides evidence from city names that higher frequency of a to-ponym increases the frequency of case agreement. As previously described (Graudina et al. 1976, Rozental et al. 1998), plural and two-word toponyms are less likely to agree in case. She also observes that unexpectedly, adjectival city names are less likely to agree in case.

The RNC example (32a) strongly implies that the hearer is not familiar with the two towns in question, yet case agreement is grammatical there. More convincingly, perhaps, the toponym *Pavlovsk* cannot be referential in example (32b) because two places with such a name exist in the context, and the same is true in (13).

Nonetheless as the presupposition of countability applies to all toponyms it seems reasonable to view semantic features here as derived from the denotation. The situation is more complex where it comes to gender.

5.5 Inanimate gender as a formal feature

To account for case agreement on the condition of gender congruence (§3.2) I propose that, contrary to what happens to inanimates in general, gender features of toponyms may be interpretable. Independent evidence for this comes from indeclinable toponyms and common nouns. While inanimate nouns in Russian are generally assigned gender on the basis of their declension class, the gender of indeclinable toponyms is often the same as the gender of their hypernym (Rozental et al. 1998, Doleschal 1996, Murphy 2000, Matushansky 2022, a.o.), which strongly suggests that inanimate gender can also be interpretable at LF.[22] If, as corpus searches reveal, along with the neuter expected for inanimates the indeclinable *Zimbabve* 'Zimbabwe' can be feminine (because *strana* 'country' is feminine) and *Bol'šoj Zimbabve* 'Great Zimbabwe' can be masculine (because *gorod* 'city' is), nothing excludes that morphologically declinable toponyms can also have semantic gender. If their gender is systematically determined by their declension class (as can be seen from their agreement outside close apposition), then for case agreement this semantic/formal gender of a toponym would have to match the gender of the sortal along the same lines as discussed for animacy and number.

The hypothesis that formal gender features can be semantically interpretable (as is needed to explain toponyms requiring gender congruence for case agreement (§3.2)) entails that gender features of toponyms requiring number congruence only (§3.1) should also be interpretable. Where does the difference come from?

By our prior reasoning toponyms are non-mass, so semantic agreement in number is possible for all toponyms and seems to be required for case agreement. To explain the role of gender it is necessary to assume that when gender

---

[22]Indeclinable common nouns can also be assigned semantic gender on the basis of their hypernym (see Wang 2014, Baranova 2016, Chuprinko et al. submitted, a.o.), both in online computation and prescriptively.

is semantically interpretable, semantic agreement just for number is insufficient. The question then arises why gender is interpretable for some toponyms (§3.2) but not for others (§3.1) and how come it suddenly becomes so for the latter in cases like (13).

The crucial property of (13) is obviously the paucal cardinal. The cardinal assigns a formal plural (or paucal) value to the number features of the sortal and the toponym, which both are morphologically singular and, following Ionin & Matushansky (2006, 2018), semantically atomic, even though the denotatum is semantically plural. Furthermore, the toponym, being in the scope of the cardinal, is not referential. Which of these factors (number mismatches or non-referentiality) can explain the more restricted character of toponyms discussed in §3.2 remains an open question.

5.6 Adjectival toponyms

To conclude the proposed sketch of a solution, it is necessary to explain why case agreement with a phi-congruent adjectival toponym is possible for any sortal. The core intuition should rely on the fact that adjectives normally do not have any underlying phi-features as all. As metalinguistic as it sounds, it seems reasonable that adjectival toponyms come with a strong intuition of what the sources of their valued phi-features are, i.e., with some presupposition about their sortals. While it is unlikely that the hypernym is syntactically represented, it can function as the source of semantic phi-features, enabling the toponym to establish semantic agreement with its sortal.

6 Conclusion

We have seen that Russian proper names fall into three categories in function of how they behave with respect to case agreement in close apposition. Proper names of human and other animate entities necessarily agree in case with the sortal. Names of inanimate entities that are not locations conversely never agree in case with the sortal. Finally, toponyms fall into the intermediate category: they may fail to agree in case with the sortal or allow case agreement on the condition of congruence in number (§3.1) or in number and gender (§3.2). While we have not looked at kind names in detail, they seem to pattern either as city/country names or as non-toponymic names (Footnote 3).

I propose that the crucial distinction between toponyms and other inanimate proper names is that toponyms may introduce interpretable phi-features in close

apposition. The advantages of this hypothesis are that, on the one hand, it does not need to assume that any semantic factors distinguish between toponyms agreeing and not agreeing in case, and on the other, that the introduction of interpretable phi-features can be naturally linked to frequency: more frequent toponyms would be more clearly identified with some presuppositions.

Many questions remain. For the time being we have no principled explanation for why there are these three classes of toponyms, or why internal syntactic complexity of proper names influences case agreement.[23] We have not explored adjectival proper names in sufficient detail and only sketched a possible solution for the apparently obligatory gender congruence with cardinals. Likewise, we have not addressed the fact that close apposition may involve restrictive or non-restrictive interpretation of the sortal and did not make clear how agreement (or congruence) in phi-features can enable agreement in case (which is, after all, a purely syntactic operation).

The entire phenomenon of phi-congruence in case agreement in toponymic close apposition, which we have encoded by hypothesizing that inanimate proper names may acquire semantic gender features, might instead be regarded as an argument in favor of treating agreement as matching rather than valuation. Irrespective of the eventual implementation, the issue of phi-congruence in case-agreement raises a number of problems for standard approaches to both proper names and agreement.

Abbreviations

2	second person	GEN	genitive
a-DECL	the declension class of nouns ending in a in the nominative	INF	infinitive
		INS	instrumental
ACC	accusative	LOC	locative
ADJ	adjective	LOC-II	locative-II
C-DECL	the declension class of masculine nouns ending in a consonant in the nominative	M	masculine
		N	neuter
		NEG	negation
		NOM	nominative
DAT	dative	PL	plural
F	feminine	SL	singular

---

[23]One possible answer might be that internally complex toponyms are simply less frequent, but this hypothesis requires independent confirmation.

Ora Matushansky

Acknowledgments

References

Acquaviva, Paolo. 2009. Roots and lexicality in Distributed Morphology. In Alexandra Galani, Daniel Redinger & Norman Yeo (eds.), *Fifth York-Essex Morphology Meeting (YEMM), 9th February and 10th February 2008*, 1–21. York: University of York. http://hdl.handle.net/10197/4148.

Asarina, Alya. 2008. Gender and adjectives in numeral constructions in Russian. MIT. http://web.mit.edu/alya/www/russian-gender.pdf.

Baranova, Anna S. 2016. Otraženie variativnosti rodovoj harakteristiki nesklonjaemyh suŝestvitel'nyh v sovremennyh slovarjah. *Vestnik Sankt-Peterburgskogo Universiteta. Jazyk i literatura* 9(1). 45–57. http://hdl.handle.net/11701/2745.

Bateman, John A., Joana Hois, Robert Ross & Thora Tenbrink. 2010. A linguistic ontology of space for natural language processing. *Artificial Intelligence* 174(14). 1027–1071. DOI: 10.1016/j.artint.2010.05.008.

Bierwisch, Manfred. 1988. On the grammar of local prepositions. In Manfred Bierwisch, Wolfgang Motsch & Ilse Zimmermann (eds.), *Syntax, Semantik und Lexikon* (Studia Grammatica XXIX), 1–65. Berlin: Akademie.

Brown, Dunstan. 2007. Peripheral functions and overdifferentiation: The Russian second locative. *Russian Linguistics* 31(1). 61–76. http://www.jstor.org/stable/40160837.

Burton-Roberts, Noel. 1975. Nominal apposition. *Foundations of Language* 13(3). 391–419. http://www.jstor.org/stable/25000923.

Chuprinko, Kirill, Magomedova Varvara & Slioussar Natalia. Submitted. *Gender variation in Russian indeclinable nouns: Optimality over structuralism.* Submitted to Acta Linguistica Academica.

Corbett, Greville G. 1979. The agreement hierarchy. *Journal of Linguistics* 15. 203–224. DOI: 10.1017/S0022226700016352.

Crockett, Dina B. 1976. *Agreement in contemporary standard Russian.* Bloomington, IN: Slavica Publishers, Inc.

Doleschal, Ursula. 1996. Gender assignment revisited. In Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen & Mirja Saari (eds.), *Gender in grammar and cognition: I: Approaches to gender*, 117–166. Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110802603.117.

Fara, Delia Graff. 2015. Names are predicates. *The Philosophical Review* 124(1). 59–117. DOI: 10.1215/00318108-2812660.

Gorpinič, Volodimir O. 1987. Skladni vipadki vidminjuvannja toponimiv. *Kul'tura slova* 32. 69–73.

Graudina, Ljudmila Karlovna, Viktor Aleksandrovič Ickovič & Lia Pavlovna Katlinskaja. 1976. *Grammatičeskaja pravil'nost' russkoj reči. Stilističeskij slovar' variantov.* Moscow: Nauka.

Gray, Aidan. 2015. Lexical individuation and predicativism about names. *Thought: A Journal of Philosophy* 4(2). 113–123. DOI: 10.1002/tht3.164.

Haspelmath, Martin. 2019. Differential place marking and differential object marking. *STUF - Language Typology and Universals* 72(3). 313–334. DOI: 10.1515/stuf-2019-0013.

Haugen, Einar. 1953. On resolving the close apposition. *American Speech* 28(3). 165–170. DOI: 10.2307/454129.

Ionin, Tania & Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23(4). 315–360. DOI: 10.1093/jos/ffl006.

Ionin, Tania & Ora Matushansky. 2018. *Cardinals: The syntax and semantics of cardinal-containing expressions.* Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/8703.001.0001.

Itkin, Il'ja B. 2016. Udivitel'naja istorija vtorogo predložnogo padeža. Paper presented at «Publičnye lekcii "Polit.ru"», May 19, 2016.

Jackendoff, Ray. 1984. On the phrase *the phrase* 'the phrase'. *Natural Language & Linguistic Theory* 2(1). 25–37. DOI: 10.1007/BF00233712.

Keizer, Evelien. 2005. The discourse function of close appositions. *Neophilologus* 89(3). 447–467. DOI: 10.1007/s11061-004-0963-9.

Kihm, Alain. 2005. Noun class, gender, and the lexicon-syntax-morphology interfaces: A comparative study of Niger-Congo and Romance languages. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, 459–512. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780195136517.013.0011.

Kracht, Marcus. 2002. On the semantics of locatives. *Linguistics and Philosophy* 25(2). 157–232. DOI: 10.1023/A:1014646826099.

Ora Matushansky

Kramer, Ruth. 2015. *The morphosyntax of gender*. Oxford: Oxford University Press.

Lasersohn, Peter. 1986. The semantics of appositive and pseudo-appositive NPs. In Fred Marshall, Ann Miller & Zheng-sheng Zhang (eds.), *Proceedings of the Eastern States Conference on Linguistics 3*, 311–322. Columbus, OH: Department of Linguistics, the Ohio State University.

Logvinova, Natal'ja N. 2018. Appozitivnye konstrukcii s imenami sobstvennymi v russkom jazyke. Ms., High School of Economics (St. Petersburg). https://www.academia.edu/42922134/.

Logvinova, Natal'ja N. in press. Imennye appozitivnye konstrukcii v russkom jazyke. Materialy k korpusnoj grammatike russkogo jazyka. http://rusgram.ru/pdf/appositive_logvinova_181020_final.pdf.

Lowenstamm, Jean. 2007. On little n, $\sqrt{}$, and types of nouns. In Jutta Hartmann, Veronika Hegedüs & Henk Van Riemsdjik (eds.), *The sounds of silence: Empty elements in syntax and phonology*, 105–144. Amsterdam: Elsevier.

Matushansky, Ora. 2008. On the linguistic complexity of proper names. *Linguistics and Philosophy* 31(5). 573–627. DOI: 10.1007/s10988-008-9050-1.

Matushansky, Ora. 2019. The case of restricted locatives. In M.Teresa Espinal, Elena Castroviejo, Manuel Leonetti, Louise McNally & Cristina Real-Puigdollers (eds.), *Proceedings of Sinn und Bedeutung 23*, vol. 2, 161–178. Barcelona: Universitat Autònoma de Barcelona. DOI: 10.18148/sub/2019.v23i2.604.

Matushansky, Ora. 2022. On the semantics of inanimate gender. In Linnaea Stockall, Luisa Martí, David Adger, Isabelle Roy & Sarah Ouwayda (eds.), *QMUL Occasional Papers in Linguistics 47*. London: Queen Mary University of London. https://www.qmul.ac.uk/sllf/media/sllf-new/department-of-linguistics/hagit-borer-celebration/Matushansky.pdf.

McCawley, James D. 1996. An overview of "appositive" constructions in English. In Marek Przezdziecki & Lindsay J. Whaley (eds.), *Proceedings of the Eastern States Conference on Linguistics*, 195–211. Ithaca, NY: Cornell University.

McCawley, James D. 1998. *The syntactic phenomena of English*. Chicago, IL: University of Chicago Press.

Murphy, Dianna L. 2000. *The gender of inanimate indeclinable common nouns in Modern Russian*. Columbus, OH: Ohio State University. (Doctoral dissertation). http://rave.ohiolink.edu/etdc/view?acc_num=osu1488194825666262.

Nesset, Tore. 2004. Case assignment and image schemas: Russian temporal adverbials. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 28(2). 285–319. DOI: 10.1075/sl.28.2.02nes.

Nikunlassi, Ahti. 2000. On gender assignment in Russian. In Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen & Mirja Saari (eds.), *Gender in grammar and cognition*, 771–792. Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110802603.771.

Noailly, Michèle. 1991. 'L'énigmatique Tombouctou': nom propre et position de l'épithète. *Langue française* 92(1). 104–112. DOI: 10.3406/lfr.1991.6214.

Percus, Orin. 2011. Gender features and interpretation: A case study. *Morphology* 21(2). 167–196. DOI: 10.1007/s11525-010-9157-2..

Pesetsky, David. 2013. *Russian case morphology and the syntactic categories.* Cambridge, MA: The MIT Press.

Plungjan, Vladimir A. 2002. K semantike russkogo lokativa («vtorogo predložnogo» padeža). *Semiotika i informatika* 37. 229–254.

RNC. 2003—2022. *The Russian National Corpus.* https://ruscorpora.ru/.

Rothstein, Robert A. 1980. Gender and reference in Polish and Russian. In Catherine V. Chvany & Richard D. Brecht (eds.), *Morphosyntax in Slavic*, 79–97. Columbus, OH: Slavica.

Rozental, Ditmar E., Evgenija V. Džandžakova & Natalija P. Kabanova. 1998. *Spravočnik po pravopisaniju, proiznošeniju, literaturnomu redaktirovaniju.* Moscow: ČeRo.

Sloat, Clarence. 1969. Proper nouns in English. *Language* 45(1). 26–30. DOI: 10.2307/411749.

van Riemsdijk, Henk. 1998. Categorial feature magnetism: The endocentricity and distribution of projections. *The Journal of Comparative Germanic Linguistics* 2(1). 1–48. DOI: 10.1023/A:1009763305416.

Wang, Qiang. 2014. *Gender assignment of Russian indeclinable nouns.* Eugene, OR: University of Oregon. (MA thesis). https://core.ac.uk/download/pdf/36692535.pdf.

Wunderlich, Dieter. 1991. *Linguistics* 29(4). 591–622. DOI: 10.1515/ling.1991.29.4.591.

Zribi-Hertz, Anne & Loïc Jean-Louis. 2017. Les grammaires des noms de pays en martiniquais et en haïtien et la question du "prototype créole". *Etudes créoles* 35(1–2). 64–91. https://www.lpl-aix.fr/wp-content/uploads/Etudes_cr%5C%C3%5C%A9oles/zribi-hertz.pdf.

Zribi-Hertz, Anne & Loïc Jean-Louis. 2018. General locative marking in Martinican Creole (Matinitjè): A case study in grammatical economy. *Quaderni di Linguistica e Studi Orientali* 4. 151–176. DOI: 10.13128/QULSO-2421-73.

Zribi-Hertz, Anne & Loic Jean-Louis. 2014. Crosslinguistic studies on noun phrase structure and reference. In Patricia Cabredo Hofherr & Anne Zribi-Hertz (eds.), 269–315. Leiden: Brill. DOI: 10.1163/9789004261440_011.

Ora Matushansky

Zwarts, Joost & Yoad Winter. 2000. Vector space semantics: A model-theoretic analysis of locative prepositions. *Journal of Logic, Language and Information* 9(2). 169–211. https://www.jstor.org/stable/40180207.

Chapter 10

Simple imperfective verbs, the sequence of similar events interpretation, and Slavic aspectual composition

Stefan Milosavljević

University of Graz

The paper examines the so-called sequence of similar events (SSE) interpretation in Serbo-Croatian (SC), which emerges with telic predicates expressed by imperfective verbs in the presence of bare plural objects. I show that this is an interpretation that, just as in English, allows the use of both durative adverbials (DurAds) and time-span adverbials (TSAds) at the same time. I argue that TSAds, as standardly assumed, modify a telic event predicate, while DurAds merge once the predicate has been made homogeneous/atelic by the plural operator (contra MacDonald's 2008 claim that DurAds combine with telic predicates in such cases). The fact that the SSE interpretation is available in SC (or Slavic more generally) for imperfective verbs – including simple ones – suggests that in Slavic there is a syntactic projection responsible for telicity analogous to that in English, and telicity of a verbal predicate can be triggered by the quantity properties of its internal arguments.

Keywords: simple imperfective verbs, sequence of similar events interpretation, telicity, Serbo-Croatian, Slavic, English

1 Introduction

The temporal modification test (TMT) is one of the most standard diagnostics for (a)telicity, according to which durative adverbials (DurAds), often referred to as *for*-adverbials, modify atelic predicates, whereas time-span adverbials (TSAds),

Stefan Milosavljević

widely known as *in*-adverbials, modify telic predicates – but not vice versa, as in (1) from English.[1]

(1)   a.   John ran *for an hour* / **in an hour.*   ATELIC

      b.   John wrote a letter *in an hour* / **for an hour.*   TELIC

According to MacDonald (2008), these two adverbials can be combined in English under the so-called sequence of similar events (SSE) interpretation, illustrated in (2). The SSE interpretation, as analyzed in MacDonald (2008), emerges when a predicate is telic, with bare plurals (BPs) contributing an indefinite number of objects that can participate in each of the iterated subevents. The BP bears the feature [+q] (akin to the +SQA feature in Verkuyl 1972, 1999, standing for the specified quantity), and telicity emerges due to the so-called object-to-event mapping (OTEM) (in the sense of Verkuyl 1972). The contribution of DurAds, under such a view, amounts to assigning an indefinite number of repetitions to the telic event.

(2)   The guy drank cans of beer *in ten seconds for an hour straight.*

Given that DurAds and TSAds are expected to be in complementary distribution, as the same predicate cannot be both telic and atelic at the same time, their combination is (at least at first glance) unexpected. MacDonald (2008: 36) claims that in such cases (i.e. under the SSE reading), it is possible to combine DurAds with telic predicates,[2] rejecting the widely accepted generalization that DurAds require atelicity, also known as the homogeneity requirement (see Borer 2005, Csirmaz 2009, Landman & Rothstein 2010, 2012a,b, a.o.).

MacDonald (2008) claims that the SSE interpretation is available in English, but not in Russian (/ Slavic).[3] His argumentation, based on the analysis of Russian simple imperfective verbs and (prefixed) perfective verbs, proceeds in the following way: the SSE interpretation requires telic predicates, in Russian only

---

[1]The TMT is probably the most widely used test for telicity since it is employed regardless of the exact way telicity is approached – in terms of the event-argument homomorphism (e.g. Dowty 1991, Krifka 1992), the result state component (e.g. Pustejovsky 1995), atomicity (e.g. Rothstein 2008a,b), non-homogeneity/quantity (e.g. Borer 2005), scale features (e.g. Hay et al. 1999); for an overview of different approaches to telicity see, e.g., Arsenijević et al. (2013).

[2]MacDonald (2008: 33) also refers to other works arguing that DurAds are compatible with telic predicates under the iterative interpretation (Alsina 1999, Jackendoff 1996, Schmitt 1996, Tenny 1987, Vanden Wyngaerd 2001).

[3]MacDonald analyzes only Russian data, but many of his claims about Russian hold for Serbo-Croatian, which is why I generalize some of his claims in the present paper.

perfective verbs are telic, but perfectives are incompatible with the SSE interpretation. Imperfectives, on the other hand, are always atelic, and bare plurals, when combined with an imperfective verb, have a vague denotation associated with the mass noun interpretation (a group interpretation in which different parts of all the objects are affected at the same time), hence they never induce the SSE interpretation (MacDonald 2008: 147). He takes this (purported) difference between SSE in English and Russian as one of the main arguments for the claim that aspectual composition in these languages is radically different. Except for the SSE interpretation, aspectual composition in Russian differs from that in English in the unavailability of OTEM, i.e. in Russian an NP cannot affect the aspectual interpretation, and, consequently, this language lacks the syntactic projection responsible for inner aspect. Namely, in Russian, as stated by MacDonald, inner aspectual properties are determined through the event features, and only perfective verbs are equipped with the feature specifying the endpoint of the event, which triggers telicity. This feature is determined in the lexicon, before entering the narrow syntax, and can be either brought about by the lexical prefixes through the lexical derivational process, or lexically specified (in the case of simple perfective verbs). In English, on the other hand, there is an AspP between the vP and the vP[4] with which (features of) NPs interact. An NP yielding telicity has the quantity feature [+q], while an NP that fails to induce telicity has the [−q] feature. The relation between the NP and the AspP is established via Agree, leading to telicity if the NP is [+q], or atelicity if the NP is [−q] (MacDonald 2008, 2010, 2012). The idea that in Slavic, unlike in English, internal arguments of the verb (incremental direct objects and/or goal PPs) do not contribute to telicity is a fairly standard one (see e.g. Łazorczyk 2010, Rothstein 2016, Fleischhauer & Gabrovska 2019, among many others). Instead, it is a common view that telicity is triggered by prefixation (Borer 2005, Nossalik 2007, Łazorczyk 2010, Svenonius 2004a,b, Slabakova 2005, Arsenijević 2007b, Ramchand 2008, Fleischhauer & Gabrovska 2019, a.o.).

In this paper, I adopt the view of telicity as computed based on the quantity properties along the lines of Borer (2005): a predicate is telic (= Quantity) if it is non-homogeneous, i.e. if it is quantized or non-cumulative. The Quantity is assigned in the projection specifying the value of inner aspect – Asp$_Q$ in Borer (2005), or Q(uantification)P(hrase) in Arsenijević (2006b, 2007a, 2013). I assume that the presence of TSAds signals that a predicate is telic (non-homogeneous/bounded), i.e. that the QP is activated, following standard analyses (e.g. Krifka 1998, Borer 2005, Arsenijević 2006b, MacDonald 2008, Mittwoch 2010, 2013, 2019, among many others).

---

[4]The vP hosts the external argument in MacDonald's approach.

Stefan Milosavljević

Based primarily on data from Serbo-Croatian (SC), I show that the SSE interpretation is available in Slavic, and emerges in the presence of bare plurals when a telic predicate is expressed by an imperfective verb, as in (3).[5] I offer an analysis according to which the QP in SC/Slavic can be triggered by the quantity properties of internal arguments, and the SSE interpretation emerges once a telic predicate has been made homogeneous/atelic by the (covert) plural operator (in the sense of van Geenhoven 2004, 2005, Arsenijević 2006a). Under such an approach, TSADs, as expected, modify the QP, while DurAds combine with a homogeneous (plural) predicate. This is in line with the standard view that DurAds always combine with atelic/homogeneous structures (e.g. Borer 2005, Csirmaz 2009, Mittwoch 2010, Landman & Rothstein 2010, 2012a,b), and contra MacDonald's claim that in the case of SSE interpretation DurAds are compatible with telic predicates. The proposed analysis also implies that both Slavic and English employ a syntactic projection responsible for telicity (contra MacDonald 2008), i.e. aspectual composition in these languages is not radically different in this regard. Within the proposed system, prefixes are argued to be specifiers of singularity which combine with telic predicates rather than introducing telicity/perfectivity by themselves, as commonly assumed.

(3) Mika je pet minuta   pio   / iz-pi-ja-o
     Mika AUX five minute.GEN.PL drink.PTCP.M PREF-drink-SI-PTCP.M
     limenke   piva   za deset sekundi.
     can.ACC.PL beer.GEN.SG in$_{za}$ ten   second.GEN.PL
     'Mika drank cans of beer in ten seconds for an hour.'

The paper is organized as follows. In §2, I briefly introduce and discuss the relationship between (im)perfective verbs and telicity in Slavic. In §3, I analyze the SSE interpretation in SC. §4 addresses the broader picture, in particular how prefixed and biaspectual verbs fit into the proposed model of Slavic aspectual composition. §5 concludes the paper.

2 Slavic (im)perfectivity vs. telicity

The question of how telicity is assigned in Slavic is tightly related to the ongoing debate on the relationship between (im)perfectivity and telicity in this group of languages. As is well-known, in Slavic languages, verbs are traditionally divided into two classes: imperfective verbs (IVs) and perfective verbs (PVs).[6] A typical

---

[5]Unless explicitly indicated otherwise, all Slavic examples in the paper are from SC.

[6]In all examples from SC the superscripts [I] and [P] stand for IVs and PVs, respectively.

way in which aspect morphology is expressed in SC is illustrated in (4). The verb in (4a) consists of just a root, a theme vowel and an inflectional ending. Most such verbs are imperfective and can be perfectivized by prefixation, as in (4b). The prefixed verb can be imperfectivized by a secondary imperfectivizing suffix, as in (4c). Finally, an imperfective verb derived in this way can be made perfective again by prefixation, as illustrated in (4d). (The same holds, mutatis mutandis, for other Slavic languages.)

(4) a. vrš-i-ti^I
         perform-TV-INF
         'to perform'
     b. iz-vrš-i-ti^P
         PREF-perform-TV-INF
         'to perform/execute'
     c. iz-vrš-ava-ti^I
         PREF-perform-SI-INF
         'to perform/execute'
     d. po-iz-vrš-ava-ti^P
         PREF-PREF-perform-SI-INF
         'to perform/execute all'

According to one of the most standard tests, if a verb can be used as a complement of a phasal verb, it is imperfective; otherwise, it is perfective, as in (5); see Borik (2006), Łazorczyk (2010), Zinova (2021) for discussion of different tests. This will be the main diagnostics applied in this paper as well.

(5) Jovan je počeo   da   peva^I / *od-peva^P
     Jovan AUX begin.TV.PTCP.M COMP sing.PRS.3SG   PREF-sing.PRS.3SG
     pesmu.
     song.ACC.SG
     'Jovan began to sing a song.'

However, the exact status of PVs and IVs is largely debated. Probably the most common view is that they are grammaticalized forms of the (perfective and imperfective) grammatical (viewpoint/outer) aspect in Slavic (cf. e.g. Pereltsvaig 2005, Borik 2006, Ramchand 2008, Rothstein 2016, Minor et al. 2022). Łazorczyk (2010) and Tatevosov (2011, 2015) argue for separating grammatical aspect from the verb, since it can only emerge once the clausal architecture is fully established (given that the viewpoint depends on the interaction between the event time and

Stefan Milosavljević

the reference time).[7] I adopt Łazorczyk's and Tatevosov's view on divorcing the Slavic verb (aspectual morphology included) from grammatical aspect, and in the remainder of the paper I will not go into a deeper discussion of how grammatical aspect is to be analyzed.

Łazorczyk (2010) argues that IVs and PVs in Slavic are better accounted for in terms of telicity (PVs) vs. atelicity (IVs).[8] Typically, indeed, Slavic IVs and PVs are used as counterparts of English atelic vs. telic predicates, as shown in (6), with SC equivalents of English examples from (1) above.

(6) a. Džon je trčao[I]   sat   vremena   / *za
         John AUX run.TV.PTCP.M hour.ACC.SG time.GEN.SG in$_{za}$
         sat   vremena.
         hour.ACC.SG time.GEN.SG
         'John ran for an hour/ *in an hour.'

     b. Džon je na-pisao[P]   pismo   za sat
         John AUX PREF-write.TV.PTCP.M letter.ACC.SG in$_{za}$ hour.ACC.SG
         vremena   / *sat   vremena.
         time.GEN.SG   hour.ACC.SG time.GEN.SG
         'John wrote a letter in an hour/ *for an hour.'

In some contexts, however, IVs are compatible with TSAds, e.g. in habitual and general-factual uses, illustrated in (7–8) from SC. There are also some PVs that combine with DurAds, e.g. those with the delimitative prefix *po-*, as in (9). Strictly relying on the TMT, IVs in (7) and (8) could be treated as telic, while PVs like those in (9) should be atelic. These types of contexts have led some researches to claim that (im)perfectivity and telicity are independent systems in Slavic (e.g. Borik 2006, Gehrke 2008b,a, Ramchand 2008, Stanojević 2012, Fleischhauer &

---

[7]One of the classical definitions is that of Comrie (1976: 16), for whom "perfectivity indicates the view of a situation as a single whole, without distinction of the various separate phases that make up that situation; while the imperfective pays essential attention to the internal structure of the situation". According to a more formal definition, inspired by work of Reichenbach (1947), imperfective viewpoint arises when the Reference Time interval is included in the Event time interval (hence, we look at the event "from the inside"), whereas perfective viewpoint stands for the Event Time interval being contained within the Reference Time interval (hence the event is seen "from the outside") (cf. Klein 1994, Bhatt & Pancheva 2005, Łazorczyk 2010); for recent overviews, see Arche (2014a,b), Rothstein (2016).

[8]Such a view is assumed in MacDonald (2008) as well. Borer (2005) also analyzes Slavic perfectivity as Quantity/telicity and simple IVs as atelic, but she treats secondary imperfectives as species of outer aspect (in the sense of Verkuyl 1972).

Gabrovska 2019). In the remainder of the paper, I focus on IVs in telic environments.[9] Some authors argue that telicity in such cases is possible only with secondary imperfectives, claiming that it is the prefix that is responsible for telicity of imperfective verbs (e.g. Stanojević 2012, Fleischhauer & Gabrovska 2019). Yet, examples like (8) show that telic readings emerge also in the absence of prefixes (and see Pereltsvaig 2000, Szucsich 2000, 2001, Braginsky & Rothstein 2008, Arsenijević 2023 for similar kinds of examples).

(7) Pera je uvek iz-pad-a-o[I]   iz   igre   za par
     Pera AUX always out-fall-SI-PTCP.M from game.GEN.SG in$_{za}$ couple
     minuta.
     minute.GEN.PL
     'Pera has always been out of the game in a couple of minutes.'

(8) Žika se već peo[I]   na to   brdo   za pola sata.
     Žika REFL already climb.PTCP.M on that hill.ACC.SG in$_{za}$ half hour.GEN.SG
     'Žika (has) already climbed that hill in half an hour.'

(9) Mika je juče   po-sedeo[P]   kod nas par   sati.
     Mika AUX yesterday DEL-sit.TV.PTCP.M at us couple hour.GEN.PL
     'Mika stayed at our place for two hours yesterday.'

One way to account for the diversity of readings IVs are associated with is to assume that they are unspecified for telicity, rather than atelic (as in Łazorczyk 2010). In other words, what is traditionally referred to as an imperfective verb is just a verbalized structure unspecified for both telicity and grammatical aspect. This stance is similar in spirit to the proposal of Arsenijević (2018) according to which Slavic IVs are unmarked for grammatical aspect, i.e. ambiguous between imperfective and perfective aspect. In this paper, I focus on simple forms, but the analysis can be extended to secondary imperfectives straightforwardly once secondary imperfectivizing suffixes are analyzed as re-verbalizing morphemes (Arsenijević 2018), i.e. sequences of theme vowels (Simonović et al. 2022).

3 The sequence of similar events interpretation

Having removed the obstacle presented by the view that IVs are incompatible with telicity, we are in a position to revisit the claim that the SSE interpretation

---

[9]See Milosavljević (2022) for a detailed analysis of perfectives with the delimitative prefix *po-* as telic predicates.

Stefan Milosavljević

is not available in Slavic. Examples with both simple (10–11) and prefixed verbs (12) show that the SSE interpretation can arise in SC as well. Just as in the case described in MacDonald (2008), in all such examples, there is an indefinite number of telic events iterated within the time interval specified by DurAds. Actually, using IVs is the only available way to express the SSE interpretation in SC in the presence of DurAds, since perfective forms cannot be combined with DurAds in such contexts.[10]

(10) Kandidat je dva sata   gubio[I]   [partije   šaha]
     candidate AUX two hour.PCL lose.TV.PTCP.M game.ACC.PL chess.GEN.SG
     od   velemajstora   za manje od dva minuta.
     from grandmaster.GEN.SG in_za less than two minute.PCL
     'The candidate lost [chess games] to the grandmaster in less than two
     minutes for two hours.'

(11) Ana je za Luninu   svadbu   ceo   dan
     Ana AUX for Luna.POSS wedding.ACC.SG whole day.ACC.SG
     pravila[I]   torte   za manje od pola sata.
     make.TV.PTCP.F cake.ACC.PL in_za less than half hour.GEN.SG
     'For Luna's wedding, Ana made cakes in less than half an hour the whole
     day.'

(12) Pera je dva minuta   iz-pi-ja-o[I]   limenke   piva
     Pera AUX two minute.PCL out-drink-SI-PTCP.M can.ACC.PL beer.GEN.SG
     za   deset sekundi.
     in_za ten   second.GEN.PL
     'Pera drank cans of beer in ten seconds for two minutes.'

---

[10]An anonymous reviewer suggests that the SSE reading with simple IVs is only marginally acceptable in Russian (i.e. possibly admissible in some contexts), and that it slightly improves when a secondary imperfective is used. The reviewer points out that a more natural way to express the SSE interpretation in Russian is when the argument introduces distributivity and not just plurality, as in (i). (The progressive form as a translation of the IV *čitat* 'read' is provided by the reviewer. MacDonald 2008 consistently uses simple past forms for such readings in English.)

(i) Nedelju   čital[I]   po vypusku   za čas.
    week.ACC.SG read.PTCP.M on issue.DAT.SG in_za hour.ACC.SG
    'For a week I was reading issues in an hour.'

Crucially for the purposes of the present paper, these examples once again show that there is no ban on using simple IVs to express telic predicates. This further suggests that the differences in the degree of acceptability of bare plurals with the SSE interpretation between SC and Russian (and possibly other Slavic languages) are not to be sought in the impossibility of IVs to express telic predicates, as analyzed in MacDonald (2008).

Before moving on to the exact analysis of the SSE interpretation in SC, a few clarification points are in order. The availability of the SSE interpretation in examples like (10–12) does not mean that other interpretations of IVs with BPs are impossible. For instance, there are at least three possible interpretations of example (13): (i) the SSE interpretation, with the distributive interpretation of the BP (one song per event); (ii) the iterative interpretation in which Mika recites a set of songs repeatedly, but each set is different; (iii) the iterative interpretation in which Mika recites the same set of songs repeatedly. In both (ii) and (iii) the BP is interpreted collectively. (Later in this section, we will see how the difference between the distributive and the collective interpretation of the BP reflects its different syntactic status.) The third type of interpretation is an instance of the multiple event interpretation which MacDonald (2008: 41) labels the sequence of identical events interpretation (= SIE interpretation), since the same object is implicated in each of the iterated subevents (i.e. the BP is interpreted specifically/definitely). The SIE interpretation is also available with singular specific objects, as in (14).

(13) Mika je recitovao   pesme   sat   vremena.
      Mika aux recite.tv.ptcp.m song.acc.pl hour.acc.sg time.gen.sg
      'Mika recited songs for an hour.'

(14) Mika je recitovao   pesmu   sat   vremena.
      Mika aux recite.tv.ptcp.m song.acc.sg hour.acc.sg time.gen.sg.
      'Mika recited a song for an hour.'

Finally, outside of multiple event interpretations discussed above, BPs in SC, just as in English, may receive a vague denotation which MacDonald (2008: 46, 147) refers to as a M(ass)N(oun) interpretation. This is illustrated by the SC example (15) similar to those discussed for English in MacDonald (2008: 46). Under the MN interpretation of (15), it does not have to be the case that Mika made multiple dragons – actually, (15) would still be true if he worked on making only one dragon without ever finishing it. As stated in MacDonald (2008: 46), predicates in examples like (15) are interpreted as activities, with a taste of a habitual interpretation.

(15) Mika je u slobodno vreme   pravio   papirne
      Mika aux in free   time.acc.sg make.tv.ptcp.m paper.poss
      zmajeve.
      dragon.acc.pl
      'Mika made paper dragons (in his free time).'

Stefan Milosavljević

3.1 The SSE interpretation and plural telic predicates

In this subsection, I propose an analysis of telicity in SC as computed on the basis of the quantity properties of internal arguments, which straightforwardly captures the possibility to get the SSE interpretation with simple IVs, i.e. in the absence of prefixes. For the sake of simplicity, I focus on the derivation from the point at which the *v*P is instantiated. I use the *v*P as a verbalizing projection (i.e. devoid of external arguments, cf. Harley 2013), assuming that theme vowels in Slavic are verbalizers (with Svenonius 2004a, Biskup 2019, Kovačević et al. 2022, Milosavljević & Arsenijević 2022). I will primarily use examples with measuring-out direct objects – traditional incremental themes, since these are the most typical cases where the aspectual role of internal arguments can be observed, and they are the main kind of examples used by MacDonald (2008) to illustrate the SSE interpretation in English.

When the verb merges with an incremental theme object equipped with the [+q] feature (in the sense of Verkuyl 1972, MacDonald 2008), the projected *v*P is culminative, i.e. it denotes a culminative predicate, as in Figure 1. Otherwise, the *v*P is non-culminative, see Figure 2. Examples of culminative predicates include *praviti tortu* 'make a cake', *gubiti meč* 'lose the match', *peti se na brdo* 'climb the hill', whereas non-culminative *v*Ps are those without a bounded internal argument, e.g. *jesti šećer* 'eat sugar' (with an object interpreted as mass), or typical intransitive activities such as *trčati* 'run', *spavati* 'sleep'. Many of non-culminative predicates can easily be turned into culminative ones, providing the [+q] internal argument is composed with a given verb, e.g. *trčati maraton* 'run a marathon' or *spavati popodnevnu dremku* 'sleep an afternoon nap'.[11]

$$v\text{P}_{[+\text{Cul}]}$$

NP$_{[+q]}$ $\quad$ *v*′

|

v
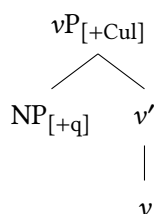
Figure 1: Culminative *v*P

Culminative *v*Ps give rise to telicity (i.e. the projection of the QP) by default. This is achieved by the movement of the accusative object from its base-generated

---

[11]Culminativity in this sense is close in spirit to telicity at the level of *v*P (as the locus of the telic description) in Arsenijević (2006b), the lexical aspect in the sense of Rothstein (2016), or completability in Janda (2011).

vP[−Cul]

(NP[−q]) v'

v
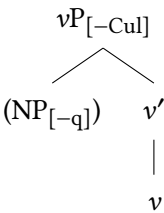
Figure 2: Non-culminative vP

position (Spec$v$P) to the specifier position of the QP, where it checks the [+q] feature (in the sense of Pereltsvaig 1999, 2000, see also Travis 2005), as illustrated in Figure 3. (Culminative predicates fail to trigger the projection of QP in progressive contexts, as briefly discussed in §3.2.)

QP

NP[+q]   Q$'$

Q vP

t$_{NP[+q]}$ v'

v
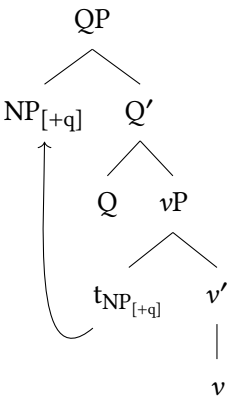
Figure 3: QP

Once the QP is projected, the derivation can proceed in two ways, both of which lead to the projection of the Num(eral)P, a phrase responsible for number in the verbal domain: the QP composes with the plural operator, yielding a plural telic predicate, or it composes with the prefix, giving rise to a singular telic predicate. The former option is how the SSE interpretation arises, and it is addressed in detail in the remainder of this subsection. The singular telicity is briefly analyzed in §4.1, since it sheds light on the overall system of the computation of telicity in Slavic.

The structure of the plural telic predicate is shown in Figure 4. Here I build on the insights of van Geenhoven (2004, 2005) and Arsenijević (2006a), who propose that distributive multiple event interpretations (referred to as SSE and SIE

Stefan Milosavljević

interpretations in this paper, following MacDonald 2008) are instances of verbal plurality, or (silent) pluractionality, which is a verbal counterpart of nominal plurality.[12] Although many languages, including English and SC, do not make use of the overt plural marking directly on the verb, there are languages with such a morphological makeup, e.g. West Greenlandic, discussed in van Geenhoven (2004).
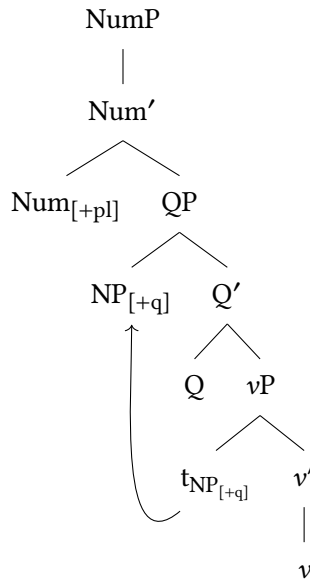
Figure 4: Plural NumP

I adopt Arsenijević's (2006a) analysis according to which in the case of the SSE interpretation the plural gets lexicalized on the noun. The relation between the plurality head and the plural marking on the noun is established via a binding relation. This is possible because the object NP, being unspecific, does not establish its referential properties outside the eventuality it is bound by, including the number specification (see Arsenijević 2006a for technical details).

Under this approach, the plurality is responsible for the homogenizing effects, enabling DurAds to combine with such a predicate. As also pointed out by van Geenhoven (2004: 142–143), plural (pluractional in her terminology) predicates are like mass nouns (i.e. cumulative and divisive), which makes them unbounded, i.e. non-homogeneous/atelic.

---

[12]For related ideas, see also Landman (2000), Rothstein (2004, 2008a), and references therein.

There are several advantages of the proposed analysis of the SSE interpretation. Let me start by comparing MacDonald's and the approach proposed here. According to MacDonald (2008: 50), BPs (i) must be [+q] in order to trigger telicity, and they (ii) introduce an indefinite number of objects, while DurAds (i) combine with a telic predicate, and (ii) contribute an indefinite number of repetitions of the telic event since they force the event to continue for the amount of time they specify. This division of labor between BPs and DurAds in contributing the SSE interpretation implies that this type of multiple event interpretation is not available in the absence of DurAds, contrary to the fact: DurAds only make it more prominent, i.e. pragmatically salient. In my approach, just as in MacDonald's, the internal argument contributes the [+q] feature, but it is the plural in the verbal domain that is responsible for the multiple events interpretation, bringing about the homogeneity effects in this way. DurAds then provide the time interval within which these multiple events occur. While I remain agnostic with respect to the exact way DurAds should be represented in this case,[13] the crucial point is that they do not compose with a telic predicate, rather – they enter the derivation once the plural homogeneous predicate has been formed. Consequently, my proposal preserves the standard analyses of both TSAds (which modify telic predicates) and DurAds (which compose with atelic predicates). In addition, the proposal preserves the view that a bounded internal argument contributes the [+q] feature and that the bare plural makes a predicate homogeneous, with the difference that in this case the plurality applies directly in the verbal domain.

The proposed analysis straightforwardly captures the difference between the distributive and collective interpretation of BPs in contexts sketched in (13) above: they are instances of the event plurality and the object plurality, respectively. Namely, under the collective interpretation, the plural is interpreted on the noun, and the plurality operator scopes over it, which delivers interpretations according to which multiple objects are affected within every counting unit of a plural event. In addition, we will see in §4.1 that only BPs which reflect the NP plurality can be used in the scope of prefixes – just as expected if prefixes, as assigners of singularity, are in complementary distribution with plural operators.

An anonymous reviewer raises the question of how the proposed analysis of the verb plurality as lexicalized on the noun under the SSE interpretation captures the fact that there are plurality interpretations dissociated from plural morphology on the noun, e.g. the SIE interpretation in the sense of MacDonald (2008: 41); recall that this is a multiple events interpretation in which the same object is implicated in each of the iterated subevents, illustrated in (16) from SC.

---

[13]A plausible candidate would be an aspectual projection responsible for repetitivity immediately above the NumP, i.e. the $AspP_{repetitive}$ in the sense of Cinque (1999).

Stefan Milosavljević

(16) Mika je recitovao pesmu sat vremena (za
Mika AUX recite.TV.PTCP.M song.ACC.SG hour.ACC.SG time.GEN.SG in$_{za}$
pet minuta).
five minute.GEN.PL
'Mika recited a song for an hour (in five minutes).'

I propose that in this case the aspectual composition proceeds in the same way as under the SSE interpretation: telicity is triggered by the specified quantity brought about by the internal argument, and the telic *v*P (= QP) is then pluralized by the (covert) plural operator. Unlike in the case of SSE interpretation, in the SIE contexts the plural fails to be lexicalized on the noun since in this case the object NP is specific, i.e. it establishes referential properties independently of the eventuality, including its own number specification (cf. Arsenijević 2006a).

3.2 Culminative *v*Ps and "failed" telicity

The default pattern sketched in Figure 3 – culminative *v*Ps yielding telicity – fails to be established only if the progressive-like kind of operator intervenes, yielding a stative interpretation in the sense of Ramchand (2018) (see also Parsons 1990). Ramchand (2018: 58–59) proposes an *ing*P projection above the *v*P, still within the first phase (i.e. within the domain of event description) for English progressive constructions, thus moving away from the standard analyses of the progressive as an instantiation of grammatical aspect (see also Ramchand & Svenonius 2014, Ramchand 2017). In analogy with this proposal, examples with culminative *v*Ps that have the interpretation analogous to the English progressive (as in (17)) can be accounted for by assuming a (null) progressive operator immediately above the *v*P, as in Figure 5, which blocks the projection of the QP.

(17) Maja je juče dva sata (*za dva sata) pravila[I]
Maja AUX yesterday two hour.PCL in$_{za}$ two hour.PCL make.TV.PTCP.F
sneška, kad je sneg odjednom počeo[P] da se
snowman.ACC.SG when AUX snow suddenly begin.TV.PTCP.M COMP REFL
topi[I] i prekinuo[P] njen poduhvat.
melt.PRS.3SG and interrupt.TV.PTCP.M her endeavor.ACC.SG
'Yesterday, Maja had been making a snowman for two hours when the snow suddenly began to melt and interrupted her endeavor.'

Hence, *v*Ps like *praviti sneška* 'make a snowman' in the progressive contexts are culminative, but they are not telic, since the projection of the QP fails. I assume

StateP
|
State′
State$_{[Prog]}$ vP
NP$_{[+q]}$ v′
|
v
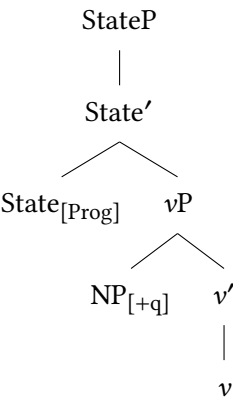
Figure 5: StateP

with Ramchand (2018: 58) that for every event description P, the progressive (operator) introduces an Identifying State as "a stative eventuality that manifests sufficient cognitive/perceptual identifiers of the event property P", which is why I label such a projection StateP in Figure 5. The proposed view straightforwardly explains why culminative predicates in SC in examples like (17) can be used with DurAds, but cannot be modified by TSAds: TSAds require the projection of the QP, which fails in this case. DurAds, on the other hand, are felicitous, since in progressive contexts they can be analyzed as scoping over the progressive operator, modifying the Identifying State of a snowman building event, as also pointed out by an anonymous reviewer.

4 Broadening the picture: singular telic predicates

The analysis presented in §3 enables accounting for telicity in Slavic and Germanic languages in a unified way: telicity *can* be triggered by the properties of internal arguments. In other words, it is not the case that in Germanic languages properties of internal arguments are crucial in computing telicity, whereas in Slavic they have no effect whatsoever, as standardly assumed (see e.g. MacDonald 2008, Łazorczyk 2010, Rothstein 2016). It should be emphasized, however, that the proposed analysis does not imply that internal arguments with a specified quantity are the only way to assign telicity: e.g. it can be triggered by some measure adverbials (cf. e.g. Pereltsvaig 2000 for Russian, Milosavljević 2022 for SC). This again is similar with what we find in Germanic languages, where various

types of adverbials can trigger the projection of QP (see e.g. Borer 2005). However, the role of internal arguments in affecting telicity in SC described in the previous section was constrained only to plural contexts, which, at first glance, contrasts with the state of affairs we find in English.[14] I propose that internal arguments retain their role in aspectual composition in Slavic in singular contexts as well. This is achieved by analyzing Slavic prefixes as scoping over the QP triggered by internal arguments, as proposed in §4.1. Another context where singular telicity emerges in the absence of prefixes productively is with biaspectual verbs, which will be briefly discussed in §4.2.

4.1 Prefixes and singular telic predicates

Prefixless incremental theme verbs discussed in previous subsections usually have prefixed variants, and such pairs are typically referred to as aspectual pairs, which have the same meaning and differ only with respect to the aspectual value. Some aspectual pairs from SC are provided in (18).

(18)   a.   graditi[I]   kuću   / sa-graditi[P]   kuću
            build.TV.INF house.ACC.SG   with-build.TV.INF house.ACC.SG
            'build a house'

        b.   praviti[I]   tortu   / na-praviti[P]   tortu
            make.TV.INF cake.ACC.SG   on-make.TV.INF cake.ACC.SG
            'make a cake'

        c.   gubiti[I]   meč   / iz-gubiti[P]   meč
            lose.TV.INF match.ACC.SG   out-lose.TV.INF match.ACC.SG
            'lose a match'

        d.   čitati[I]   knjigu   / pro-čitati[P]   knjigu
            read.TV.INF book.ACC.SG   through-read.TV.INF book.ACC.SG
            'read a book'

These prefixes are often labeled as purely perfectivizing prefixes (PPPs) and are

---

[14]I assume that other syntactic contexts in which IVs are used in telic environments (e.g. habitual and general-factual uses) include a (potential) repetition of the same (telic) event type/kind (see Milosavljević 2019), hence they are also based on the plurality of telic *v*Ps (but see Arsenijević 2023 for a different view). However, the exact analysis of these cases goes beyond the scope of the present paper. For a unified treatment of habitual and general-factual readings of imperfectives in Russian, see Minor (2019). For accounts of the general-factual meaning that employ the notion of event kind, see Mehlig (2013), Mueller-Reichau (2013, 2015).

typically analyzed as semantically empty.[15] In this subsection, I propose that PPPs compose with telic predicates, and that they are specifiers of the projection responsible for number in the verbal domain, where they specify a telic verbal predicate for singularity (via specifier-head agreement in the sense of Borer 2005), as shown in Figure 6. I opt for an analysis of prefixes as specifiers rather than heads building on Milosavljević's (in prep) proposal that the semelfactive suffix *-nu* is an exponent of the head of this projection, and the two morphemes can be combined (e.g. *od-gur-nu-ti* [PREF-push-SEM-INF] 'push away').[16]
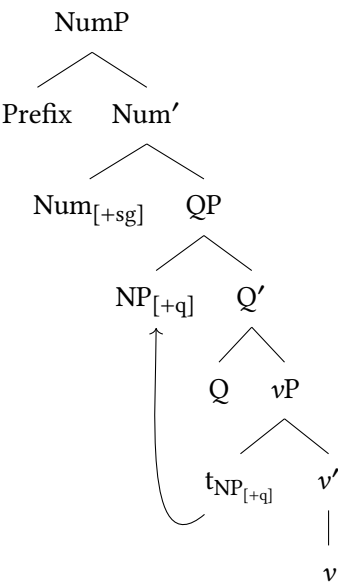
Figure 6: Singular NumP

Let me situate this proposal against some common analyses in the literature. As is well known, the object of PVs gets an obligatorily bounded interpretation. On the common view, such an interpretation is usually analyzed as brought about either by the prefix or the perfective aspect, a process inverse to what we see in English: instead of the object determining the interpretation of the verbal predicate, the verbal predicate determines the properties of the object (see Szucsich 2001, 2002, Łazorczyk 2010, MacDonald 2008, Rothstein 2016). IVs, on

---

[15]I use the term PPPs descriptively here – it does not necessarily mean that these prefixes are devoid of meaning; for detailed semantic analyses of prefixes traditionally claimed to be semantically empty, see e.g. Endresen et al. (2012), Janda & Lyashevskaya (2013), Miljković (2021).

[16]See Svenonius (2008) for additional arguments in favor of the analysis of prefixes as specifiers.

the other hand, do not impose restrictions on the interpretation of the direct object, i.e. it may be both unbounded and bounded and can be optional with same verbs, e.g. *pisati (pismo)* 'to write (a letter)' or *čitati (knjigu)* 'to read (a book)'. To account for this difference in the status of objects of PVs and IVs, Basilico (2008), for instance, proposes that they are introduced by different heads at different points in the syntactic derivation: the direct object of PVs is introduced by the (affixed) Root, while the direct object of IVs is introduced by the *v* categorizing head.

My approach to prefixation is closer to an alternative view, suggested in Krifka (1992: 50) and Verkuyl (1999: 102). For these authors, prefixes, as perfective operators, require the *v*P they combine with to be quantized/terminative (which is possible only if the object NP is bounded). According to Verkuyl (1999: 126–127), until the ASP-node, which hosts a prefix, merges, the derivation of the verb has not yet been completed, and the bounded object, though necessary, is not itself sufficient to bring about the terminative/bounded *v*P. Only after the perfective prefix is added, the perfective terminative (= telic) *v*P arises. Hence, in this approach, although the prefix merges with a terminative/quantized/telic *v*P, such a *v*P is always realized only in perfective contexts, after the prefix has been merged.

The view according to which prefixes scope over bounded/telic predicates has several advantages. First, it recognizes the role of internal arguments in affecting telicity in both English and Slavic, without a need for specifying the inverse operation for the latter group of languages. Second, the object NP of IVs and PVs need not to be analyzed as generated in different ways (as in Basilico 2008), since, as we have seen, its obligatory nature with PVs follows from the fact that the prefix picks out the *v*P with a bounded NP object. In this way, PVs are actually aspectual counterparts of IVs with a bounded object.[17] Finally, if the QP has its telic aspectual status independently prior to merging with the prefix, we expect to find it in some other syntactic contexts as well. The SSE interpretation, analyzed in §3, provides exactly the kind of context that employs the QP divorced from prefixes. Hence, while I share with Krifka (1992) and Verkuyl (1999) the view that prefixes scope above complex (telic) *v*Ps, in my approach prefixation is not the only syntactic context that enables telic predicates to show up. Prefixes are specifiers of singularity, and as such they are in complementary distribution with plural telic predicates presented in §3. For instance, BPs with prefixes in SC cannot give rise to the SSE interpretation, rather – they always receive a collective interpretation.

---

[17]E.g. it is not the case that the verbs *pisati*[I] and *na-pisati*[P] 'to write' are themselves aspectual pairs, rather *na-pisati*[P] + $NP_{[+q]}$ is a counterpart of *pisati*[I] + $NP_{[+q]}$.

This is expected if the BP giving rise to the SSE interpretation reflects the plurality of events, while under the collective interpretation it reflects the plurality of objects. As expected, in the latter case the prefix is able to compose with a predicate whose object is expressed by a BP when the BP is bounded (which is usually contextually provided), as in (19).

(19) Pera je   na-pravio[P]   torte.
      Pera aux PREF-make.TV.PTCP.M cake.ACC.PL
      'Pera made the cakes.'

Except for their complementary distribution with plural predicates, I prefer the analysis of prefixes as markers of singularity rather than markers of perfectivity, as in Krifka (1992) and Verkuyl (1999) (see also Slabakova 2005), because the prefix does not guarantee perfectivity, since in many cases the prefixed QP can undergo secondary imperfectivization (and the prefixed QP is realized as perfective only upon the inclusion of the reference time). Moreover, the view of prefixes as singulative morphemes also accords well with some recent approaches to prefixes as (morphemes of the same kind as) numeral classifiers (see Dickey & Janda 2015).[18]

4.2 Biaspectual verbs and telicity

Biaspectual verbs (BVs) are traditionally analyzed as verbs that can be either perfective or imperfective, depending on the syntactic context (see Janda 2007, Kolaković 2018, Zinova 2021, Starý 2017, a.o.). In terms of the system presented in this paper, BVs can be used in both singular and plural telic environments, as in (20) and (21) from SC. Since they are simple, i.e. unprefixed forms, BVs can be taken as additional evidence that telicity in Slavic can emerge in the absence of prefixes. Some extensive corpus-based studies show that BVs are based on culminative *v*Ps (see Grickat 1957/8, Janda 2007, Kolaković 2018), which also supports the view that telicity is based on culminativity, which is in turn based on the contribution of internal arguments, as proposed in §3.

---

[18]In this section, I focused on PPPs with incremental theme verbs. In Milosavljević (2022, in prep), I argue that Slavic prefixes generally compose with telic predicates. In short, just as internal arguments are not the only way to trigger telicity, the proposal that prefixes combine with telic predicates does not mean that they must combine with telic predicates whose telicity is triggered by internal arguments. For instance, in Milosavljević (2022, in prep) an analysis of the delimitative prefix *po-* in Slavic is proposed according to which this prefix combines with the QP triggered by DurAds or some contextually provided quantity.

Stefan Milosavljević

(20) Pera je malopre downlodovao film za 15
Pera AUX just now download.TV.PTCP.M movieACC.SG in_za 15
minuta.
minute.GEN.PL
'Pera just downloaded a movie in 15 minutes.'

(21) Pera je ceo dan downlodovao filmove za 15
Pera AUX whole day.ACC.SG download.TV.PTCP.M movie.ACC.PL in_za 15
minuta.
minute.GEN.PL
'Pera downloaded movies in 15 minutes the whole day.'

While the plural telicity emerges when the QP is combined with the plural operator, it remains an open question how the singular reading emerges in the absence of prefixes (or the semelfactive suffix). A possible solution is to assume that singularity is triggered by a variable-like anaphoric element – following the argumentation in Stanley (2000), Stanley & Szabo (2000), a.o., that all effects of extra-linguistic context on the truth-condition are represented at LF.[19]

5 Conclusion

In this paper, I examined the so-called sequence of similar events interpretation in Serbo-Croatian SC, which emerges in the presence of bare plural objects when a telic predicate is expressed by an imperfective verb. I showed that this is an interpretation that, just as in English, allows the use of both durative adverbials and time-span adverbials at the same time. I proposed that, as standardly assumed, TSAds modify a telic event predicate, while DurAds in such cases merge once the predicate has been made homogeneous/atelic by the plural operator (contra Mac-Donald's 2008 claim that DurAds combine with telic predicates in such cases). The fact that the SSE interpretation is possible in Serbo-Croatian (and at least some other Slavic languages), and is realized by employing imperfective verbs – including simple ones (i.e. those without prefixes) – suggests that in Slavic there is a syntactic domain responsible for telicity analogous to that in English (contra MacDonald 2008).

---

[19]An alternative option would be to assume a null prefix to account for singular telic uses or "perfective" uses of bi-aspectuals, as suggested in Grickat (1957/8, 1966/7), Łazorczyk (2010).

Abbreviations

ACC = accusative	PCL = paucal
AUX = auxiliary	POSS = possesive
BP = bare plural	PTCP = participle
BV = biaspectual verb	PREF = prefix
COMP = complementizer	PV = perfective verb
DAT = dative	REFL = reflexive
DEL = delimitative (prefix)	SC = Serbo-Croatian
DurAds = durative adverbials	SG = singular
GEN = genitive	SEM = semelfactive
F = feminine	SI = secondary imperfectivizing (suffix)
IV = imperfective verb	SIE = sequence of identical events
LOC = locative	SSE = sequence of similar events
M = masculine	TSAds = time-span adverbials
OTEM = object-to-event mapping	TV = theme vowel
PL = plural	

Acknowledgments

References

Alsina, Alex. 1999. On the representation of event structure. In Tara Mohanan & Lionel Wee (eds.), *Grammatical semantics: Evidence for structure in meaning*, 7–122. Stanford: CSLI Publications.

Arche, María J. 2014a. About the primitives of aspect across languages. *Natural Language and Linguistic Theory* 32(3). 711–733. DOI: 10.1007/s11049-014-9242-z.

Arche, María J. 2014b. The construction of viewpoint aspect: The imperfective revisited. *Natural Language and Linguistic Theory* 32(3). 791–831. DOI: 10.1007/s11049-013-9209-5.

Stefan Milosavljević

Arsenijević, Boban. 2006a. Event-argument homomorphism as a concord phenomenon. In Paul J. E. Dekker & Hedde H. Zeijlstra (eds.), *Concord and the syntax semantics interface*, 45–56. Amsterdam / Tübingen: ESSLLI.

Arsenijević, Boban. 2006b. *Inner aspect and telicity*. Utrecht: LOT Publications.

Arsenijević, Boban. 2007a. A unified analysis of two classes of Slavic verb prefixes. In Sylvia Blaho, Luis Vicente & Erik Schoorlemmer (eds.), *Proceedings of ConSOLE XIV*, 21–36. Leiden: ConSOLE.

Arsenijević, Boban. 2007b. Slavic verb prefixes are resultative. *Cahiers Chronos* 17. 197–213. DOI: 10.1163/9789401204446_013.

Arsenijević, Boban. 2013. The *for*-phrase. *Philologia Mediana* 5. 515–538.

Arsenijević, Boban. 2018. *Serbo-Croatian suffix -va and the illusion of the imperfective specification*. Presented at Ambigo, University of Göttingen, July 4-6, 2018.

Arsenijević, Boban. 2023. Specification of telicity in Serbo-Croatian, without null prefixes. In Petr Biskup, Marcel Börner, Olav Mueller-Reichau & Iuliia Shcherbina (eds.), *Advances in formal Slavic linguistics 2021*, ?–? Berlin: Language Science Press.

Arsenijević, Boban, Berit Gehrke & Rafael Marín. 2013. The (de)composition of event predicates. In Boban Arsenijević, Berit Gehrke & Rafael Marín (eds.), *Studies in the composition and decomposition of event predicates*, 1–26. Dordrecht: Springer. DOI: 10.1007/978-94-007-5983-1_1.

Basilico, David. 2008. The syntactic representation of perfectivity. *Lingua* 118(11). 1716–1739. DOI: 10.1016/j.lingua.2007.02.012.

Bhatt, Rajesh & Roumyana Pancheva. 2005. *LSA 130: The syntax and semantics of aspect*. Lecture notes. http://web.mit.edu/rbhatt/www/lsa130/l1.pdf.

Biskup, Petr. 2019. *Prepositions, case and verbal prefixes: The case of Slavic*. Amsterdam: John Benjamins. DOI: 10.1075/la.255.

Borer, Hagit. 2005. *The normal course of events: Structuring sense Vol. II*. Oxford: Oxford University Press.

Borik, Olga. 2006. *Aspect and reference time*. Oxford: Oxford University Press.

Braginsky, Pavel & Susan Rothstein. 2008. Vendlerian classes and the Russian aspectual system. *Journal of Slavic Linguistics* 16(1). 3–55. DOI: 10.1353/jsl.0.0009.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

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

Csirmaz, Anikó. 2009. Homogeneity and flexibility in temporal modification. In José M. Brucart, Anna Gavarró & Jaume Solà (eds.), *Merging features: Computation, interpretation and acquisition*, 235–253. Oxford: Oxford University Press.

Dickey, Stephen M. & Laura A. Janda. 2015. Slavic aspectual prefixes and numeral classifiers: Two kinds of lexico-grammatical unitizers. *Lingua* 168. 57–84. DOI: 10.1016/j.lingua.2015.09.005.

Dowty, David. 1991. Thematic proto-roles and argument structure. *Language* 67(3). 547–619. DOI: 10.2307/415037.

Endresen, Anna, Laura A. Janda, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nesset & Svetlana Sokolova. 2012. Russian purely aspectual prefixes: Not so empty after all? *Scando-Slavica* 58(2). 231–291. DOI: 10.1080/00806.740247.

Fleischhauer, Jens & Ekaterina Gabrovska. 2019. Perfectivity and atelicity – the role of perfective aspect in aspectual composition. In Emmanuelle Roussel, Adeline Patard & Rea Peltola (eds.), *Cross-linguistic perspectives on the semantics of grammatical aspect*, 97–126. Leiden / Boston: Brill.

Gehrke, Berit. 2008a. *Ps in motion: On the semantics and syntax of P elements and motion events*. Utrecht: LOT Publications.

Gehrke, Berit. 2008b. *Ps in motion: On the semantics and syntax of P elements and motion events*. Utrecht: LOT. https://www.lotpublications.nl/ps-in-motion-ps-in-motion-on-the-semantics-and-syntax-of-p-elements-and-motion-events.

Grickat, Irena. 1957/8. O nekim vidskim osobenostima srpskohrvatskog glagola. *Južnoslovenski filolog* XXII. 65–130.

Grickat, Irena. 1966/7. Prefiksacija kao sredstvo gramatičke (čiste) prefiksacije (načelna razmatranja i savremena srpskohrvatska građa). *Južnoslovenski filolog* XXVII(1–2). 185–221.

Harley, Heidi. 2013. External arguments and the mirror principle: On the distinctness of Voice and *v*. *Lingua* 125. 34–57. DOI: 10.1016/j.lingua.2012.09.010.

Hay, Jennifer, Christopher Kennedy & Beth Levin. 1999. Scalar structure underlies telicity in "Degree Achievements". In Tanya Matthews & Devon Strolovitch (eds.), *Proceedings of SALT 9*, 127–144. Ithaca, NY: CLC Publications. DOI: 10.3765/salt.v9i0.2833.

Jackendoff, Ray. 1996. The proper treatment of measuring out, telicity, and perhaps even quantification in English. *Natural Language & Linguistic Theory* 14(2). 305–354. DOI: 10.1007/BF00133686.

Janda, Laura A. 2007. What makes Russian bi-aspectual verbs special? In Dagmar Divjak & Agata Kochanska (eds.), *Cognitive paths into the Slavic domain*, 83–109. Berlin / New York: Walter de Gruyter.

Stefan Milosavljević

Janda, Laura A. 2011. Completability and Russian aspect. In Marcin Grygiel & Laura A. Janda (eds.), *Slavic linguistics in a cognitive framework*, 13–37. Frankfurt am Main: Peter Lang.

Janda, Laura A. & Olga Lyashevskaya. 2013. Semantic profiles of five Russian prefixes: *Po-, s-, za-, na-, pro-. Journal of Slavic Linguistics* 21(2). 211–258. https://www.jstor.org/stable/pdf/24602127.pdf.

Klein, Wolfgang. 1994. *Time in language.* London: Routledge.

Kolaković, Zrinka. 2018. *Dvoaspektni glagoli – razlike između (p)opisa u priručnicima i stanja u korpusu s posebnim osvrtom na uporabu izvornih govornika.* Regensburg / Zagreb: Universität Regensburg / University of Zagreb. (Doctoral dissertation).

Kovačević, Predrag, Stefan Milosavljević & Marko Simonović. 2022. *Theme-vowel minimal pairs show argument structure alternations.* Under review in *Journal of Linguistics.*

Krifka, Manfred. 1992. Thematic relations as links between nominal reference. In Ivan A. Sag & Anna Szabolsci (eds.), *Lexical matters*, 29–53. Stanford, CA: CSLI Publications.

Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (ed.), *Events and grammar*, 197–235. Dordrecht: Kluwer.

Landman, Fred. 2000. *Events and plurality: The Jerusalem lectures.* Dordrecht: Kluwer.

Landman, Fred & Susan Rothstein. 2010. Incremental homogeneity in the semantics of aspectual *for*-phrases. In Malka Rappaport Hovav, Edit Doron & Ivy Sichel (eds.), *Lexical semantics, syntax, and event structure*, 229–251. Oxford: Oxford University Press.

Landman, Fred & Susan Rothstein. 2012a. The felicity of aspectual *for*-phrases – part 1: Homogeneity. *Language and Linguistics Compass* 6(2). 85–96. DOI: 10.1002/lnc3.324.

Landman, Fred & Susan Rothstein. 2012b. The felicity of aspectual *for*-phrases – part 2: Incremental homogeneity. *Language and Linguistics Compass* 6(2). 97–112. DOI: 10.1002/lnc3.323.

Łazorczyk, Agnieszka. A. 2010. *Decomposing Slavic aspect: The role of aspectual morphology in Polish and other Slavic languages.* Los Angeles, CA: University of Southern California. (Doctoral dissertation).

MacDonald, Jonathan E. 2008. *The syntactic nature of inner aspect: A minimalist perspective.* Amsterdam / Philadelphia: John Benjamins Publishing Company.

MacDonald, Jonathan E. 2010. The aspectual influence of the noun: (A)telicity, (a)symmetry, incrementality and universality. *Language and Linguistics Compass* 4(9). 831–845. DOI: 10.1111/j.1749-818X.2010.00227.x.

MacDonald, Jonathan E. 2012. Nouns affect aspect syntactically. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz & Alex Trueman (eds.), *Coyote papers 20: Proceedings of the poster session of the West Coast Conference on Formal Linguistics 29*, 73–80. Somerville, MA: Cascadilla Proceedings Project. http://hdl.handle.net/10150/253423.

Mehlig, H. R. 2013. Obščefaktičeskoe i edinično-faktičeskoe značenija nesoveršennogo vida v russkom jazyke. *Vestnik Moskovskogo Universiteta. Ser 9. Filologija* 4. 19–47.

Miljković, Vanja. 2021. *Kognitivnolingvističko istraživanje glagolskih prefiksa u savremenom srpskom jeziku*. Belgrade: University of Belgrade. (Doctoral dissertation).

Milosavljević, Stefan. 2019. O upotrebi iterativnih glagola u perfektu za označavanje izbrojivog mnoštva radnji i njihovoj konkurenciji sa svršenim glagolima (na primerima iz novinarskog stila srpskog jezika). *Naučni sastanak slavista u Vukove dane* 48(3). 149–191. DOI: 10.18485/msc.2019.48.3.ch8.

Milosavljević, Stefan. 2022. The delimitative prefix *po*-, durative adverbials, and Slavic aspectual composition. In Annie Holtz, Iva Kovač & Rasmus Puggaard-Rode (eds.), *Proceedings of ConSOLE XXX*. Leiden: Leiden University Centre for Linguistics.

Milosavljević, Stefan. in prep. *Specification of event duration and aspectual composition in Slavic*. In preparation. Graz: University of Graz. (Doctoral dissertation).

Milosavljević, Stefan & Boban Arsenijević. 2022. What differentiates Serbo-Croatian verbal theme vowels: content or markedness? *Glossa: a journal of general linguistics* 7(1). 1–36. DOI: 10.16995/glossa.8535.

Minor, Serge. 2019. *The semantics of grammatical aspect in Slavic (and beyond)*. Handout for a class given at EGG 2019, Wrocław. http://www.eggschool.org/classes/serge-minor-the-semantics-of-grammatical-aspect-in-slavic-and-beyond/.

Minor, Serge, Natalia Mitrofanova & Gillian Ramchand. 2022. Fine-grained time course of verb aspect processing. *Plos One* 17(2). e0264132. DOI: 10.1371/journal.pone.0264132.

Mittwoch, Anita. 2010. Event measurement and containment. In Malka Rappaport Hovav, Edit Doron & Ivy Sichel (eds.), *Lexical semantics, syntax, and event structure*, 252–266. Oxford: Oxford University Press.

Mittwoch, Anita. 2013. On the criteria for distinguishing accomplishments from activities, and two types of aspectual misfits. In Boban Arsenijević, Berit Gehrke & Rafael Marín (eds.), *Studies in the composition and decomposition of event predicates*, 27–48. Dordrecht: Springer. DOI: 10.1007/978-94-007-5983-1_2.

Stefan Milosavljević

Mittwoch, Anita. 2019. Aspectual classes. In Robert Truswell (ed.), *The Oxford handbook of event structure*, 31–49. Oxford: Oxford University Press. DOI: 10.1093/oxfordhb/9780199685318.013.4.

Mueller-Reichau, Olav. 2013. Why *kratnost'*? On Russian factual imperfectives. In Uwe Junghanns, Dorothee Fehrmann, Denisa Lenertová & Hagen Pitsch (eds.), *Formal description of Slavic languages: The ninth conference. Proceedings of FDSL 9, Göttingen 2011*, 191–210. Frankfurt am Main: Peter Lang.

Mueller-Reichau, Olav. 2015. Pseudo-incorporation in Russian? Aspectual competition and bare singular interpretation. In Olga Borik & Berit Gehrke (eds.), *The syntax and semantics of pseudo-incorporation*, 262–295. Leiden: Brill. DOI: 10.1163/9789004291089_009.

Nossalik, Larissa. 2007. Slavic perfective prefixes: Are they telicity markers? In Milica Radišić (ed.), *Proceedings of the 2007 annual conference of the Canadian Linguistic Association*, 1–12. Ottawa, Ontario: Canadian Linguistic Association. https://cla-acl.ca/actes/actes-2007-proceedings.html.

Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.

Pereltsvaig, Asya. 1999. Cognate objects in Russian: is the notion 'cognate' relevant for syntax? *Canadian Journal of Linguistics* 44(3). 267–291. DOI: 10.1017/S0008413100017333.

Pereltsvaig, Asya. 2000. On accusative adverbials in Russian and Finnish. In Artemis Alexiadou & Peter Svenonius (eds.), *Adverbs and adjunction*, 155–176. Potsdam: Universitätsverlag Potsdam.

Pereltsvaig, Asya. 2005. Aspect lost, aspect regained: Restructuring of aspectual marking in American Russian. In Paula Kempchinsky & Roumyana Slabakova (eds.), *Aspectual inquiries*, 369–395. Springer: Dordrecht.

Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.

Ramchand, Gillian C. 2008. Perfectivity as aspectual definiteness: Time and the event in Russian. *Lingua* 118(11). 1690–1715. DOI: 10.1016/j.lingua.2007.03.008.

Ramchand, Gillian C. 2017. The event domain. In Roberta D'Alessandro, Irene Franco & Ángel J. Gallego (eds.), *The verbal domain*, 233–254. Oxford: Oxford University Press.

Ramchand, Gillian C. 2018. *Situations and syntactic structures: Rethinking auxiliaries and order in English*. Cambridge, MA: MIT Press.

Ramchand, Gillian C. & Peter Svenonius. 2014. Deriving the functional hierarchy. *Language Sciences* 46. 152–174. DOI: 10.1016/j.langsci.2014.06.013.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. Berkeley, CA: University of California Press.

Rothstein, Susan. 2004. *Structuring events: A study in the lexical semantics of aspect.* Malden, MA: Blackwell.

Rothstein, Susan. 2008a. Telicity, atomicity and the Vendler classification of verbs. In Susan Rothstein (ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*, 43–77. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Rothstein, Susan. 2008b. Two puzzles for a theory of lexical aspect: Semelfactives and degree achievements. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds.), *Event structures in linguistic form and interpretation*, 175–197. Berlin / New York: Walter de Gruyter.

Rothstein, Susan. 2016. Aspect. In Maria Aloni & Paul Dekker (eds.), *The Cambridge handbook of formal semantics*, 342–368. Cambridge: Cambridge University Press.

Schmitt, Cristina J. 1996. *Aspect and the syntax of noun phrases.* College Park, MD: University of Maryland. (Doctoral dissertation).

Simonović, Marko, Stefan Milosavljević & Boban Arsenijević. 2022. Serbo-Croatian secondary imperfectivisers consist of theme vowels. *Journal of Slavic Linguistics.* Accepted for publication pending some revisions.

Slabakova, Roumyana. 2005. Perfective prefixes: What they are, what flavors they come in, and how they are acquired. In Steven Franks, Frank Gladney & Mila Tasseva-Kurktchieva (eds.), *Formal Approaches to Slavic Linguistics 13: The South Carolina meeting*, 324–341. Ann Arbor, MI: Michigan Slavic Publications.

Stanley, Jason. 2000. Context and logical form. *Linguistics and Philosophy* 23(4). 391–434. DOI: 10.1023/A:10057.

Stanley, Jason & Zoltán Szabo. 2000. On quantifier domain restriction. *Mind and Language* 15(2). 219–261. DOI: 10.1111/1468-0017.00130.

Stanojević, Veran. 2012. Perfektivnost i teličnost u srpskom jeziku sa osvrtom na situaciju u francuskom jeziku. *Naučni sastanak slavista u Vukove dane* 41(1). 141–155.

Starý, Zdeněk. 2017. Biaspectuals revisited. *Studie z aplikované lingvistiky – Studies in Applied Linguistics* 8(1). 111–123.

Svenonius, Peter. 2004a. Slavic prefixes and morphology: An introduction to the Nordlyd volume. *Nordlyd* 32(2). 177–204. DOI: 10.7557/12.67.

Svenonius, Peter. 2004b. Slavic prefixes inside and outside VP. *Nordlyd* 32(2). 205–253. DOI: 10.7557/12.68.

Svenonius, Peter. 2008. Russian prefixes are phrasal. In Gerhild Zybatow, Luka Szucsich, Uwe Junghanns & Roland Meyer (eds.), *Proceedings of Formal Descriptions of Slavic Languages 5*, 526–537. Frankfurt am Main: Peter Lang.

Stefan Milosavljević

Szucsich, Luka. 2000. *Nominale Adverbiale im Russischen. Syntax, Semantik und Informationsstruktur*. Leipzig: Universität Leipzig. (Doctoral dissertation).

Szucsich, Luka. 2001. Adjunct positions of nominal adverbials in Russian. In Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn & Luka Szucsich (eds.), *Current issues in formal Slavic linguistics* (Linguistik International 5), 106–116. Frankfurt am Main: Peter Lang.

Szucsich, Luka. 2002. Case licensing and nominal adverbials in Slavic. In Toman Jindřich (ed.), *Formal Approaches to Slavic Linguistics 10: The Ann Arbor meeting 2001*, 249–270. Ann Arbor, MI: Michigan Slavic Publications.

Tatevosov, Sergei. 2011. Severing perfectivity from the verb. *Scando-Slavica* 57(2). 216–244. DOI: 10.1080/00806765.2011.631782.

Tatevosov, Sergei. 2015. Severing imperfectivity from the verb. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebova (eds.), *Slavic grammar from a formal perspective: The 10th anniversary FDSL conference*, 465–494. Frankfurt am Main: Peter Lang. DOI: 10.3726/978-3-653-05335-7.

Tenny, Carol Lee. 1987. *Grammaticalizing aspect and affectedness*. Cambridge, MA: MIT. (Doctoral dissertation).

Travis, Lisa de Mena. 2005. Articulated *v*ps and the computation of aspectual classes. In Paula Kempchinsky & Roumyana Slabakova (eds.), *Aspectual inquiries*, 69–93. Springer: Dordrecht.

Vanden Wyngaerd, Guido. 2001. Measuring events. *Language* 77(1). 61–90. https://www.jstor.org/stable/pdf/3087028.pdf.

van Geenhoven, Veerle. 2004. *For*-adverbials, frequentive aspect, and pluractionality. *Natural Language Semantics* 12. 135–190. DOI: 10.1023/B:NALS.0000031389.36427.af.

van Geenhoven, Veerle. 2005. Atelicity, pluractionality, and adverbial quantification. In Henk J. Verkuyl, Henriëtte de Swart & Angeliek van Hout (eds.), *Perspectives on aspect*, 107–124. Dordrecht: Springer.

Verkuyl, Henk J. 1972. *On the compositional nature of the aspects*. Dordrecht: Springer.

Verkuyl, Henk J. 1999. *Aspectual issues: Studies on time and quantity*. Stanford, CA: CSLI publications.

Zinova, Yulia. 2021. *Russian verbal prefixation: A frame semantic analysis*. Berlin: Language Science Press.

Chapter 11

Slavic ʟ-periphrases: Linguistic change and variation

Hagen Pitsch

Georg-August-Universität Göttingen

The present paper addresses the variation in ʟ-periphrases looking at a broad range of modern Slavic languages. Based on a thorough description, a typological division between AUXILIARY LANGUAGES and PARTICLE LANGUAGES is proposed. The difference between them is then motivated by sketching diachronic scenarios of linguistic change and subsequently given syntactic analyses. In sum, the paper reveals a remarkable variation that has so far been widely disregarded from a theoretical point of view.

Keywords: past, perfect, preterit, conditional, tense, mood, linguistic change

1 Introduction

A common thread of present-day Slavic languages is that they use ʟ-PERIPHRASES to express specific tenses and moods, namely the future, the preterit (perfect or generalized past), and the conditional; see (1), (2), and (3), respectively.[1,2,3]

(1)   ʟ-future

    a. Md-ã   pisa-ł-a.
       FUT-1SG write-ʟ-SG.F
       'I shall be writing.'   (Kashubian)

---

[1]Unless otherwise indicated, examples are constructed by myself.

[2]Kashubian has several stem variants for its future auxiliary. Besides *md-* illustrated in (1a), the stem can be *będ-*, *bãd-*, or *bd-* (see Stone 1993: 776–777).

[3]BCMS has ʟ-future forms only in temporal and conditional clauses like (1c). If perfective, they are interpreted as a future perfect, otherwise as a simple future.

Hagen Pitsch

b. Będ-ą   prosi-ł-y o pokój   na wyższych piętrach.
   FUT-3PL ask-L-PL.F for silence.ACC on higher floors.LOC
   'They shall ask for silence on the higher floors.'   (Polish)

c. Kad bude-mo govori-l-i s   Marijom, sve će   biti   jasno.
   when FUT-1PL speak-L-PL.M with Maria.INS all FUT.3SG be.INF clear
   'When we speak with Marija (in the future), everything will be clear.'
                                                    (BCMS; Browne 1993: 331)

d. Prosi-l-a bo-š   za dopust.
   ask-L-SG.F FUT-2SG for vacation.ACC
   'You shall apply for leave.'   (Slovene)

(2) L-preterit

a. Ima-l-a   je   razgovor sa   psihologom.
   have-L-SG.F be.3SG talk.ACC with psychologist.INS
   'She had a talk with her psychologist.'   (BCMS)

b. Wona je   dźěła-ł-a   jako bibliotekarka.
   she be.3SG work-L-SG.F as   librarian.SG.F
   'She has been working as a librarian.'   (Upper Sorbian)

c. Ma-l   veľké šťastie.
   have-L.SG.M big luck.ACC
   'He had enormous luck.'   (Slovak)

d. Koly ty narody-l-a-s'?
   when you give-birth-L-SG.F-REFL
   'When were you born?'   (Ukrainian)

(3) L-conditional

a. Ima-l-a   bi-h   sigurno napad   panike.
   have-L-SG.F COND-1SG certainly attack.ACC panic.GEN
   'I would certainly have a panic attack.'   (BCMS)

b. Da   bi   se   v žlici   vode   utopi-l!
   PART COND REFL in spoon.LOC water.GEN drown-L.SG.M
   'May you drown in a spoonful of water!'   (Slovene; Priestly 1993: 431)

c. Ma-l-a   by   som   ísť   do postele.
   have-L-SG.F COND be.1SG go.INF to bed.GEN
   'I should go to bed.'   (Slovak)

d. Ja {b} c'oho   ne skaza-v   {by}.
   I COND this.GEN NEG say-L.SG.M COND
   'I would not have said that.'   (Ukrainian; Amir-Babenko 2007: 158)

298

The general format of L-periphrases is given in (4).[4]

(4)   (AU) V~L~

In (4), "AU" and "V~L~" stand for AUXILIARY UNIT and VERBAL L-FORM, respectively. I prefer "AU" over the more familiar notion "auxiliary" due to its being more neutral: Saying "auxiliary", one usually thinks of an inflected verb form. While AUs in Slavic L-periphrases can indeed be inflected verb forms – and always are in L-futures –, they may also be noninflected, in which case they are commonly called PARTICLES. This is why in (5), which shows the general morphological makeup of Slavic AUs, I put the agreement categories in brackets.

(5)   AU: stem(-PERSON/NUMBER)

Finally, (6) depicts the general morphological structure of V~L~s.

(6)   V~L~: stem-L-NUMBER(/GENDER)

The variation between absent and present and – if present – inflected and noninflected AUs is the main issue of the present paper. It aims at (i) giving a detailed description of this cross-Slavic variation, (ii) reconstructing it from a theoretical point of view, (iii) integrating the perspective of linguistic change, and (iv) putting forward a syntax-based formalization of the auxiliary/particle distinction as manifested in Slavic L-periphrases, most notably the L-preterit and the L-conditional.[5]

To that end, §2 gives a detailed description of the relevant L-periphrases. In §3, I sketch a set of diachronic scenarios of linguistic change which are likely to have given rise to the present-day situation. Finally, §4 presents my claims as to the syntax of Slavic L-periphrases. §5 summarizes the paper.

2 Description

This part describes the cross-Slavic variation in the L-preterit and L-conditional, leaving aside the L-future due to the fact that it does not display any variation in the languages that have it.

---

[4]As the auxiliary unit (AU) is absent in a subset of cases, "periphrasis" seems to be partly inadequate to characterize the verb forms in question. Lateron, however, I will show that, syntactically, all cases are indeed bipartite/analytic.

[5]The motivation for having such a formalization is that the linguistic notions of auxiliary and particle, while ubiquitous in the literature, are still very vague.

Hagen Pitsch

2.1 L-preterit

2.1.1 The general picture

All present-day Slavic languages exhibit verb forms related to the Late Proto-Slavonic present-perfect periphrasis, which consisted of a present-tense form of the BE-auxiliary (showing person and number; e.g., Old Church Slavic *jesmĭ* 'am', *jesi* 'are', etc.) and the main verb V_L. While these forms retain their original present-perfect meaning in Bulgarian and the standard varieties of BCMS, Macedonian, and Sorbian, they have developed into a general(ized) past in the remaining languages/varieties. To conflate these notions, I use the term L-PRETERIT.

On the other hand, the modern languages show considerable variation concerning the shape of the AU: Some have clitic BE-auxiliaries inflected for person and number throughout the paradigm; see Table 1.[6] A smaller subset of languages has inflected auxiliaries everywhere in the paradigm except for the 3rd person; see Table 2. Finally, East Slavic languages and Kashubian spoken by younger speakers lack AUs altogether; see Table 3.

Table 4 provides an overview. It shows that the variation cuts across the traditional division between South and West Slavic, while the East Slavic languages behave uniformly. Kashubian comes in two varieties: Kashubian-A (literary language and older speakers) aligns with "minor" West Slavic languages, whereas Kashubian-B (younger speakers) resembles East Slavic. In §2.1.2 and §2.1.3, respectively, I describe the diachronic background underlying the absence of AUs in the L-preterit in East Slavic and Kashubian-B. On the other hand, Polish AUs stand out from AUs in the remaining languages in that they seem to be suffixes. Again, some diachronic background is supplied not only to track the changes underlying the present-day situation but also to arrive at assumptions about the syntax of the relevant AUs. That background is presented in §2.1.4.

2.1.2 Auxiliary loss in East Slavic

Beginning with the 11th century, Old East Slavic gradually lost the present-tense paradigm of *byti* 'be' (see, a.o., Issatchenko 1940; Ivanov 1964: 391; Borkovskij & Kuznecov 1965: 298; Sokolova 2017). First of all, this process affected the third-person forms (3SG *jest´*, 3PL *sjat´*), the remaining forms (1SG *esm´*, 2SG *esi*, 1PL *esme*,

---

[6]As to Kashubian, Stone (1993: 776) notes that the variant AU + V_L ("Kashubian-A") is widely used in the literature and characteristic in the spoken language of the older generation, while elsewhere, the preterit consists of V_L only ("Kashubian-B" in Table 3). See also Menzel (2013) and Bartelik (2015). Note that descriptions vary. Thus, Lubaś (2002: 268) and Breza (2009: 174) make no mention of the AU-less variant.

Table 1: L-preterit with inflected AU throughout

| | | SG | | PL | |
		AU	V_L	AU	V_L
BCMS	1	*sam*	*pisala*	*smo*	*pisale*
'write'	2	*si*	*pisala*	*ste*	*pisale*
	3	*je*	*pisala*	*su*	*pisale*
Bulgarian	1	*săm*	*čela*	*sme*	*čeli*
'read'	2	*si*	*čela*	*ste*	*čeli*
	3	*e*	*čela*	*sa*	*čeli*
Slovene	1	*sem*	*pohvalila*	*smo*	*pohvalile*
'praise'	2	*si*	*pohvalila*	*ste*	*pohvalile*
	3	*je*	*pohvalila*	*so*	*pohvalile*
Lower Sorbian	1	*som*	*słyšała*	*smy*	*słyšali*
'hear'	2	*sy*	*słyšała*	*sćo*	*słyšali*
	3	*jo*	*słyšała*	*su*	*słyšali*
Upper Sorbian	1	*sym*	*dźěłała*	*smy*	*dźěłali*
'work'	2	*sy*	*dźěłała*	*sće*	*dźěłali*
	3	*je*	*dźěłała*	*su*	*dźěłali*
Kashubian-A	1	*jem*	*robiła*	*jesmë*	*robiłë*
'make, work'	2	*jes*	*robiła*	*jesta*	*robiłë*
	3	*je*	*robiła*	*są*	*robiłë*

2PL *este*) following suit.[7] As a consequence, the present-perfect paradigm, formerly periphrastic, lost the AU without substitution, turning it effectively into a synthetic form consisting exclusively of V_L. The scheme in (7) depicts this change, using the 2nd singular preterit of the verb *čitati* 'read' as an illustration.

(7) AU + V_L ⟶ V_L
    jesi čitala *čitala*

The absence/loss of the AU had further implications: For one thing, the by now solitary V_L, once a participle, acquired the role of the finite verb. Nonetheless,

---

[7]The same sequence of changes can be traced for Old Polish (see Decaux 1955: 127–128; Migdalski 2006: 41).

Hagen Pitsch

Table 2: L-preterit without inflected AU in the 3rd person

| | | SG | | PL | |
		AU	V$_L$	AU	V$_L$
Macedonian	1	*sum*	*molela*	*sme*	*molele*
'ask'	2	*si*	*molela*	*ste*	*molele*
	3		*molela*		*molele*
Czech	1	*jsem*	*udělala*	*jsme*	*udělaly*
'make'	2	*jsi*	*udělala*	*jste*	*udělaly*
	3		*udělala*		*udělaly*
Slovak	1	*som*	*volala*	*sme*	*volali*
'call'	2	*si*	*volala*	*ste*	*volali*
	3		*volala*		*volali*
Polish	1	*-m*	*prosiła*	*-śmy*	*prosiły*
'ask'	2	*-ś*	*prosiła*	*-ście*	*prosiły*
	3		*prosiła*		*prosiły*

Table 3: L-preterit without AU throughout

| | | SG | PL |
		V$_L$	V$_L$
Belarusian	1	*čytala*	*čytali*
'read'	2	*čytala*	*čytali*
	3	*čytala*	*čytali*
Russian	1	*skazala*	*skazali*
'say'	2	*skazala*	*skazali*
	3	*skazala*	*skazali*
Ukrainian	1	*bula*	*buly*
'be'	2	*bula*	*buly*
	3	*bula*	*buly*
Kashubian-B	1	*robiła*	*robiłë*
'make, work'	2	*robiła*	*robiłë*
	3	*robiła*	*robiłë*

Table 4: Cross-Slavic variation in the ʟ-preterit

| | AU in … person | | |
	1st	2nd	3rd
BCMS	•	•	•
Slovene	•	•	•
Bulgarian	•	•	•
Macedonian	•	•	
Czech	•	•	
Slovak	•	•	
Polish	•	•	
Lower Sorbian	•	•	•
Upper Sorbian	•	•	•
Kashubian-A	•	•	•
Kashubian-B			
Belarusian			
Russian			
Ukrainian			

it retained its original nominal agreement (number and gender), thus leaving person unexpressed in the verbal domain. This in turn added significance to (the use of) overt personal pronouns to avoid ambiguity (see Issatchenko 1940: 193).[8]

AU-lessness is characteristic of all present-day East Slavic languages. To capture it for Russian, Junghanns (1995) claims a new agreement pattern in the past tense with the person feature underspecified; see (8a) as opposed to the "canonical" non-past pattern in (8b).

(8) a. $[-\text{PAST}] \rightarrow [\alpha\text{PERSON}, \beta\text{NUMBER}, \varnothing\text{GENDER}]$
     b. $[+\text{PAST}] \rightarrow [\varnothing\text{PERSON}, \beta\text{NUMBER}, \gamma\text{GENDER}]$

(see Junghanns 1995: 174)

Due to the loss of the present-tense paradigm of *byti*, present-day East Slavic languages are also "copula-less".[9] In emphatic (verum and contrastive focus; see

---

[8]It is not clear from the literature whether the loss of the *byti*-forms fostered the more frequent use of personal pronouns or whether it was the other way around. Fortunately, this issue is of minor importance for the present investigation.

[9]Issatchenko (1940: 192) applies Leonard Bloomfield's term EQUATIONAL PREDICATIONS to the resulting copula structures.

Hagen Pitsch

Geist 2007: 127) contexts, however, the former 3sg form – Belarusian *ёsc´*, Russian *est´*, Ukrainian *je* – survives but has lost its agreement specification and thus occurs in all persons and numbers. This leads Issatchenko (1940: 192) to the statement that Russian *est´* "has lost its verbal character; it has become an impersonal particle."[10] Moreover, equational and identificational clauses involve a (de-pronominal) particle: Belarusian *hèta*, Russian *èto*, Ukrainian *ce*. Crucially, all these particles can by no means function as AUs in the ʟ-preterit. These observations will be taken up in §4.

2.1.3 Auxiliary drop in Kashubian-B

There are two ways to form the ʟ-preterit in Kashubian: Either V_L is combined with an inflected be-auxiliary as schematized in (9a) or V_L is used alone as in (9b) (see Breza & Treder 1981: 130–134).[11]

| (9) | a. | AU V_L | (Kashubian-A) |
| | b. | V_L | (Kashubian-B) |

Rittel (1970: 100) assumes that the situation in (9b) was fostered by the increased use of personal pronouns (allegedly induced by language contact with German; see also Nomachi 2014), which resembles the development described for East Slavic languages in §2.1.2. It is fair to assume that the co-existence of the two patterns documents a linguistic change in progress which parallels the change in Old East Slavic sketched in (7). An analogous scheme for Kashubian-B is given in (10) using the 1pl preterit of the verb *robic* 'make, work' as an illustration.

| (10) | AU | + V_L | $\longrightarrow V_L$ |
| | *jesmë* | *robiłë* | *robiłë* |

Summarizing so far, East Slavic languages and Kashubian-B share the AU-less type of ʟ-preterit due to the loss or drop, respectively, of the be-auxiliary. Their current ʟ-preterit consists exclusively of V_L and lacks overt person agreement.

The next section shows that the diachronic reshaping of the present-tense be-paradigm can give rise to yet another, rather peculiar, situation.

---

[10]Especially speakers from the Western Ukraine may employ *je* in place of the zero copula in all persons and numbers. Elsewhere, the zero copula is the default choice.

[11]According to Stone (1993: 776), pattern (9a) is characteristic of older speakers, while younger speakers prefer (9b). See Menzel (2013) for a corpus-based discussion. Crucially, there is no Polish-like variant of the preterit with reduced ("suffixal") agreement markers (see §2.1.4).

2.1.4 Auxiliary reduction in Polish

Polish reshaped the present-tense forms of its ʙᴇ-verb far more profoundly than the remaining Slavic languages. To put it informally, Polish reduced the inherited present-tense forms of *być* 'be' to such an extent that their modern reflexes function as mere agreement markers. While this state of affairs is well-investigated (see, a.o., Decaux 1955; Rittel 1970; Andersen 1987; Piskorz et al. 2013), the actual nature of the "new" agreement markers is still a subject of controversy (see, a.o., Booij & Rubach 1987; Borsley & Rivero 1994; Embick 1995; Franks & Bański 1999; an overview and discussion is provided in Abramowicz 2008: 5–9).

The relevant changes seem to have started in the 14th century (Rittel 1970: 103; Migdalski 2006: 41). While most Slavic languages reduced the present-tense forms of their ʙᴇ-verb – especially when used as an auxiliary – to clitics, their reduction went even further in Polish. This process gave rise to two coexisting sets of present-tense ʙᴇ-forms in Old Polish dubbed ᴏʀᴛʜᴏᴛᴏɴɪᴄ and ᴀᴛᴏɴɪᴄ, respectively, by Andersen (1987). Modern Polish retains only the latter. Table 5 (from Migdalski 2006: 41; see also Rittel 1970: 99–103; Andersen 1987: 24; Embick 1995: 3) summarizes the diachronic development.

Table 5: Diachronic development of Polish present-tense *być*-forms

		16th century orthotonic	atonic	today
	1	*jeśm*	*-(e)śm/-(e)m*	*-(e)m*
ꜱɢ	2	*jeś*	*-(e)ś*	*-(e)ś*
	3	*je/jest/jeść*	—	—
	1	*jesm(y)*	*-(e)smy*	*-(e)śmy*
ᴘʟ	2	*jeść*	*-(e)śće*	*-(e)ście*
	3	*są*	—	—

Embick (1995) emphasizes that the modern atonic forms are restricted to the ʟ-preterit (and ʟ-conditional; see §2.2), which is illustrated in Table 6.

The lost orthotonic forms were compensated for by a completely new present-tense paradigm for *być* based on the former third-singular form *jest* suffixed with the "new" atonic agreement markers from Table 5; see Table 7.[12]

---

[12]As shown in the rightmost column, some Polish dialects employ the original third-plural form *są-* as plural stem (Migdalski 2006: 42).

Hagen Pitsch

Table 6: The Polish ʟ-preterit

	SG	PL
1	*prosiła-m*	*prosiły-śmy*
2	*prosiła-ś*	*prosiły-ście*
3	*prosiła*	*prosiły*

Table 7: Modern Polish present-tense *być*-forms (full verb)

	SG	PL	(dialectal)
1	*jest-em*	*jest-eśmy*	(*są-śmy*)
2	*jest-eś*	*jest-eście*	(*są-śće*)
3	*jest*	*są*	

Moreover, the atonic agreement markers occur on *powinien* 'should, ought', a former predicative adjective that developed into a modal quasi-verb; see (11a). Rarely, they can also fulfil the function of the copula as in (11b).

(11) a. Nie powinn-a-m   (by-ł-a)   jechać.   (Polish)
          NEG obliged-SG.F-1SG be-ʟ-SG.F go.INF
          'I should not (have) go(ne).'

      b. Zmęczony-m i   głodny(-m).
          tired-1SG   and hungry.1SG
          'I am tired and hungry.'   (Migdalski 2006: 234)

Migdalski (2006: 275–276) claims that the third-person forms of the full verb *być* – i.e. *jest* and *są* – do not specify any person feature. Moreover, judging from their combinatorial potential, *jest* is completely underspecified ([αNUMBER]), whereas *są* is marked as plural ([PL]).[13] From these facts, Migdalski (2006: 275) concludes that *jest* and *są* are in a lower syntactic position as compared to the other forms of the paradigm, and that they have to raise in the structure to adjoin to the relevant person/number marker (*-m*, *-ś*, etc.). Only in the 3rd person do they always remain in situ, as there is no (overt) agreement marker to adjoin to. Finally,

---

[13]This becomes apparent by the fact that *jest* can combine with any person and number marker, whereas *są* is restricted to the plural as shown in Table 7.

considering data like (11b), it seems fair to assume that *jest*/*są* may also be absent or left unpronounced under specific circumstances.

In §2.1.2, I referred to Issatchenko (1940: 192) who claims that Russian *est´* "has become an impersonal particle." I suggest that the facts about Polish *jest* and *są* just mentioned point in the same direction, though Polish seems to be in an intermediate stage: While in isolation, *jest* seems to have lost its verbal character (agreement) just like Russian *est´*, it can still be "upgraded" into a (composite) verb by merging it with an agreement marker. On the other hand, *są* retains number agreement anyway. The parallels and differences allow the determination of the syntactic positions of the elements in question in §4.

To sum up thus far, Polish reduced its original present-tense BE-forms to atonic agreement markers which occur in the L-preterit but also in the "new" present-tense paradigm of *być*, and which are likely to be located in a relatively high (functional) syntactic position.

2.2 L-conditional

2.2.1 The general picture

Unlike the L-preterit, the L-conditional has an AU in all Slavic languages. However, variation obtains in the presence or absence, respectively, of person/number agreement on it. Moreover, if there is agreement, there is variation as to its shape.

Languages with inflected conditional AUs are illustrated in Tables 8 and 9: The AUs in Table 8 are clearly synthetic. Most of these AUs are inherited from Late Proto-Slavic, which employed aorist BE-forms as auxiliaries in the periphrastic conditional.[14] On the other hand, the AUs in Table 9 are apparently analytic, as they seem to contain a noninflected particle *bi/by* accompanied by one of the clitic BE-auxiliaries familiar from the L-preterit (see Table 2).[15,16]

Polish occurs in Table 9 since its characteristic agreement markers are, at least diachronically, reduced BE-auxiliaries (see §2.1.4). The same applies to the variants of Kashubian (see §2.1.3).[17]

---

[14]"BCMS-A" and "Czech-A" stand for the standard (written) varieties of these languages. As to Čakavian, see Panzer (1967: 35), Nehring (2002: 248–249), and Lisac (2009: 17–27). Note that the Čakavian forms display analogy-based present-tense endings (1SG -*n* is the regular reflex of -*m*; some dialects feature a 3PL *biju*).

[15]As to colloquial/dialectal Czech ("Czech-B"), see Toman (1980: 310) and Franks & King (2000: 92). The writing of the Czech-B AUs in one word is Toman's.

[16]"Macedonian⁺" marks the special case when speakers use *bi* plus a present-tense form of *sum* 'be' to disambiguate or emphasize the grammatical person (see Kramer 1986: 110–111). Else-

Hagen Pitsch

Table 8: L-conditional with inflected synthetic AU

| | | SG | | PL | |
		AU	V_L	AU	V_L
BCMS-A	1	*bih*	*pisala*	*bismo*	*pisale*
'write'	2	*bi*	*pisala*	*biste*	*pisale*
	3	*bi*	*pisala*	*bi*	*pisale*
Čakavian	1	*bin*	*bila*	*bimo*	*bili*
'be'	2	*biš*	*bila*	*bite*	*bili*
	3	*bi*	*bila*	*bi*	*bili*
Bulgarian	1	*bix*	*čela*	*bixme*	*čeli*
'read'	2	*bi*	*čela*	*bixte*	*čeli*
	3	*bi*	*čela*	*bixa*	*čeli*
Czech-A	1	*bych*	*udělala*	*bychom*	*udělaly*
'make'	2	*bys*	*udělala*	*byste*	*udělaly*
	3	*by*	*udělala*	*by*	*udělaly*
Upper Sorbian	1	*bych*	*dźěłała*	*bychmy*	*dźěłali*
'work'	2	*by*	*dźěłała*	*byšće*	*dźěłali*
	3	*by*	*dźěłała*	*bychu*	*dźěłali*

Table 10 shows those languages or varieties that have a noninflected AU.[18]

2.3 Garde's (1964) observation

In his paper on the Slavic conditional, Garde (1964: 88) makes an interesting note: Only Polish and the East Slavic languages have a particle in the conditional, and it is only these languages that can use more than only V_L in conditional clauses. While Garde does not provide any evidence supporting his former claim, the

---

where, *bi* alone is used (see Table 10).

[17]Stone (1993: 778) refers to Breza & Treder (1981: 134) when stating that *bë* "may or may not acquire a personal ending". Duličenko (2005: 392–393) adds that the "inflected" variants of the AU (*bë-m*, *bë-s*, etc.), which I dub "Kashubian-A1", are influenced by Polish, and that the "Kashubian-A2" AU-type is an archaism. Given this, the "Kashubian-B" variants in Table 10 are the modern standard.

[18]"BCMS-B" stands for colloquial/dialectal varieties (see Panzer 1967: 39; Kramer 1986: 105; Browne 2004: 253; Xrakovskij 2009: 276). As to Burgenland Croatian, see Tornow (2002: 240).

Table 9: L-conditional with inflected analytic AU

		SG		PL	
		AU	V$_L$	AU	V$_L$
Czech-B	1	*bysem*	*udělala*	*bysme*	*udělaly*
'make'	2	*bysi*	*udělala*	*byste*	*udělaly*
	3	*by*	*udělala*	*by*	*udělaly*
Slovak	1	*by som*	*volala*	*by sme*	*volali*
'call'	2	*by si*	*volala*	*by ste*	*volali*
	3	*by*	*volala*	*by*	*volali*
Polish	1	*by-m*	*prosiła*	*by-śmy*	*prosiły*
'ask'	2	*by-ś*	*prosiła*	*by-ście*	*prosiły*
	3	*by*	*prosiła*	*by*	*prosiły*
Kashubian-A1	1	*bë-m*	*miała*	*bë-smë*	*miałë*
'have'	2	*bë-s*	*miała*	*bë-sta*	*miałë*
	3	*bë*	*miała*	*bë*	*miałë*
Kashubian-A2	1	*bë jem*	*miała*	*bë jesmë*	*miałë*
'have'	2	*bë jes*	*miała*	*bë jesta*	*miałë*
	3	*bë je*	*miała*	*bë są*	*miałë*
Macedonian[+]	1	*bi sum*	*molela*	*bi sme*	*molele*
'ask'	2	*bi si*	*molela*	*bi ste*	*molele*
	3	*bi*	*molela*	*bi*	*molele*

latter one is valid and needs to be extended to Kashubian. The examples in (12)–(14) illustrate some alternative verb forms in the conditional periphrasis of the relevant languages.

(12) a. ..., (że-)by   przeczyta-ć książkę.
            that-COND read-INF   book.ACC
       '..., (in order) to read the book.'

    b. Włączo-no   by   radio.
       turn.on-IMPS COND radio.ACC
       'One would switch on the radio.'   (Polish; Migdalski 2006: 253)

Hagen Pitsch

Table 10: ʟ-conditional with noninflected AU

	AU		SG	PL
BCMS-B	1–3	*bi*	*pisala*	*pisale*
Burgenland Croatian	1–3	*bi*	*gledala*	*gledale*
Slovene	1–3	*bi*	*pohvalila*	*pohvalile*
Macedonian	1–3	*bi*	*molela*	*molele*
Kashubian-B	1–3	*b(ë)*	*miała*	*miałë*
Lower Sorbian	1–3	*by*	*słyšała*	*słyšali*
Belarusian	1–3	*b(y)*	*čytala*	*čytali*
Russian	1–3	*b(y)*	*skazala*	*skazali*
Ukrainan	1–3	*b(y)*	*bula*	*buly*

(13) …, że-bë   mie-c   jednã klasã   wëżi.
      that-COND have-INF one class.ACC more
  '…, (in order) to have one more class.'
          (Kashubian; www.odroda.kaszubia.com/01-07/edukacja.htm)

(14) a. Pospa-t' by!   (Russian)
      sleep-INF COND
      'If I could only sleep a little while!'   (Isačenko 1962: 346)

    b. …, čto-by   spa-t'.
       that-COND sleep-INF
      '…, (in order) to sleep.'

    c. Ne skaž-i (by) on mne   ėtogo   vo-vremja, …
      NEG say-IMP COND he me.DAT this.ACC in.time
      'If he had not told me that in time, …'   (Panzer 1967: 22)

In addition to infinitives and imperatives, Russian combines *by* with the present tense, participles, adverbs, and even nominals (see Issatchenko 1940: 195; Panzer 1967: 21–23).

As indicated by round brackets in Table 10, East Slavic and Kashubian (see Panzer 1967: 26) exhibit a reduced particle variant *b*. The same holds for Polish,

albeit in colloquial (presumably dialectal) contexts; see (15).[19],[20]

(15) a. Prosi   mnie   raz, że-b   ja z   nim   nad   rzekę
         ask.3SG me.ACC once that-COND I   with him.INS above river.ACC
         poszed-ł.
         go-L.SG.M
         'Once he asks me to go to the river with him.'

                                              (H. Auderska: *Babie lato*, 1974)

      b. [D]o końca   walczyliśmy, że-b   awansować do Ligi
         to   end.GEN fight.L.1PL   that-COND ascend.INF to league.GEN
         Mistrzów.
         champion.GEN.PL
         'We fought to the end to ascend to the Champions League.'

                                         (W. Batko: *Dramat pod Akropolem*, 2005)

      c. Wróciwszy   wczoraj z   zakupów   usiadła ja na
         having.returned yesterday from shopping.GEN.PL sit.L.SG.F I   on
         kanapie z   kubkiem melisy   w ręku,   że-b   się
         sofa.LOC with cup.INS melissa.GEN in hand.LOC that-COND REFL
         uspokoić.
         calm.down.INF
         'When I returned from shopping yesterday, I sat down on the sofa
         with a cup of melissa tea in my hands to calm down.'

                                 (Polish; gazetaolsztynska.pl, 2021 [accessed 4/2022])

Crucially, it is precisely the languages (plus Kashubian) that Garde (1964) claims
to possess a conditional (inflexible) particle which allow the phonological reduc-
tion of that very particle. I wish to propose that Garde's (1964) intuition is per-
fectly right, and that there is a fundamental difference between PARTICLE LAN-
GUAGES – Kashubian, Polish, Belarusian, Russian, and Ukrainian, all of which
have inflexible (or even absent) mood/tense markers and allow infinitives in the
conditional periphrasis – and AUXILIARY LANGUAGES, which have auxiliary verbs
specified for person and number. The latter holds true for all remaining lan-
guages, even if they display a particle from a descriptive point of view (BCMS-B,
Burgenland Croatian, Slovene, Macedonian, and Lower Sorbian).

   There is another phenomenon to be taken into consideration: AU-doubling in
the L-conditional.

---

[19]Examples (15a) and (15b) are taken from the National Corpus of Polish (http://nkjp.pl/).

[20]The colloquial character of (15a) also manifests in the absence of the appropriate agreement
marker -*m*. Only thanks to its absence can the particle undergo phonological reduction.

Hagen Pitsch

2.4 AU-doubling in the ʟ-conditional

In a number of Slavic languages, the conditional AU can occur twice in the same clause. While this is well-documented for Russian and older stages of Polish and Polish in early acquisition, there is only little data available on the remaining languages and varieties. This is likely to be due to the fact that AU-doubling is a phenomenon characteristic of substandard speech and considered incorrect by most grammars.

For Russian, Xrakovskij (2009) mentions the examples in (16).[21]

(16) a. Ja by   pogulja-l   by   segodnja večerom.
          I COND take.a.walk-L.SG.M COND today   evening.INS
          'I would like to take a walk tonight'

      b. Čto-by   ja tebja   by   zdes' bol'še ne vide-l.
          that.COND I you.ACC COND here more   NEG see-L.SG.M
          'So that I would not see you here again.'

                                          (Russian; Xrakovskij 2009: 277)

Rittel (1973) gives the examples in (17) and (18) from Kashubian and the Masovian dialect of Polish, respectively.

(17) jag by   úna by   odeš-ŭ-a
      how COND she COND walk.away-L-SG.F
      'as though she should have gone away'   (Kashubian)

(18) že-by   učy-ł-by   s'e   xoźić
      that.COND learn-L.SG.M-COND REFL walk.INF
      'in order for him to learn to walk'   (Masovian; Rittel 1973: 146)

By-doubling is also found in colloquial Polish as shown in (19).

(19) a. … to dziś by-m   by-ł-by   szejkiem!
              then today COND-1SG be-L.SG.M-COND sheikh.INS
          '… then today I would be a sheikh'   (wykop.pl, accessed 4/2022)

      b. nie sądzę,   by-śmy by-l-i-by   tak blisko siebie   i   tak
          NEG think.1SG COND-1PL be-L-PL.M-COND so close REFL.ACC and so
          związani   jak my, gdyby nie ten czas
          connected.PL.M as we if   NEG this time
          'I don't think we would be as close to each other and as connected as
          we are if it were not for this time'

---

[21]Hansen (2010: 331) notes that a random sample taken from the National Corpus of Russian indicates that *by*-doubling occurs "quite frequently" in Russian despite its being not accepted by the norms of the standard language.

(coll. Polish; pl.spiceend.com, accessed 4/2022)

Smoczyńska (1985: 624) notes that children acquiring Polish as their first language quite regularly double the conditional AU; see (20).

(20) A   moja mamusia też by   mia-ł-a-by
      and my   mum   also COND.3SG have-L-SG.F-COND.3SG
      ładne włoski.
      pretty hair.ACC.PL
      'My mum would also like to have pretty hair.'
                                          (Polish; from Błaszczyk 2018: 119)

Especially in subjunctive clauses, the phenomenon has also been observed in Lithuanian-Polish bilinguals; see (21).

(21) a. Teraz to-by   na pewno zaintersowani by   by-l-i.
          now   PART-COND for sure interested.PL.M COND be-L-PL.M
          'Now they would certainly be interested.'

      b. …, że-by   my nie widzie-l-i-b   co   oni gadają.
              that-COND we NEG see-L-PL.M-COND what.ACC they chatter.3PL
          '…, so that we do not know what they are chattering.'
                      (Polish; Smułkowa 1999: 58; from Błaszczyk 2018: 132)

According to Błaszczyk (2018: 132), many similar examples can be found in Grek-Pabisowa & Maryniakowa (1999), who describe the linguistic peculiarities of the dialects of the former Polish Eastern Borderlands. Zielińska (2002) does not regard such examples as the result of interference/contact but as local variants. She adds that they might well be considered archaisms, as doubling already occurs in Old Polish as documented in (22).

(22) iże-by   by   by-ł-y   wysłuchany twoje prośby
      that-COND COND be-L-PL.F heard.PL.F   your pleas.ACC.PL
      'so that your pleas might be heard' (Old Polish;*Historia Aleksandra*, 1510)
                                  (Rittel 1975: 113; from Błaszczyk 2018: 133)

There must clearly be more research as to the extent of AU-doubling in Slavic but the data allow for the following generalizations: First, AU-doubling is not a recent phenomenon. Second, it seems to be restricted to colloquial and dialectal varieties as well as speech produced in the course of early language acquisition.

Hagen Pitsch

Third, it seems to prevail in East Slavic (first of all Russian), Polish and Kashubian.[22] Potentially, AU-doubling might turn out to be another piece of evidence for the special status of the languages and varieties of the "North-Eastern group" as regards their conditional AU. In §4.4, I sketch a syntactic analysis to account for the phenomenon.

2.5 Putting the pieces together

Bringing together the pieces of information provided thus far – (i) the absence or presence of person/number agreement in the ʟ-preterit and the ʟ-conditional, (ii) the non-/availability of other forms than V_L in the conditional, and (iii) the possibility of particle reduction in those languages that (seem to) have one – gives us the overall picture in Table 11.[23]

The overall picture reveals a number of facts:

First, the variation in agreement in the ʟ-preterit is not coextensive with the one in the ʟ-conditional: Whilst in the preterit, only Kashubian and East Slavic do not express person/number agreement, this holds for far more languages in the conditional. Thus, it seems that the conditional AU is more prone to linguistic change than the AU in the ʟ-preterit.[24]

Second, there is no obvious correlation between the absence of person/number agreement in the ʟ-conditional and the availability of verb forms other than V_L.

Third, there is a robust correlation between the availability of verb forms other than V_L in the conditional and the possibility of phonologically reducing the AU.

The availability of verb forms other than V_L as well as of particle reduction clearly distinguish Kashubian, Polish, and the East Slavic languages. Crucially, in all of them diachronic change lead to the loss or reshaping of the present-tense paradigm of the (former) ʙᴇ-auxiliary (see §§2.1.2–2.1.4).

---

[22]Luka Szucsich (p.c.) reports that *bi*-doubling seems to be possible in Burgenland Croatian. The question calls for further (corpus-based) research.

[23]I omit AU-doubling. According to the data, it is possible in the same languages that allow for particle reduction. From the varieties of Kashubian, I list only Kashubian-B, as it seems to represent present-day Kashubian (regarding the conditional). "∘" signifies the possible lack of agreement in the Macedonian conditional in unmarked contexts (see footnote 16).

[24]Tentatively, this is due to the more "regular" shape of the conditional AU with the stem *bë/bi/by-* throughout its whole paradigm. By contrast, the preterit AU lacks a similarly consistent base. Reducing the conditional AU to its stem by dropping the agreement ending (and thus boiling it down to its essential grammatical meaning) seems thus more natural than in the case of the preterit AU (which can at best be dropped altogether).

Table 11: ʟ-periphrases in comparison

	agreement on AU in the ...		more	particle
	ʟ-preterit	ʟ-conditional	than V_L	reduction
BCMS-A	•	•		
BCMS-B	•			
Burgenland Croatian	•			
Čakavian	•	•		
Slovene	•			
Bulgarian	•	•		
Macedonian	•	○		
Polish	•	•	•	•
Kashubian(-B)			•	•
Czech	•	•		
Slovak	•	•		
Lower Sorbian	•			
Upper Sorbian	•	•		
Belarusian			•	•
Russian			•	•
Ukrainian			•	•

In what follows, I will sketch a number of scenarios of language change to explain the present-day situation in the Slavic languages. In doing so, I will identify four groups of languages with a distinct development each.

3 Linguistic change

The modern shape of the ʟ-preterit and the ʟ-conditional in Slavic allows reflections about what happened to the relevant periphrases in preceding centuries and has thus given rise to the current state of affairs. Four distinct diachronic scenarios emerge.

3.1 "Old symmetry"

The first scenario concerns the following South and West Slavic languages:

Hagen Pitsch

- BCMS-A
- Bulgarian
- Čakavian
- Macedonian[+]

- Czech
- Slovak
- Upper Sorbian

All of them retain the Late Proto-Slavic shape of the L-preterit and L-conditional, especially of the relevant AUs, i.e., they use inflected auxiliary verbs expressing person/number agreement. What is more, they do not allow verb forms other than V_L in the conditional, and they do not reduce their conditional AU.

No particular language change took place in these languages apart from occasional replacement of the old aorist inflections on the conditional AU with present-tense markers. In this respect, it is possible to discern two subgroups:

1. Čakavian has replaced the aorist inflections with present-tense suffixes and thus retains synthetic auxiliary verbs (e.g., 1SG *bi-n* [< *bi-m*], 2SG *bi-š*, etc.).

2. Czech-B, Macedonian[+], and Slovak have substituted the old aorist markers with the present-tense forms of their respective BE-auxiliary (e.g., Macedonian[+] *bi sum*, *bi si*, etc.; Slovak *by som*, *by si*, etc.). Minor analogies of the same type took place in BCMS-A (1PL *bi-smo*) and Czech (2SG *by-s*). I claim that these new "analytic" auxiliary forms are really (still) synthetic, i.e. that the BE-forms substituting the old aorist inflections are suffixes, not clitics. They have been carried over from the L-preterit by analogy but changed their morphosyntactic status. Thus, for instance, BCMS-A 1PL *bi-smo* and Slovak *by sme* are clearly parallel formations – irrespective of orthographic conventions.

3.2 "Asymmetry"

The second scenario concerns the following languages:

- BCMS-B
- Burgendland Croatian
- Macedonian
- Slovene

- Lower Sorbian

There is an asymmetry in that these languages feature an inflected auxiliary verb in the L-preterit but a noninflected AU (*bi* or *by*) in the L-conditional. But like the

varieties described in §3.1, they exclude any verb forms other than V_L from the conditional and lack reduced variants of their conditional AU.

A straightforward way to explain these facts goes as follows: The conditional AU is merely a "pseudo-particle", i.e. we are actually (still) dealing with an inflected auxiliary verb. This verb, however, has dropped its agreement marking at the surface, which means that it is underspecified for person and number. In other words, /bi/ should be analyzed as being associated with person/number agreement features as sketched in (23).[25]

(23)   /bi/$_{[\alpha\text{PERSON},\beta\text{NUMBER}]}$

If this is on the right track, the languages and varieties in question form a larger class with the ones addressed in §3.1, the reason being that both groups retain – even if covertly – synthetic auxiliary verbs that encode person/number agreement.

Possible causes for the loss of overt agreement are phonological reduction (drop) of inflectional endings or/and paradigm leveling (intraparadigmatic analogy). Both mechanisms seem to have been involved, for instance, in the development from BCMS-A to BCMS-B; see Table 12.

Occasionally, language contact is identified as another possible source of overt agreement loss. Thus, for instance, Panzer (1967: 24) suggests that Lower Sorbian dropped the person/number suffixes on its conditional AU due to the increased use of personal subject pronouns (induced by language contact with German). A similar explanation is put forward by Rittel (1970: 100) to derive the present-day state of the L-periphrases in Kashubian-B.

Evidence in favor of analyzing the conditional AUs in question as underspecified auxiliary verbs comes from Macedonian: In cases where speakers need to disambiguate the person feature ("Macedonian⁺"; see footnote 16), *bi* co-occurs with what looks like clitic BE-forms as used in the L-preterit, hence 1SG *bi sum*, 2SG *bi si*, etc. in place of solitary *bi*. I wish to claim that these elements do not differ from, e.g., Slovak *by som*, *by si*, etc. (see §3.1) – i.e. they are suffixes. However, different from Slovak, the Macedonian suffixes can be left unpronounced when there is no need to express the person feature on the AU. Thus, when Macedonian *bi* occurs without person/number agreement, it resembles (23). Incidentally, it does not seem too bold a claim that the step from Macedonian⁺ to Macedonian represents phonological drop (Table 12) and, thus, linguistic change in progress.

---

[25] The agreement features might also be located in a silent agreement suffix attached to the stem.

Table 12: Loss of overt agreement encoding in BCMS

		SG	PL
BCMS-A	1	*bih*	*bismo*
	2	*bi*	*biste*
	3	*bi*	*bi*

↓ phonological drop (1SG) ↓

		SG	PL
	1	*bih*	*bismo*
	2	*bi*	*biste*
	3	*bi*	*bi*

↓ paradigm leveling (1/2PL) ↓

		SG	PL
BCMS-B	1	*bi*	*bismo*
	2	*bi*	*biste*
	3	*bi*	*bi*

3.3 "New symmetry"

The third scenario concerns the following languages:

- Belarusian
- Russian
- Ukrainian
- Kashubian-B

As said in §2.1.2 and §2.1.3, respectively, present-day East Slavic and Kashubian-B lack AUs in the L-preterit and the L-conditional for individual diachronic reasons. Apart from that, they employ verb forms other than only V_L in the conditional, and they also allow the reduction of their conditional AU (*by/bë → b*).[26]

For Old East Slavic, historical grammars commonly note the significant effect the changes sketched in §2.1.2 had on the East Slavic verbal system. Thus, for instance, Issatchenko (1940: 193) writes that "[t]his change, which at first affected only the verb *byti*, shook the whole verbal system." In the same vein, Ivanov (1964: 395) states that the essence of the relevant changes consisted in the loss of

---

[26]Whereas the two variants are in complementary phonological distribution in Belarusian and Ukrainian, their choice depends primarily on stylistic factors in Kashubian-B and Russian.

(agreement on) the former AU, which in turn caused a shift of the "center of the tense/mood form" to V_L.

What the authors refer to is a shift in agreement marking and finiteness: While before the changes, Old East Slavic L-periphrases uniformly contained a finite auxiliary verb and a nonfinite L-participle, the changes turned the former into a particle encoding tense/mood (but not agreement), and the latter into a form associated with a complete set of agreement features. Initially, the change affected only the L-preterit, effectively deleting the auxiliary due to the loss of the present-tense paradigm of *byti*. As a consequence, speakers now recognized V_L as the only (finite) verb, associating it with a "hidden" (underspecified) person feature (Junghanns 1995: 88); see (8b) in §2.1.2.

Only after the L-preterit had thus turned into a synthetic form, the change spread to the L-conditional: By analogy, speakers now also perceived V_L in the L-conditional as finite. As a clause can only contain one finite verb, a finite auxiliary became redundant. This paved the way for dropping person/number agreement on the conditional AU, which thus turned into a mere mood particle. This chain of events is schematized in Table 13, using the 1SG of *čitati* 'read' as an illustration.

Table 13: Diachronic change in East Slavic

	I			II	
L-preterit:	*jesmĭ* finite	*čitala* nonfinite	→	∅ PART	*čitala* finite
			↓		
L-conditional:	*bychŭ* finite	*čitala* nonfinite	→	*by* PART	*čitala* finite

I suggest that by and large the same took place in Kashubian – though at a later time –, giving rise to the situation in present-day Kashubian-B. In §4, I put forward a syntactic account to explain the availability of verb forms other than V_L in the conditional. This account builds upon the presence of a particle in the relevant languages, i.e., of a tense/mood operator in the functional domain of the clause. With one important addition, the analysis also captures present-day Polish, which I turn to in the following section.

Hagen Pitsch

3.4 "Demolition and reconstruction"

As outlined in §2.1.4, the present-tense forms of Polish *być* 'be', inherited from Late Proto-Slavic, completely vanished due to their reduction to atonic forms and concomitant repurposing as person/number markers, which compensated for the lost auxiliaries in the ʟ-preterit. These very markers have subsequently also been used to form an utterly new present-tense paradigm for the copula *być* (*jest-em*, *jest-eś*, etc.). Finally, they also occur on the Modern Polish conditional AU as shown in (24).

(24) a. Ja by-m   pisa-ł-a.   (Polish)
        I COND-1SG write-L-SG.F
        'I would be writing.'

     b. … że=by-m   ja pisa-ł-a.
        that=COND-1SG I write-L-SG.F
        '… that I would be writing.'

However, different from the ʟ-preterit, the atonic agreement markers are syntactically immobile once they show up on conditional *by*. This raises the question if the members of the paradigm of the Polish conditional AU are not simply synthetic forms with agreement endings that merely "imitate" the atonic markers from the ʟ-preterit by analogy. In §4, I will argue against this view and claim that the monolithic nature of *by-m*, *by-ś*, etc. is due to the fact that the Polish atonic markers are generated in the specifier of the functional head I^0 (occupied by *by*) and subsequently "m-merge" (Matushansky 2006, Pietraszko 2021) with it. It follows that, ultimately, both form a single and inseparable unit.

Additional evidence for treating Polish *by* as a particle that is initially separate from the atonic agreement markers comes from language acquisition (a.o., Smoczyńska 1985; Błaszczyk 2018; Dogil & Aguado 1989); see (25).

(25) a. pisał-em-by   (Polish)
        write.L.M.SG-1SG-COND

     b. Ja by   pisał-em
        I COND write.L.M.SG-1SG
        'I would be writing' (Smoczyńska 1985: 640; from Błaszczyk 2018: 118)

The data show that children frequently "mix up" the canonical positions of *by* and the agreement markers, respectively. Apparently, they do so by analogy with the ʟ-preterit, where the latter mostly attach directly to V_L. However, in §4, I will try to show that Embick (1995) is right in claiming that the direct attachment of

the agreement markers to V_L is an illusion. Underlyingly, the L-preterit involves a silent past-tense operator in I^0, and it is this operator which the agreement marker adjoins to.

My scenario for Polish is thus the following: The demolition of the inherited present-tense BE-paradigm led to a situation where Old Polish was very close to East Slavic and Kashubian-B (§3.3): It had effectively lost the inflected auxiliary verb in the L-preterit and would at some later point in time face the same situation in the conditional. But unlike East Slavic and Kashubian-B, Polish did not entirely dispose of the old BE-forms but re-utilized them as agreement markers. Combining insights of Embick (1995), Matushansky (2006), and Pietraszko (2018, 2021), I argue that these markers are clitic heads generated in SpecTP from where they adjoin to I^0, which is silent in the L-preterit (realis mood, past tense) but overt (*by*) in the L-conditional (irrealis mood).

4 Towards a syntactic analysis

Based on the preceding observations, I wish to argue that there are two major classes of Slavic languages with regard to L-periphrases. The difference between them concerns the category of their AUs.

4.1 The framework

With modifications, I rely on the framework developed in Pietraszko (2018, 2021) who argues that, in periphrases, T^0 (= I^0) has an uninterpretable (i.e. selectional) feature [uV] which cannot be checked against the interpretable (categorial) feature [iV] of V^0 due to an intervening functional projection, namely AspP.[27] As a consequence, an auxiliary verb (Aux) with its own [iV] is generated in the specifier of I^0 where it satisfies the selectional requirement; see Figure 1.[28]

Unlike Pietraszko, I claim that the crucial (type of) feature in Slavic L-periphrases is not [V] but rather [φ], i.e. verb-subject agreement. This modification is motivated by the fact that, no matter whether or not AspP is assumed in the syntax of Slavic languages, a constellation like Figure 1 is unlikely to arise: If AspP is projected, it is so in general, hence each and every Slavic clause should be periphrastic. On the other hand, if AspP is not assumed (because viewpoint aspect is taken to be a lexical rather than a grammatical category), it should again be

---

[27]According to Pietraszko, this constellation underlies, e.g., English progressive tenses.

[28]Circle-ended lines mark Agree relations, checked features are struck out. Pietraszko uses the framework of Bare Phrase Structure, so in her tree the auxiliary is generated next to I^0, which equals SpecIP under X-bar assumptions.

Hagen Pitsch
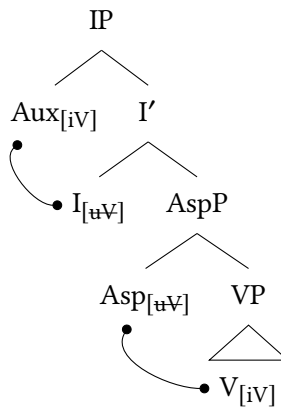
Figure 1: Configuration giving rise to periphrasis (see Pietraszko 2021: 11)

absent in general, which eliminates Pietraszko's (2021) structural motivation for periphrasis.

Verb-subject agreement is a more plausible candidate: If the verb in V^0 comes from the lexicon equipped with a complete set of φ-features (φ^+), there is no need to project any auxiliary, which gives rise to a synthetic structure. On the other hand, if V^0 is occupied by a verb with an incomplete set of φ-features (φ^-), the missing features have to be supplied by an auxiliary. Crucially, for a φ-set to be incomplete, one of the following conditions has to be complied: Either the set lacks a person feature (participles) or it is completely empty (infinitives).[29]

Depending on the class a language belongs to, it either has or has not available "true" auxiliary verbs (in Aux^0 or/and I^0) that come with φ^+. If it has, I^0 owns or receives (via percolation; see Pietraszko 2018) φ^+ and can thus enter into an Agree relation with the subject. If it has not, one of two scenarios are possible: In Polish and Kashubian-A1, φ^+ is generated in SpecIP in the form of an atonic agreement marker and subsequently fused (via m-merger; see Matushansky 2006) with I^0. On the other hand, in East Slavic and Kashubian-B, V_L comes from the lexicon with a complete set of φ-features (see §2.1.2), so I^0 can establish an Agree relation with the subject without the intervention of an auxiliary or agreement marker.

What the "North-Eastern group" of Slavic languages have in common is that I^0 is a mere particle (Garde 1964), which is due to the diachronic reduction or loss,

---

[29]See Pitsch (2015) for a formal account of the finite/nonfinite distinction in Slavic resting on a prominent role of grammatical person.

respectively, of the present-tense paradigm of 'be'. All remaining languages retain "true" auxiliary verbs. I address both these classes in the following sections.

4.2 Auxiliary languages

The first class is constituted by the languages discussed in §§3.1–3.2, i.e. BCMS (both varieties), Bulgarian, Burgenland Croatian, Čakavian, Macedonian (both varieties), Slovene, Czech, Lower Sorbian, Slovak, and Upper Sorbian. All retain auxiliary verbs specified for person and number, hence φ^+. On the other hand, the participle in V^0 only specifies number and possibly also gender, hence φ^-.

Crucially, I claim that it is the verbiness of auxiliaries that allows them to select V_L in V^0, which is therefore the only verb form available in L-periphrases.

According to Pietraszko (2018, 2021), verbs carry [iV], while I^0 has [uV], which is checked against the closest [iV] (see Svenonius 1994; Chomsky 1995; Julien 2002; Adger 2003; Cowper 2010). Additionally, I argue that auxiliary verbs carry both [uV] and [iV], so they select (a verb in V^0) and are selected (by I^0) at the same time.[30] In a subset of periphrases, said auxiliary verbs are generated as the head of an AuxP between IP and VP as shown in (26).

(26) IP > AuxP > VP

This spine underlies, for instance, the L-preterit in BCMS with the 3SG *je* 'is' (see Tomić 1996: 838) as well as the BCMS and Polish L-future (see Browne 1993: 331; Migdalski 2006: 275). The auxiliary in Aux^0 selects V_L in V^0 and adds a person feature (φ^+). By contrast, V_L is φ-incomplete (φ^-). Following Pietraszko (2018), the φ-probe undergoes feature percolation under V-checking, i.e., from V^0 (number/gender) and Aux^0 (adding person) to I^0. Only in its percolated position does the probe become active and enters in an Agree relation with a subject; see (27a) and illustrations from BCMS in (27b) and (27c) = (1c).[31]

(27) a.

---
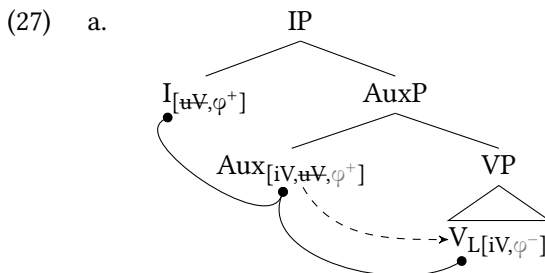
[30]The feature [uV] of the auxiliary merely requires a verbal category in its complement domain. In addition, the auxiliary comes with a feature requiring that this verb be a V_L.

[31]The dashed arrow indicates the selection of V_L by Aux^0. Inactive φ-features are gray.
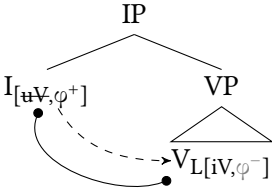
Hagen Pitsch

    b. [$_{IP}$ ∅ [$_{AuxP}$ Ivana je   [$_{VP}$ govori-l-a   ]]]
             PST   I.   AUX-3SG   speak-L-SG.F
        'Ivana (has) spoke(n)'

    c. Kad   [$_{IP}$ ∅ [$_{AuxP}$ *pro* bude-mo [$_{VP}$ govori-l-i   …]]]
        when   FUT   1PL AUX-1PL   speak-L-PL.M
        'When we will speak …'   (BCMS)

It is crucial that the auxiliary in Aux^0 selects (thanks to its verbiness) V$_L$ in V^0. As a consequence, any other verb form in V^0 is excluded.[32]

    However, besides Aux^0, auxiliary verbs may also reside in I^0. According to Tomić (1996: 838), this holds for so-called weak pronouns in BCMS (all except 3SG *je*, i.e. *sam*, *si*, etc.). Migdalski (2006: 275) makes a similar claim for Polish (see §4.3.3). By and large the same is likely to be true for Bulgarian and Macedonian. In Czech and Slovak, the placement of the negation *ne* relative to the forms of the BE-auxiliary and V$_L$ provides evidence that auxiliaries are generally merged in I^0. By contrast, the full-verb (copular) forms of *být/byť* 'be' (which also figure in the participial passive) are best analyzed as being generated in Aux^0, whereas ordinary full verbs – including V$_L$ – are in V^0.

    The structure with auxiliaries generated directly in I^0 is shown in (28a), with a Czech illustration in (28b).[33] A complete set of φ-features is present in I^0 since it is occupied by the auxiliary (here: *jsem*). Feature percolation is thus confined to a possible gender feature on V^0 and may in fact rather amount to an Agree relation between the two φ-sets in V^0 and I^0. Quite like in (27a), the auxiliary selects V$_L$ in V^0.

(28)   a.   IP

    b. [$_{IP}$ Já jsem   [$_{VP}$ ⟨já⟩ pracova-l-a]].
        1SG PST.1SG   work-L-SG.F
        'It is me who (has) worked.'   (Czech)

---

[32]The Polish L-future may also contain an infinitive in V^0. Arguably, the *będ*-auxiliary has a (more) flexible selectional frame.

[33]Possibly, the subject pronoun *já* in (28b) does not merely go to SpecIP but adjoins to IP to be interpreted as contrastive or verum focus (see Junghanns & Zybatow 2009).

One way or the other, auxiliary languages have verbal auxiliaries with a complete φ-set that select V_L in the main verb slot, which is why other verb forms (like the infinitive) are unavailable in this position. It is therefore that "impersonal" conditionals/subjunctives are not attested.

4.3 Particle languages

The second class is constituted by the "North-Eastern group", i.e. Kashubian, Polish as well as Belarusian, Russian, and Ukrainian. These languages have a particle both in the L-preterit and in the L-conditional. This situation is the result of the diachronic reshaping or loss, respectively, of the present-tense paradigm of 'be'.

Using Tomić's (2000) terminology, the relevant particles are OPERATORS, as they are in a high functional position – I^0 – from where they supply the proposition as a whole with their tense/mood semantics. They have developed from former auxiliaries which lost their "verbal character" (Issatchenko 1940). In other words, they do not specify person/number agreement anymore and may even be silent in some cases (as the East Slavic L-preterit).

4.3.1 East Slavic and Kashubian-B

For the East Slavic languages and Kashubian-B, diachronic changes had at least two crucial consequences:

1. Since there was no other way left to encode agreement, V_L, hitherto a participle specified only for number and gender, was reinterpreted as a fully-fledged (finite) form (see Tseng 2009: 757), i.e., it was additionally associated with an underspecified person feature (Junghanns 1995: 174; see §2.1.2). In other words, V_L enters the syntactic derivation equipped with a complete set of φ-features (φ^+).

2. Not being a verbal category, the particle in I^0 fails to select a specific form in V^0. As a consequence, V_L is not the only choice, at least in the conditional/subjunctive.[34] On feature checking, φ^+ percolates from V^0 to I^0, allowing the latter to establish an Agree relation with the subject.

---

[34]The past tense always and exclusively contains V_L. I suspect that this is due to the fact that no other verb form could possibly reflect the presence of the silent past-tense operator in I^0 (note that the languages in question have long-since lost past-tense aorist and imperfect forms).

Hagen Pitsch

The corresponding syntactic structure with V^0 being occupied by a V_L is given in (29a). Two Ukrainian examples are shown in (29b) (past tense) and (29c) (conditional), repeated from (2) and (3), respectively.[35] In the glosses, I indicate that V_L is equipped with an implicit person feature matching the subject.

(29) a.

```
            IP
          /    \
  I[u̶V̶,φ⁺]     VP
    •         /\
     \       /  \
      \   V[iV,φ⁺]
       \____•
```

    b. Koly [IP ty Ø [VP ⟨t̶y̶⟩ narody-l-a-s'   ]]?
        when   2SG PST   give-birth-L-[2]SG.F-REFL
        'When were you born?'

    c. [IP Ja b [NegP c'oho ne [VP ⟨j̶a̶⟩ skaza-v ⟨c̶'̶o̶h̶o̶⟩]]].
        1SG COND   this.GEN NEG   say-L.[1]SG.M
        'I would not have said that.'   (Ukrainian)

But V^0 can also be occupied by an infinitive. As infinitives lack φ-features, there is nothing to percolate to I^0, thus the only possible subject is φ-less PRO. The resulting syntactic structure in (30a) is what we find in irrealis conditionals (subjunctives) like (30b).[36],[37]

(30) a.

```
            IP
          /    \
   I[u̶V̶]       VP
    •         /\
     \       /  \
      \  V_INF[iV]
       \____•
```

---

[35]In (29b), I stay agnostic about the base and target positions of the wh-word. Arguably, in (29c), the subject pronoun moves further to adjoin to IP (contrastive focus), while the direct object *c'oho* has moved to SpecNegP.

[36]In §4.4.2, I will argue that *by* goes from I^0 to C^0 and fuses with it.

[37]Willis (2000) argues that Russian *by* is generated in C^0 as a result of grammaticalization. He claims that it was originally merged in I^0, from where it frequently moved to C^0 in Old East Slavic. Speakers then reanalyzed its derived position as underlying. I am hesitant to agree, mainly due to *by*-doubling (§4.4). A theoretical possibility is that there are two homophonous instances: by_I (conditional mood) and by_C (subjunctive clause type). The same might be true for Polish *by*, which can introduce subjunctive clauses even without a complementizer (an advocate for a by_C is Jędrzejowski 2020: 108). Still, I prefer a movement/copying analysis with only one *by* (like Migdalski 2006: 259).

b. [$_{CP}$ čto [$_{IP}$ by   [$_{VP}$ PRO rabota-t']]]
    that   COND   work-INF
  '(in order) to work'   (Russian)

4.3.2 Polish and Kashubian-A1

Polish and Kashubian-A1 possess atonic person/number agreement markers. I adopt the view that these markers are clitics. Embick (1995) proposes that they are generated as adjuncts to I^0. Unlike in East Slavic and Kashubian-B, this ensures that person/number agreement is encoded in a position distinct from V^0, so V_L is not in need of an underspecified person feature – it is a participle proper. In other words, Polish and Kashubian-A1 have substituted their former auxiliary verbs (once in Aux^0) with composite items in I^0. These items consist of a particle ($\emptyset$ in the past tense, *by/bë* in the conditional) plus an agreement marker; see Figure 2.

I

I   AGR
{$\emptyset$/*by*}
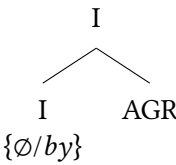
Figure 2: Analytic I^0 according to Embick (1995)

However, Embick's (1995) analysis has one crucial theoretical disadvantage: It assumes that two heads are base-generated as adjuncts to each other, which involves the danger of overgeneralization. To avoid this problem, I again follow Pietraszko (2018, 2021) who argues against the common claim that auxiliaries in periphrases are necessarily generated as heads within the clausal spine (i.e. Aux^0 or I^0). They can also be generated in SpecIP. To implement this alternative, Pietraszko adopts Matushansky's (2006) idea of "m-merger": A head is merged in the specifier of a functional head (here: I^0) and subsequently adjoins to that head to form an inseparable unit; see Figure 3a and Figure 3b, respectively.

This analysis eliminates the danger of overgeneralization inherent to Embick's (1995) approach as there is a clearly defined motivation for merging the agreement marker in SpecIP: It compensates for the missing person feature in V^0. At the same time, the analysis yields the same syntactic configuration as in Figure 2 – i.e. Figure 3b – and thus preserves its advantages.[38]

---

[38]See Abramowicz (2008) for a survey of the advantages of Embick's (1995) analysis as compared

Hagen Pitsch

(a) Base generation in SpecIP · · · · · · (b) Subsequent adjunction to I[0]
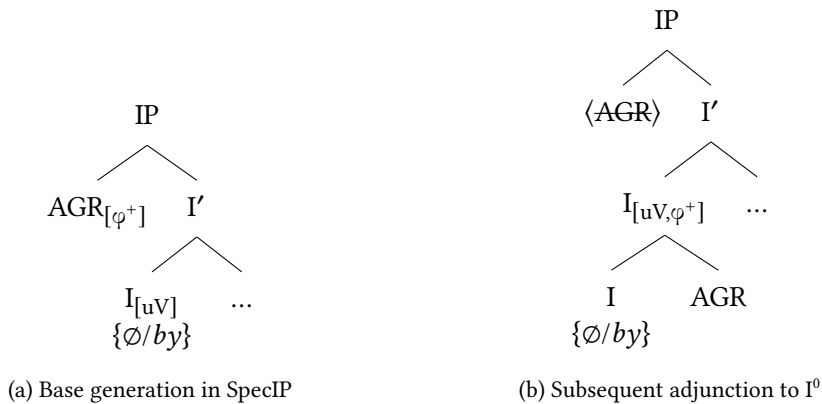
Figure 3: M-merger of an agreement marker (AGR)

As a result of the adjunction in Figure 3b, the particle in I[0] and the agreement marker fuse into a complex I[0] specified with [uV] and [φ^+]. This gives a constellation very much similar to (28a). Put differently, Polish and Kashubian-A1 ""reconstruct" an analytic auxiliary verb in I[0].

Another advantage of the view that the atonic agreement markers are syntactic heads is that they can also be absent. If there is no agreement marker generated in SpecIP, I[0] stays a mere tense/mood particle and the clause lacks φ-features. Like in East Slavic and Kashubian-B, this makes it possible to have an infinitive or impersonal *no/to*-form in V[0] and, consequently, a PRO subject as shown for Russian in (30a) and (30b).

4.3.3 Full verb 'be' in East Slavic and Polish

The present proposal should also be able to deal with the full verb (copular) forms of 'be' in Polish and East Slavic.

As to Polish, I agree with Migdalski (2006) that *jest* and *są* do not specify any person feature (also see Tomić 1996 on BCMS *je*). Like Migdalski (2006: 275), I analyze them as heading an AuxP; see (31).[39]

(31) I[0]   > Aux[0]   > V[0]   > XP
      MOOD/TENSE/AGR   {*jest/są*}   zero copula   predicate nominal

---

to alternative approaches. Note that Pietraszko's (2018) approach is not restricted to the X-bar framework but also works under Bare Phrase and Labeling Algorithm assumptions.

[39]In copular clauses, V[0] is silent but introduces a situation argument as well as argument slots for the predicate nominal and the subject. This silent head corresponds to Bowers's (1993) Pr(ed)[0] (see also Bailyn 2001, 2012; Markman 2008), Citko's (2008) π^0, or den Dikken's (2006) Rel[0].

Except for the 3rd person, *jest* (only in dialects also *są*) usually raises to I^0 – a silent present-tense operator – to adjoin to the agreement marker m-merged with that operator; see (32a). However, though nowadays rarely, *jest/są* can also stay *in situ*; see (32b). In the 3rd person, *jest/są* always stay *in situ*; see (33). Finally, Aux^0 may be absent as in (34).

(32) a. [$_\text{IP}$ ja [$_\text{I}$ jest+[$_\text{I}$ ∅+(e)m]] [$_\text{AuxP}$ ⟨j̶e̶s̶t̶⟩ [$_\text{VP}$ ⟨j̶a̶⟩ ∅$_\text{Cop}$ [$_\text{AP}$ głodny]]]]]
      b. [$_\text{IP}$ ja [$_\text{I}$ ∅+m] [$_\text{AuxP}$ jest [$_\text{VP}$ ⟨j̶a̶⟩ ∅$_\text{Cop}$ [$_\text{AP}$ głodny]]]]]
      'I am hungry'

(33) [$_\text{IP}$ Anna ∅$_\text{I}$ [$_\text{AuxP}$ jest [$_\text{VP}$ ⟨A̶n̶n̶a̶⟩ ∅$_\text{Cop}$ [$_\text{AP}$ głodna]]]]]
      'Anna is hungry'

(34) [$_\text{IP}$ ja [$_\text{I}$ ∅+m] [$_\text{VP}$ ⟨j̶a̶⟩ ∅$_\text{Cop}$ [$_\text{AP}$ głodny]]]
      'I am hungry'

The situation is different in East Slavic: I follow Issatchenko (1940) in that Russian *est´* has become a particle, and argue that this translates into a shift from Aux^0 to I^0. In other words, Belarusian *ësc´*, Russian *est´*, and Ukrainian *je*, respectively, are the overt variant of an otherwise silent I-head encoding the present tense. Their being overt nicely matches the fact that they are, unlike Polish *jest*, emphatic (verum or contrastive focus; see Geist 2007: 127); see (35).[40]

(35) [$_\text{IP}$ Anna est´$_\text{I}$ [$_\text{VP}$ ⟨A̶n̶n̶a̶⟩ ∅$_\text{Cop}$ [$_\text{AP}$ golodna]]]
      'Anna IS hungry'

4.4 The doubling issue

The analysis proposed in §4 covers all "standard" examples including those with reduced particles.[41] However, the phenomenon of particle doubling in conditional clauses described in §2.4 still calls for a syntactic explanation.

For the time being, the data suggest that said doubling is characteristic of the "North-Eastern group", i.e. Kashubian, Polish, and the East Slavic languages. Therefore, I suspect that there is a connection between particle doubling and the syntactic peculiarities of the relevant languages. More precisely, I suggest that it is the existence of a conditional particle that enables its reduplication.

---

[40]Taking into consideration that "usual" verbs are emphasized by means of contrastive intonation, the existence of an overt present-tense I^0 specifically for copular clauses is likely to be due to the lack of overt present-tense BE-forms in East Slavic.

[41]The reason particles can be reduced is that their final segments do not encode grammatical information and can thus be dropped on phonological grounds.

Hagen Pitsch

Any syntactic analysis designed to capture the doubling phenomenon has to ensure that there can be two instances of the conditional particle in the same clause. Furthermore, the second instance must be semantically vacuous, as doubling affects only the surface form, not meaning and interpretation (there is no doubling of the irrealis semantics in the sense of decreased probability, counterfactuality, or the like).

4.4.1 Multiple copies

The most straightforward way to achieve these goals is provided by the Copy Theory of Movement (see, a.o., Chomsky 1993, Nunes 1995, Corver & Nunes 2007). According to this theory, the syntactic trace left behind of a moved element ("α") is a copy of that very element; see (36).

(36) $[_{XP} \, \alpha \, [_{YP} \, \alpha \,]]$

As a rule, only one copy is pronounced. The choice is mostly considered a matter of phonology (PF). Thus, either the lower or the higher copy of α is deleted at PF; see (37a) and (37b), respectively.

(37) a. $[_{XP} \, \alpha \, [_{YP} \, \not\alpha \,]]$
      b. $[_{XP} \, \not\alpha \, [_{YP} \, \alpha \,]]$

However, there is evidence that more than one copy of α can be pronounced within the same clause (see, e.g., Bošković & Nunes 2007 on so-called wh-copying constructions in, i.a., German, Afrikaans, and Romani). Slavic doubling data such as (16) = (38) show (i) that *by* is indeed copied, (ii) that both copies are within the same clause, and (iii) that both copies are pronounced.

(38) a. Ja by   pogulja-l   by   segodnja večerom.   (Russian)
          I COND take.a.walk-L.SG.M COND today   evening.INS
          'I would like to take a walk tonight'

      b. Čto-by   ja tebja   by   zdes' bol'še ne vide-l.
          that.COND I you.ACC COND here more   NEG see-L.SG.M
          'So that I would not see you here again.'

It is noteworthy that the number of copies of the particle does not exceed two. A straightforward way to account for this fact is that *by* in (38) occupies two distinct syntactic positions, and that there are no more than two such positions available to host its copies.

4.4.2 One particle in I^0 and C^0

From the analysis in §4.3, it follows that one of these positions must be I^0, the basic position of the conditional particle. The second position must be higher in the tree, which makes C^0 a strong candidate. The fact that in the majority of doubling examples – see (38b) – the higher copy of the particle is adjacent to a complementizer, confirms this location.[42]

Additional evidence for C^0 as the target position of the conditional particle comes from examples without doubling. The particle occurs in two alternative positions in all languages under discussion. An example is the Russian minimal pair in (39).[43]

(39) a. Ja vypi-l   by   stakan moloka.
         I drink.up-L.SG.M COND glas   milk.GEN
         'I would like a glass of milk.'   (Xrakovskij 2009: 277)

     b. Ja by   vypi-l   stakan moloka.
         I COND drink.up-L.SG.M glas   milk.GEN
         'As for me, I would like a glass of milk.'   (Russian)

Moreover, constituents that appear clause-initially and left of the conditional particle are interpreted as topic or focus (see, a.o., Willis 2000: 327; Migdalski 2006: 230–231), which holds for sentences with and without doubling; see the Russian examples in (39b) and (38a), respectively.

Finally, Polish provides additional evidence for *by* in C^0: As shown in (40), in subjunctive clauses the conditional particle (with or without agreement) occurs in a sentence-initial position either adjacent to a complementizer or alone. (It must not occur adjacent to V_L.)

(40) Każda matka chce,   (że-)by   jej syn chodzi-ł do
      every mother want.3sg that-COND her son go-L.SG.M to
      przedszkola.
      kindergarten.GEN
      'Every mother wants her son to go to the kindergarten.'
                        (Polish; Jędrzejowski 2020: 109)

---

[42]See also Szucsich (2009: 413) who argues that it is only through "subjunctive raising" of *by* from I^0 to C^0 that the irrealis feature of *by* becomes visible to the embedding matrix verb.

[43]It is an open question which of the two positions is more frequent. Based on a "somewhat restricted data sample" (Hansen 2010: 330), Hacking (1998: 60) spots a tendency for Russian *by* to occur immediately after the subordinating conjunction in the protasis but immediately after the verb in the apodosis of conditional sentences. Hansen (2010) notes that it is not difficult to detect counterexamples and that a more refined corpus-based empirical investigation is necessary to verify the distribution of *by*.

Hagen Pitsch

A doubling example with "solitary" subjunctive *by* is (19) = (41).[44]

(41) nie sądzę,   by-śmy   by-l-i-by   tak blisko siebie   …
      NEG think.1SG COND-1PL be-L-PL.M-COND so close   REFL.ACC
      'I don't think we would be as close to each other …'   (Polish)

To summarize, it is a plausible claim that the conditional particle in particle languages is base-generated in I^0 and can subsequently be copied to C^0.

4.4.3 Fusion and doubling

Following Nunes (2004), cases of simultaneous pronunciation of multiple copies are always due a morphological reanalysis of one of the copies as part of a bigger unit ("word"). He argues that this reanalysis corresponds to a syntactic operation combining two terminal nodes into one, i.e. FUSION (see Halle & Marantz 1993; Muñoz Pérez 2018: §3.1). Crucially, although fusion is the prerequisite for multiple-copy pronunciation, there is no mutual dependence: Fusion can well take place without only one overt copy.

I believe that Nunes' claim is in accordance with the Slavic data: Thus, in (38a), the higher copy of Russian *by* is likely to fuse with C^0 which allows the pronunciation of both *by*-copies. Presumably, the fact that *by* is copied from I^0 to C^0 in the first place is linked to the information-structural status of the subject *ja* 'I': Its interpretation as topic depends on its being in a sentence-initial position and left of *by*, so *ja* itself has to go to SpecCP, while *by* is copied to C^0; see (42).[45]

(42) [$_{CP}$ ja $\varnothing_C$+by [$_{IP}$ poguljal$_V$+by$_I$ [$_{VP}$ segodnja večerom [$_{VP}$ ~~ja poguljal~~]]]]

The same can be stated about (39b), with the exception that here the lower copy of *by* is deleted at PF, which complies with what prescriptive grammars require.

Fusion of *by* with C^0 takes also place in (38b), and again it enables the doubling of *by*; see (43). Since C^0 hosts the complementizer *čto* 'that', the result is the complex C^0 *čtoby*, which "is sometimes treated as an independent lexeme and sometimes as a syntactic combination of two lexemes" (Hansen 2010: 329). I wish to claim that both views are justified: Before the fusion, there are two lexemes. After it, they have become one element.[46]

---

[44]Note that the agreement marker in (41) occurs only on the higher copy. See §4.4.3.

[45]The verb *poguljal* adjoins to I^0 for information-structural reasons, namely to leave the adverbial *segodnja večerom* stranded in a clause-final position (information focus). The verb (meaning) itself is thus presented as (presupposed) background information. Additionally, the verb functions as a phonological host for the enclitic lower copy of *by* in I^0.

[46]Both the subject *ja* and the object *tebja* adjoin to IP to be backgrounded, so the (negated) verb is focused. I ignore the internal structure of the IP (NegP/VP) and V-to-Neg movement in (43).

(43) [$_{CP}$ čto$_C$+by [$_{IP}$ ja [$_{IP}$ tebja [$_{IP}$ by$_I$ zdes' bol'še ne ~~ja~~ videl ~~tebja~~]]]]

As mentioned above, Polish subjunctive clauses can be introduced with or without a complementizer. In other words, *by* alone may, in addition to its basic conditional meaning, assume the function of a complementizer in subjunctive clauses. I suggest that both variants – with and without a "true" complementizer – have the same underlying syntax, the only difference being that C^0 is overtly filled in the former but silent in the latter case; see (44a) and (44b), respectively.

(44) a. [$_{CP}$ że$_C$+by [$_{IP}$ ~~by~~ [$_{VP}$ jej syn chodził do przedszkola]]]
      b. [$_{CP}$ $\emptyset_C$+by [$_{IP}$ ~~by~~ [$_{VP}$ jej syn chodził do przedszkola]]]

In both variants does fusion of C^0 with *by* take place, yielding a complex C-head encoding subjunctive mood (Migdalski 2006: 251). Thus, in a sense, Jędrzejowski (2020: 109) is right in claiming that in (40), "[i]t is [...] *by* which introduces the embedded clause and marks its illocutionary force as well its subordinate status." Crucially, however, the latter is due to the silent C-head fused with *by*.

There is another issue that calls for an explanation: In Polish doubling examples such as (41) with the verb in the first or second person, person/number agreement occurs only once, namely on the higher copy of *by*. If *by* is copied from I^0 to C^0 and subsequently fuses with it, allowing both copies to be pronounced, why does only the higher copy encode agreement? Following the Copy Theory of Movement, the way to account for this pattern is to say that, while *by* is pronounced in both positions, the agreement marker is deleted in the lower one. This is shown in (45).

(45) [$_{CP}$ $\emptyset_C$+byśmy [$_{IP}$ byli$_V$+~~byśmy~~$_I$ [$_{VP}$ *pro* ~~byli~~ tak blisko siebie]]]

As to the reason for the deletion of the agreement marker, I propose that it follows from economy: There is simply no need to pronounce it twice. Note that, in (45), *by* is pronounced in I^0 to reveal the movement (and concomitant backgrounding) of the verb *byli* from V^0 to I^0. There is no need, however, to also pronounce the agreement marker in I^0 since the particle alone is perfectly sufficient to accomplish the task.

5 Summary

This paper provides evidence for a typological division of the Slavic languages into auxiliary languages and particle languages based on the kind of auxiliary

unit used in the L-preterit and L-conditional. Where the members of the former group have inflected auxiliary verbs that encode person/number agreement, the latter have noninflected particles lacking any agreement whatsoever.

The group of particle languages is constituted by Polish, Kashubian, and East Slavic. In the East Slavic languages and Kashubian-B, the particle is generated in I^0 where it encodes the irrealis mood. Crucially, it does not select any specific verb form in V^0 which allows this position to be filled not only with an L-form but also with other forms, most prominently the infinitive.

Polish and Kashubian-A1 are similar in that they, too, have a particle in I^0. However, they stand out within the Slavic branch due to the availability of mobile inflections. In the present paper, these markers are analyzed as syntactic heads generated in SpecIP and subsequently m-merged with I^0, thus yielding a complex inflectional unit encoding tense/mood and agreement (see Embick 1995).

Put differently, Polish and Kashubian-A1 are able to furnish their tense/mood particle in I^0 with person/number agreement, whereas East Slavic and Kashubian-B are not. From this it follows that present-day East Slavic and Kashubian-B have L-forms associated with an underspecified person feature (see Junghanns 1995), while Polish and Kashubian-A1 – on a par with the remaining Slavic languages – have L-participles (number and gender only).

Moreover, the claim that Polish agreement markers are syntactic heads that are initially generated independently of the particle in I^0 provides a straight-forward explanation for why Polish allows, besides *l*-participles, infinitives and *no/to*-forms: The agreement marker may simply not be part of the numeration. If this is the case, the structure is impersonal (lack of person/number agreement).

Crucially, the analysis put forward explains the observation that the conditional in auxiliary languages is limited to verbal L-forms, whereas it is not in particle languages: Auxiliaries retain their "verbal character" (Issatchenko 1940) including the capacity to select specific verb forms in their complement position. By contrast, particles are no verbal categories anymore, which is why there is no selection, hence the wider range of possible forms in V^0.

Finally, the phenomenon of particle doubling attest in colloquial particle languages receives a syntactic explanation: They can be copied from I^0 to C^0, and both copies can be pronounced under specific circumstances (mostly related to information structure).

The present paper shows that there is a remarkable cross-Slavic variation, which is especially true of the auxiliary unit in the conditional periphrasis: While some languages either retain the inherited suffixes or replaced them with present-tense inflections, others developed a pseudo-particle (an underspecified auxiliary verb with a silent agreement suffix), while still others use analogy-based suffixes

which look like the clitic BE-auxiliaries from the L-preterit ("pseudo-clitics"). Despite these differences, all relevant languages possess inflected auxiliary verbs, which distinguishes them from particle languages. An overview is given in Table 14.

Table 14: Auxiliary and particle languages

auxiliary languages with			particle languages with	
inflectional suffixes	pseudo-clitics	silent suffixes	no agreement	mobile agreement
BCMS-A	Czech-B	BL Croatian	Kashubian-B	Kashubian-A
Bulgarian	Macedonian[+]	Macedonian	Belarusian	Polish
Čakavian	Slovak	Slovene	Russian	
Czech-A		L. Sorbian	Ukrainian	
U. Sorbian				

Theoretically, the present paper argues in favor of a formalization of the auxiliary/particle distinction in morphosyntactic terms: Whereas the former are verbal categories generated in Aux^0 or I^0, the latter lost their verbal character and are particles (i.e. tense/mood operators) generated in I^0.

Overall, the paper reveals that the variation in auxiliary units in Slavic periphrases raises a bulk of empirical and theoretical questions. For some, I hope to have provided convincing proposals.

Hagen Pitsch

Abbreviations

1/2/3	first/second/third person	L	*-l*-suffix
ACC	accusative	LOC	locative
AU	auxiliary unit	M	masculine
COND	conditional	N	neuter
DAT	dative	NEG	negation
F	feminine	NOM	nominative
FUT	future	PART	particle
GEN	genitive	PST	past
IMP	imperative	PL	plural
IMPS	impersonal	REFL	reflexive marker
INF	infinitive	SG	singular
INS	instrumental	V_L	verbal L-form
IRR	irrealis		

Acknowledgements

References

Abramowicz, Łukasz. 2008. Using sociolinguistic data to illuminate a theoretical debate: The case of person/number marking in Polish. *University of Pennsylvania Working Papers in Linguistics* 14(1). https://repository.upenn.edu/pwpl/vol14/iss1/1.

Adger, David. 2003. *Core syntax: A minimalist approach.* Oxford: Oxford University Press.

Amir-Babenko, Svetlana. 2007. *Lehrbuch der ukrainischen Sprache.* Hamburg: Buske.

Andersen, Henning. 1987. From auxiliary to desinence. In Martin Harris & Paolo Ramat (eds.), *Historical development of auxiliaries.* Berlin: Mouton de Gruyter. DOI: 10.1515/9783110856910.21.

Bailyn, John F. 2001. The syntax of Slavic predicate case. *ZAS Papers in Linguistics* 22. 1–23. DOI: 10.21248/zaspil.22.2001.99.

Bailyn, John F. 2012. *The syntax of Russian.* Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511984686.

Bartelik, Piotr. 2015. Das Verb *bëc* ‚sein' und seine Funktionen im kaschubischen Tempus- und Genussystem. In Michail L. Kotin & Richard J. Whitt (eds.), *To be or not to be? The verbum substantivum from synchronic, diachronic and typological perspective*, 181–209. Cambridge: Cambridge Scholars Publishing.

Błaszczyk, Izabela Maria. 2018. *Wenn man die polnische Sprache erbt...: Beschreibung und empirische Analyse zum Irrealis-Marker by im Polnischen und seiner Verwendung bei Herkunftssprechern mit Deutsch als dominanter Sprache.* Regensburg: Universität Regensburg. (Doctoral dissertation). https://epub.uni-regensburg.de/37566/1/Blaszczyk_Wenn%20man%20die%20polnische%20Sprache%20erbt%20%283%29.pdf.

Booij, Geert & Jerzy Rubach. 1987. Postcyclic versus postlexical rules in lexical phonology. *Linguistic Inquiry* 18(1). 1–44. https://www.jstor.org/stable/4178523.

Borkovskij, Viktor I. & Pëtr S. Kuznecov. 1965. *Istoričeskaja grammatika russkogo jazyka.* 2nd edn. Moskva: Nauka.

Borsley, Robert D. & María Luisa Rivero. 1994. Clitic auxiliaries and incorporation in Polish. *Natural Language & Linguistic Theory* 12(3). 373–422. DOI: 10.1007/BF01118135.

Bošković, Željko & Jairo Nunes. 2007. The copy theory of movement: A view from PF. In Norbert Corver & Jairo Nunes (eds.), *The Copy Theory of Movement*, 13–74. Amsterdam, Philadelphia: John Benjamin. DOI: 10.1075/la.107.03bos.

Bowers, John. 1993. The syntax of predication. *Linguistic Inquiry* 24(4). 591–656. https://www.jstor.org/stable/4178835.

Breza, Edward. 2009. Das Kaschubische. In Peter Rehder (ed.), *Einführung in die slavischen Sprachen (mit einer Einführung in die Balkanphilologie*, 171–177. Darmstadt: Wissenschaftliche Buchgesellschaft.

Breza, Edward & Jerzy Treder. 1981. *Gramatyka kaszubska: Zarys popularny.* Gdańsk: Zrzeszenie Kaszubsko-Pomorskie.

Browne, Wayles. 1993. Serbo-Croat. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 306–387. London: Routledge.

Browne, Wayles. 2004. Serbo-Croatian enclitics for English-speaking learners. *Journal of Slavic Linguistics* 12(1/2). 249–283. https://www.jstor.org/stable/24599873.

Hagen Pitsch

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from Building 20*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, MA: MIT Press.

Citko, Barbara. 2008. Small clauses reconsidered: Not so small and not all alike. *Lingua* 118(3). 261–295. DOI: 10.1016/j.lingua.2007.05.009.

Corver, Norbert & Jairo Nunes. 2007. *The Copy Theory of movement* (Linguistics Today 107). Amsterdam, Philadelphia: John Benjamin. DOI: 10.1075/la.107.

Cowper, Elizabeth. 2010. Where auxiliary verbs come from. In Melinda Heijl (ed.), *Proceedings of the 2010 Annual Conference of the Canadian Linguistic Association.* Toronto: Candian Linguistic Association. https://cla-acl.ca/pdfs/actes-2010/CLA2010_Cowper.pdf.

Decaux, Étienne. 1955. *Morphologie des enclitiques polonais* (Travaux de l'Institut d'Études Slaves 23). Paris: l'Institut d'Études Slaves.

den Dikken, Marcel. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas.* Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/5873.001.0001.

Dogil, Grzegorz & Miquel Aguado. 1989. Clitics in lexical phonology: Alleged counterevidence? *Linguistische Berichte* 120. 99–116.

Duličenko, Aleksandr D. 2005. Kašubskij jazyk. In Aleksandr M. Moldovan et al. (eds.), *Jazyki mira: Slavjanskie jazyki*, 383–403. Moskva: Academia.

Embick, David. 1995. Mobile inflections in Polish. *Proceedings of NELS* 25(1). article 10. https://scholarworks.umass.edu/nels/vol25/iss1/10.

Franks, Steven & Piotr Bański. 1999. Approaches to "schizophrenic" Polish person agreement. In Katarzyna Dziwirek, Herbert S. Coats & Cynthia M. Vakareliyska (eds.), *Formal Approaches to Slavic Linguistics 7: The Seattle Meeting 1998*, 123–143. Ann Arbor, MI: Michigan Slavic Publications.

Franks, Steven & Tracy H. King. 2000. *A handbook of Slavic clitics.* New York, NY/Oxford: Oxford University Press.

Garde, Paul. 1964. Problèmes du conditionnel dans les langues slaves. *Revue des études slaves* 40. 85–93. https://www.jstor.org/stable/43270547.

Geist, Ljudmila. 2007. Wieviel Verb braucht ein russischer Satz? Überlegungen zur Nullkopula. In Uwe Junghanns (ed.), *Linguistische Beiträge zur Slavistik: XIII. JungslavistInnen-Treffen Leipzig 2004* (Specimina Philologiae Slavicae 149), 123–141. München: Otto Sagner.

Grek-Pabisowa, Iryda & Irena Maryniakowa. 1999. *Współczesne gwary polskie na dawnych kresach północno-wschodnich.* Warszawa: Slawistyczny Ośrodek Wydawniczy.

Hacking, Jane F. 1998. *Coding the hypothetical: A comparative typology of Russian and Macedonian conditionals* (Studies in Language Companion Series 38). Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/slcs.38.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Hansen, Björn. 2010. Mood in Russian. In Björn Rothstein & Rolf Thieroff (eds.), *Mood in the languages of Europe* (Studies in Language Companion Series 120), 325–341. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/slcs.120.18han.

Isačenko, Alexander V. 1962. *Die russische Sprache der Gegenwart. Teil 1. Formenlehre.* Halle (Saale): Max Niemeyer.

Issatchenko, Alexander. 1940. Tense and auxiliary verbs with special reference to Slavic languages. *Language* 16(4). 189–198. https://www.jstor.org/stable/409054.

Ivanov, Valerij V. 1964. *Istoričeskaja grammatika russkogo jazyka.* Moskva: Prosveščenie.

Jędrzejowski, Łukasz. 2020. On the semantic change of evidential argument *jakoby*-clauses in Polish. In Remus Gergel & Jonathan Watkins (eds.), *Quantification and scales in change*, 85–130. Berlin: Language Science Press. DOI: 10.5281/zenodo.3929243.

Julien, Marit. 2002. *Syntactic heads and word formation.* Oxford: Oxford University Press.

Junghanns, Uwe. 1995. Funktionale Kategorien im russischen Satz. In Uwe Junghanns (ed.), *Linguistische Beiträge zur Slawistik aus Deutschland und Österreich: II. JungslawistInnen-Treffen Leipzig 1993*, 167–203. Wien: Gesellschaft zur Förderung slawistischer Studien.

Junghanns, Uwe & Gerhild Zybatow. 2009. Grammatik und Informationsstruktur. In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.), *The Slavic languages: An international handbook of their structure, their history and their investigation* (Handbook of Linguistics and Communication Science 32.1), 684–707. Berlin, New York: de Gruyter. DOI: 10.1515/9783110214475.1.9.684.

Kramer, Christina Elizabeth. 1986. *Analytic modality in Macedonian* (Slavistische Beiträge 198). München: Otto Sagner. http://library.oapen.org/handle/20.500.12657/26454.

Lisac, Josip. 2009. *Hrvatska dijalektologija* 2: *Čakavsko narječje.* Zagreb: Golden marketing – Tehnička knjiga.

Hagen Pitsch

Lubaś, Władysław. 2002. Kaschubisch. In Miloš Okuka (ed.), *Lexikon der Sprachen des europäischen Ostens*, 265–273. Klagenfurt: Wieser.

Markman, Vita. G. 2008. On Slavic semelfactives and secondary imperfectives: Implications for the split 'AspP'. In *University of Pennsylvania Working Papers in Linguistics 14(1)*, 255–268. https://repository.upenn.edu/pwpl/vol14/iss1/20.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1). 69–109. DOI: 10.1162/002438906775321184.

Menzel, Thomas. 2013. Der morphologische Ausdruck des Präteritums im Sprachkontakt: Der Fall des Lemkischen und Kaschubischen. *Wiener Slavistisches Jahrbuch* 1. 99–128. DOI: 10.13173/wienslavjahr.1.2013.0099.

Migdalski, Krzysztof. 2006. *The syntax of compound tenses in Slavic.* Utrecht: LOT. https://www.lotpublications.nl/Documents/130_fulltext.pdf.

Muñoz Pérez, Carlos. 2018. Recognizing copies: On the definition of non-distinctiveness. *Glossa: a journal of general linguistics* 3(1). 40. DOI: doi.org/10.5334/gjgl.271.

Nehring, Gerd-Diether. 2002. Čakawisch. In Miloš Okuka (ed.), *Lexikon der Sprachen des europäischen Ostens*, 247–251. Klagenfurt: Wieser.

Nomachi, Motoki. 2014. On the Kashubian past tense form *jô bél* 'i was' from a language-contact perspective. *Die Welt der Slaven* 55. 218–239. https://www.academia.edu/36365264.

Nunes, Jairo. 1995. *The copy theory of movement and linearization of chains in the Minimalist Program.* College Park: University of Maryland. (Doctoral dissertation). http://hdl.handle.net/11858/00-1735-0000-0022-5E4F-1.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement.* Cambridge, MA: MIT Press.

Panzer, Baldur. 1967. *Der slavische Konditional: Form – Gebrauch – Funktion* (Forum Slavicum 14). München: Wilhelm Fink Verlag.

Pietraszko, Asia. 2018. The distribution of ϕ-probes in the inflectional structure. *Proceedings of the Linguistics Society of America Annual Meeting* 3(58). 1–15. DOI: 10.3765/plsa.v3i1.4355.

Pietraszko, Asia. 2021. Cyclic Selection: Auxiliaries are merged, not inserted. *Linguistic Inquiry* 54(2). 350–377. DOI: 10.1162/ling_a_00439.

Piskorz, Jadwiga, Werner Abraham & Elisabeth Leiss. 2013. Doppelter Grammatikalisierungszyklus und funktionale Universalgrammatik: Am Beispiel des analytischen Perfekts und des Präteritums in der Sprachgeschichte des Polnischen. *Die Welt der Slaven* 58(2). 276–307.

Pitsch, Hagen. 2015. Finiteness, operators and auxiliaries in North Slavic. *Linguistische Berichte* 241. 49–79.

Priestly, Tom M. S. 1993. Slovene. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 388–451. London: Routledge.

Rittel, Teodozja. 1970. Polski czas przeszły i tryb przypuszczający z zaimkiem. *Rocznik Naukowo-Dydaktyczny* 38. 91–108. http://hdl.handle.net/11716/3829.

Rittel, Teodozja. 1973. Szyk enklityk werbalnych w polskich gwarach. *Rocznik Naukowo-Dydaktyczny* 47. 141–148. http://hdl.handle.net/11716/4062.

Rittel, Teodozja. 1975. *Szyk członów o obrębie form czasu przeszłego i trybu przypuszczającego*. Wrocław: Ossolineum.

Smoczyńska, Magdalena. 1985. The acquisition of Polish. In Dan Isaac Slobin (ed.), *The crosslinguistic study of language acquisition*, 595–686. Hillsdale, London: Lawrence Erlbaum Associates.

Smułkowa, Elżbieta. 1999. Sytuacja językowa w Ejszyszkach: wyniki badań sondażowych, prowadzonych na Litwie w 1997 roku. In Józef Porayski-Pomsta (ed.), *Sytuacja językowa na Wileńszczyźnie: materiały sympozjum "Socjo- i psycholingwistyczne uwarunkowania sytuacji językowej w Wilnie i na Wileńszczyźnie": Warszawa 16–18 października 1997 roku*, 54–63. Warszawa: Elipsa.

Sokolova, Anastasia. 2017. *Slavjanskije preterity s istoričeskoj točki zrenija: k istorii form perfekta v russkom jazyke (na materiale služebnych minej na maj XI–XIII vv.)* Brno: Masarykova univerzita. DOI: 10.5817/CZ.MUNI.M210-8877-2017.

Stone, Gerald. 1993. Cassubian. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*, 759–794. London: Routledge.

Svenonius, Peter. 1994. C-selection as feature checking. *Studia Linguistica* 48(2). 133–155. DOI: 10.1111/j.1467-9582.1994.tb00853.x.

Szucsich, Luka. 2009. Obviation und temporale Abhängigkeit bei Subjunktiven. *Zeitschrift für Slavische Philologie* 54(4). 398–415. DOI: 10.1524/slaw.2009.0027.

Toman, Jindřich. 1980. Weak and strong: Notes on *be* in Czech. In Gunter Brettschneider & Christian Lehmann (eds.), *Wege zur Universalienforschung*, 305–310. Tübingen: Narr.

Tomić, Olga Mišeska. 1996. The Balkan Slavic clausal clitics. *Natural Language and Linguistic Theory* 14. 811–872. DOI: 10.1007/BF00133364.

Tomić, Olga Mišeska. 2000. Operator clitics. In Birgit Gerlach & Janet Grijzenhout (eds.), *Clitics in phonology, morphology and syntax* (Linguistics Today 36), 387–404. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/la.36.15mis.

Tornow, Siegfried. 2002. Burgenlandkroatisch. In Miloš Okuka (ed.), *Lexikon der Sprachen des europäischen Ostens*, 235–245. Klagenfurt: Wieser.

Tseng, Jesse. 2009. A formal model of grammaticalization in Slavic past tense constructions. In Manghyu Pak (ed.), *Current issues in unity and diversity of languages. Collection of the papers selected from the CIL 18, held at Korea Uni-*

Hagen Pitsch

versity in Seoul, on July 21–26, 2008, 749–762. Seoul: The Linguistic Society of Korea.

Willis, David. 2000. Verb movement in Slavonic conditionals. In Susan Pintzuk, George Tsoulas & Anthony Warner (eds.), *Diachronic syntax: Models and mechanisms*, 322–348. Oxford: Oxford University Press.

Xrakovskij, Viktor S. 2009. The conditional. In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.), *The Slavic languages: An international handbook of their structure, their history and their investigation* (Handbook of Linguistics and Communication Science 32.1), 484–505. Berlin, New York: de Gruyter. DOI: 0.1515/9783110214475.1.4.275.

Zielińska, Anna. 2002. *Polska mniejszość na Litwie Kowieńskiej: studium socjolingwistyczne*. Warszawa: Slawistyczny Ośrodek Wydawniczy.

Chapter 12

Number mismatch effect and processing cataphora in a *pro*-drop language: The case of Slovenian

Matic Pavlič[a] & Arthur Stepanov[b]

[a]University of Ljubljana [b]University of Nova Gorica

Cataphora (also known as backward anaphora) is a type of pronominals that precede their antecedents linearly. Previous research on processing cataphora has explored the idea that cataphoric pronouns trigger a forward-looking active search for an antecedent to establish a coreference relation similar to a filler-gap dependency between a fronted wh-phrase and its base-generated syntactic position (Cowart & Cairns 1987). Van Gompel & Liversedge (2003) have shown that in an active search, the parser establishes a cataphoric coreference before considering pronominal phi-features. This results in a gender mismatch effect: sentences with incongruent incoming NP antecedents were more difficult to read than their congruent counterparts, as evidenced by slower reading times and eye movement regressions. In this paper, we report the results of a self-paced reading experiment in which the active search hypothesis is further tested by examining online cataphora resolution with respect to the number feature in Slovenian, a pro-drop language with a rich nominal and verbal morphology.

Keywords: cataphora, *pro*, feature mismatch effect, forward-looking active search, self-paced reading, Slovenian, psycholinguistics

1 Introduction

The interpretation of pronominal expressions such as the Slovenian reflexive *svoj* 'self's' in (1) depends on their relation to referring expressions in the context in which they are used.

Matic Pavlič & Arthur Stepanov

(1) Svoje sorodnike kliče po več krat na dan.
   SELF's relatives.PL calls.SG at several times on day
   'He calls his relatives several times a day.'   (Slovenian)

If we restrict ourselves to the syntactic context, a dependency relation is established between the base-generated position of a pronominal expression and a referring expression: The latter acts as an antecedent and c-commands the former according to principles A (reflexives) and B (pronouns) of the binding theory (Chomsky 1981; Reinhart 1983). The linear order of the two expressions may be changed so that the pronominal element is spelled out after the antecedent, as in the case of forward anaphora, or before the antecedent, as in the case of backward anaphora or cataphora, but it should not itself be in the position of c-commanding the antecedent according to Principle C (the referring expressions should not be bound at any times).

Due to the relative word order of the two expressions, anaphora and cataphora resolutions differ fundamentally in real time sentence comprehension (Lust 1986; Reinhart 1986; Blackwell 2003; Tsimpli et al. 2004; Kennison et al. 2009; Lobo & Silva 2016). In the case of an anaphora, the parser encounters a pronominal expression and simultaneously considers all possible candidates previously integrated into the incoming sentence and stored in working memory. This mechanism is closely related to the processes of memory retrieval (Chow et al. 2014). In the case of a cataphora, the parser does not find an available antecedent in the previous syntactic context. Therefore, the parser expects to find it in the incoming sentence material and considers each subsequent noun phrase as a potential antecedent. Crucially, the "active" or "impatient" parser does not wait until all potential antecedents are stored in working memory but evaluates them one by one as they are integrated into the structure. More specifically, the parser attempts to associate the cataphora with the first potential antecedent as soon as the structural requirements of its c-commanding relation to the cataphora are confirmed – but before it considers the interpretative requirements (phi-feature matching).

This analysis was first introduced in the seminal work of Cowart & Cairns (1987) who observed a strong preference for linking cataphora to the first possible noun phrase encountered by the parser. Sturt (2003) and Van Gompel & Liversedge (2003) supported this finding by measuring the gaze direction and reading time of sentences such as (2).

(2) a. When he$_i$ appeared, the king$_i$ immediately greeted the boys very warmly.

b. When they$_i$ appeared, the boys$_i$ immediately greeted the king very warmly.

c. When they$_i$ appeared, the king immediately greeted the boys$_i$ very warmly.

d. When he$_i$ appeared, the boys immediately greeted the king$_i$ very warmly.

In (2a) and (2b), the first potential NP antecedent bears cataphora-congruent phi-features, in this case, number. The parser associates the cataphora *he/they* with the referring expression, as marked in the example by matching indices. Van Gompel & Liversedge (2003) used examples such as (2a) and (2b) as a baseline for analysing the reading times and gaze directions of examples such as (2c) and (2d). For the latter, they found an effect when the number value of the first potential NP antecedent (or gender, in other experiments) did not match that of a preceding cataphor. Van Gompel and Liversedge refer to this as the mismatch effect and claim that it reflects the parser's unsuccessful attempt to establish a referential dependency between the antecedent and the cataphora. They add that the gender or number mismatch effect can only occur when the parser attempts to establish a referential dependency before comparing the features of the NP with those of the cataphora.

Kazanina (2005) and Kazanina et al. (2007) replicated the gender mismatch effect with a paradigm of self-paced reading and explored it in more detail. They attributed the slowdown to the parser's search for an antecedent, which involves predictive processes. One of these predictive processes is the active search mechanism, which was originally used to interpret filler-gap dependencies (Crain & Fodor 1985; Stowe 1986; Frazier & Clifton 1989; Frazier & Flores d'Arcais 1989; Garnsey et al. 1989; Kaan et al. 2000, Stepanov & Stateva 2015). Wh-dependencies are established between the fronted wh-phrase and its base-generated position. The search for a gap begins as soon as a wh-phrase is processed. This was demonstrated in online experiments by filling the gap position with an overt element (which prevented the parser from interpreting the wh-phrase in that position), resulting in longer processing times compared to a sentence without an overt element in the gap position (Crain & Fodor 1985; Stowe 1986; Lee 2004). Thus, the active search mechanism assumes that the parser expects a gap as soon as a wh-phrase is encountered (Frazier & Clifton 1989). In the case of pronoun interpretation, the active search mechanism predicts that a search for an antecedent

will be initiated as soon as a pronoun is encountered (Frazier & Clifton 1989; Kazanina et al. 2007; Kazanina & Phillips 2010) to resolve the interpretation of the pronoun. Although pronouns may have an antecedent outside the sentence in which they occur, the active search mechanism states that searching for an antecedent within the sentence is the default strategy when there is no preceding context.

Moreover, Kazanina et al. (2007), Aoshima et al. (2009), Kazanina & Phillips (2010), and Yoshida et al. (2014) show that the gender mismatch effect is absent in syntactic contexts where the incoming NP is not available for coreference because it cannot be bound by a c-commanding expression (Principle C of Binding Theory; Chomsky 1981; Reinhart 1983). For several types of structures containing cataphora and potential NP antecedents to which Principle C applies, the researchers found no differences in reading time between the gender-congruent and gender-incongruent conditions. In examples (3a) and (3b), the potential NP antecedent is within the c-command domain of a cataphoric pronoun. Therefore, the parser discards NP as a potential antecedent without looking at the phi-features of NP *Kathryn* (F) or *Russell* (M) and without comparing them to the cataphoric features (F).

(3) a. * She$_i$ was taking classes full-time while Kathryn$_i$ was working two jobs to pay the bills.

    b. * She$_j$ was taking classes full-time while Russell$_j$ was working two jobs to pay the bills. (Kazanina et al. 2007)

Cataphoric coreference has also been investigated using the event-related potential (ERP) technique. According to previous literature on gender error processing in this domain, frontal positivity within the P600 time window (reflecting syntactic repair) and late anterior negativity (reflecting additional load on working memory) were expected. These effects have been associated with less preferred continuations of syntactically ambiguous sentences (Osterhout & Holcomb 1992; Barber et al. 2004; Gouvea et al. 2010) and agreement errors (Hagoort & Brown 1999; Osterhout & Mobley 1995). It has also been argued that the P600 signals difficulties or errors in integrating syntactic dependencies, which are predicted according to context (Federmeier et al. 2007; Delong et al. 2011, 2014; Van Petten & Luka 2006). ERP results in reading comprehension of cataphoric dependencies in Dutch (Pablos et al. 2015; Pablos et al. 2018) showed that gender incongruence leads to P600 only in positions where the binding principles are satisfied. There was no ERP effect in the incongruent NP antecedent that would violate Principle C if coreference with cataphora had been established. According to the

researchers, the negativity in these cases reflects the cancellation of the tentative antecedent and not the gender incongruence between cataphora and antecedent.

2 Cataphora in *pro* contexts

In addition to syntactic constraints such as binding theory, pronominal resolution (both anaphoric and cataphoric) is also determined by several different language-specific factors, such as the complementary/ shared distribution of null and overt personal pronouns (Bosch et al. 2003; Wilson et al. 2009; and the references below).

The interpretative properties of null *pro* in contexts of anaphora and cataphora compared to those of overt pronouns have been studied mainly on Chinese (Yulong & Xiaodan 2007; Zhiyi 2019), Italian (Carminati 2002, 2005; Fedele & Kaiser 2014) and Spanish material (Alonso-Ovalle et al. 2002). Carminati (2002) shows that the shared syntactic distribution of null anaphora and overt anaphora causes the parser to preferentially associate null anaphora with more prominent antecedents (e.g. subject of the main clause) and to associate overt anaphora with less prominent antecedents (e.g. object of the main clause), where prominence is structurally defined, e.g. the antecedent in the highest specifier projection (SpecIP) is considered more prominent than the antecedent in the lower projections. In Italian, the subordinate pronominal expression refers to the main clause subject *Mario* when the former is realised as a null pronoun *pro* (4a). On the other hand, the subordinate pronominal expression refers to the indirect object *Giovanni* of the main clause when the former is realised as the overt pronoun *lui* (4b).[1]

(4)   a.   Mario$_i$ ha telefonato a Giovanni$_j$ quando *pro$_i$* aveva appenafinite
          Mario has telephoned to Giovanni when   had just-finished
          di mangiare.
          to eat
          'Mario called John, when he just finished eating.'

     b.   Mario$_i$ ha telefonato a Giovanni$_j$ quando lui$_j$ aveva appenafinite
          Mario has telephoned to Giovanni when he had just-finished
          di mangiare.
          to eat
          'Mario called John, when the latter just finished eating.'

                                                            (Italian; Carminati 2002)

---

[1]Note that Belletti et al. (2007) found the opposite result: Null pronouns tend to prefer the object in anaphoric configurations, while overt pronouns seem to prefer the extra-sentential referent.

In offline experiments assessing interpretations, Fedele & Kaiser (2014) extended this study to cataphora cases. These authors found that (i) null cataphors tend to prefer subject antecedents compared to overt cataphors, and (ii) overt cataphors tend to prefer object antecedents compared to null cataphors, such that the null pronominal expression *pro* in (5a) refers to the subject NP *Maria*, while the pronominal expression *lei* in (5b) refers to the object NP *Rita*.

(5)   a.   Mentre *pro*$_i$ parla del   viaggio a Londra Maria$_i$ abbraccia Rita$_j$.
         while   speaks of-the trip   to London Maria hugs   Rita
         'While speaking about the trip to London, Maria hugs Rita.'

    b.   Mentre lei$_j$ parla del   viaggio a Londra Maria$_i$ abbraccia Rita$_j$
         while she speaks of-the trip   to London Maria hugs   Rita
         'While she speaks about-the trip to London Maria hugs Rita.'
                                                   (Italian; Fedele & Kaiser 2014)

These results recall Carminati's (2002, 2005) Position of Antecedent Hypothesis (PAH) for anaphora contexts. According to PAH, null pronouns refer to a structurally prominent antecedent in a SpecIP position, and overt pronouns refer to an antecedent lower in the clause structure. The contrast in (5) is broadly consistent with PAH if the latter is construed in terms of preferences rather than absolute expectations. Importantly, when *pro* is part of the main clause (6), it acquires an extra-linguistic 'someone else' interpretation, suggesting that speakers are guided by Principle C of Binding Theory which overrides the intra-sentential referential bias.

(6)   *pro*$_k$ parla del   viaggio a Londra mentre Maria$_i$ abbraccia Rita$_j$.
       speaks of-the trip   to London while   Maria hugs   Rita
      'Somebody speaks about the trip to London, while Maria hugs Rita.'
                                                       (Fedele & Kaiser 2014)

Alonso-Ovalle et al. 2002 test the PAH in Spanish and find that it also holds for the Spanish null subject *pro*, both within sentences and across sentences. Moreover, the interpretation of the Spanish *pro* is sensitive to antecedents that are discourse topics and thus interact closely with the topic-focus system (see also Yulong & Xiaodan 2007; Zhiyi 2019).

3 Research hypothesis

In the present work we extend research on the interpretative properties of the null subject *pro* in the context of cataphora to Slovenian, another *pro*-drop language. In Slovenian, there is no overt pronoun available in configurations like

in Italian (5b) above. This is evident in example (7b), where the overt personal pronoun *ona* cannot co-refer with either the subject NP *Marija* or the object NP *Rita*.

(7) a. Ko   *pro*$_i$ pripoveduje o Londonu Marija$_i$ objema Rito$_j$.
         when   speaks   of London Marija hugs   Rita
         'While speaking/ she speaks about London, Marija hugs Rita.'

     b. Ko   ona$_k$ pripoveduje o Londonu Marija$_i$ objema Rito$_j$.
         when she speaks   of London Marija hugs   Rita
         'Somebody speaks about London, while Marija hugs Rita.'

(Slovenian)

This is in line with the Avoid Pronoun Principle (Chomsky 1981), according to which a null variant is preferred to an overt pronoun whenever possible, provided that a language has a null and an overt subject pronoun in the given syntactic environment. Consequently, Slovenian provides a good testing ground for online comprehension of null cataphors, as language-specific factors such as PAH (see the previous section) and the contrast between null and overt pronouns (Bosch et al. 2003; Wilson et al. 2009) do not clash with syntactic constraints. Moreover, Slovenian has a rich verbal inflection as well as overtly realised gender and number features on NP. The topic of cataphora processing has so far received little attention in the Slovenian psycholinguistic literature. These considerations were crucial for the focus of the present study.

Previous studies of *pro*-cataphora processing have mostly relied on global or offline evaluation metrics, such as comprehension questions. It is not clear whether the active search mechanism postulated for overt cataphoric pronouns that provide unambiguous cues works in a similar way for silent pronouns such as *pro*. Our main interest was therefore in better understanding the mechanism of establishing a cataphoric dependency in the absence of an overt pronoun cue. Specifically, building on the earlier experimental findings on the subject-oriented nature of the null cataphoric *pro*, we asked whether null *pro* triggers the parser's active search mechanism that links the *pro* to the subject of the main clause, as described above. Our second goal was to investigate the mechanism of active search in *pro*-cataphora at the level of specific phi-features by exploiting the rich Slovenian morphology. Specifically, we were interested in the mismatch effect in the context of number. Carminati (2002, 2005) argues that number is a better pronoun disambiguator than gender in ambiguous anaphora contexts in the feature hierarchy (Greenberg 1963; Silverstein 1985).[2] Van Gompel & Liversedge (2003)

---

[2]Studies such as that of Mancini et al. (2014) support the idea that features do not in fact behave

showed that number initiates a mismatch effect in a similar fashion to gender, for overt pronoun contexts. To our knowledge, the number feature has not yet been studied in the domain of the mismatch effect in cataphors in the absence of an overt cue. If the incongruence or mismatch effect holds in Slovenian with null *pro* one can also ask how it is distributed in the time course of reading the respective sentence and how different values of the number feature may modulate this effect, given that, on null hypothesis, the incongruent conditions are expected to manifest a similar performance pattern.

4 Experiment

We conducted an online self-paced reading experiment in Slovenian *pro*-cataphora sentences with *when*-subordinate clauses. In this experiment we explored the number congruency effect associated with integration of *pro* with the subject of the main clause and whether this effect is sensitive to the actual number feature of a silent subject pronoun *pro* in the function of a cataphor. In a self-paced reading task, the informant reads individual sentences on a computer screen, with stimulus sentences presented word by word in moving window mode (Just et al. 1982). When the informant presses a predefined key, the first word is displayed. The next time he presses the key, the first word disappears and the next appears. The informant continues in this way until the end of the sentence. Since the informant sees only one word at a time, he must retain the incoming information in his short-term memory. Since the participant does not receive a direct cue to the pronominal reference in the case of the silent *pro*, but must infer it from a more indirect cue, participial agreement, when he encounters the subject of the main clause, the subsequent active search procedure of "looking forward" presumably contributes to the load on short-term memory. The self-paced reading paradigm was chosen because it allows us to test the difference between two lexically identical sentences that differ in their functional elements and/or their phi-features. The method thus allows a direct comparison between two related syntactic structures, e.g. between *when*-subordinate clause with congruent and incongruent number feature on the second (i.e., main) clause subject. When reading a sentence in self-paced mode (i.e., word by word), readers show longer

---

in the same way. In self-paced reading of online processing of subject-verb agreement in Italian, where both person and number agreement factors were manipulated, results showed a greater processing penalty for violations of person agreement compared to number agreement. This was interpreted as evidence for separate access to the two features. On the other hand, Van Gompel & Liversedge (2003), among others, report a generally similar pattern of processing the number and gender features.

reading times in the region, which causes additional mental load due to syntactic repairs, less preferred readings, agreement errors, difficulties in integrating syntactic dependencies, etc.

4.1 Materials

Our chosen sentences consisted of a main clause and a subordinate clause which linearly preceded the former. The main clause was a transitive clause with a time adverbial phrase and all arguments overtly expressed by referring expressions. The subordinate clause was a copular clause headed by a connector *when*, *if* or *because* (evenly distributed across items). The subordinate (and preceding) clause always included the null subject *pro*, the auxiliary verb *be* as copula, and an attributive adjective (or an adjective-like present participle).

The target material was arranged in a 2 × 2 design crossing factors Congruency (congruent, incongruent) and Number (sg/pl) on the first (i.e., subordinate) adjective and its accompanying auxiliary verb. This resulted in four conditions in the manipulation: subordinate null subject in singular + main clause overt subject in singular ((8a); congruent), subordinate null subject in singular + main clause overt subject in plural ((8b); incongruent), subordinate null subject in plural + main clause overt subject in singular ((8c); incongruent), and subordinate null subject in plural + main clause overt subject in plural ((8d); congruent). All target sentences are grammatical in the normal everyday language.

(8) a. Ko je osamljen, stric kliče sorodnike po več
   when is.aux.sg lonely.sg uncle.sg calls.sg relatives.pl at several
   krat na dan.
   times a day
   'When he is lonely, the uncle calls relatives several times a day.'
   [+congr, +sg]

   b. Ko je osamljen, sorodniki kličejo strica po več
   when is.aux.sg lonely.sg relatives.pl call.pl uncle.sg at several
   krat na dan.
   times a day
   'When he is lonely, the relatives call the uncle several times a day.'
   [−congr, +sg]

   c. Ko so osamljeni, stric kliče sorodnike po več
   when are.aux.pl lonely.pl uncle.sg calls.sg relatives.pl at several
   krat na dan.
   times a day
   'When they are lonely, the uncle calls the relatives several times a
   day.'
   [−congr, −sg]

351

    d. Ko   so   osamljeni, sorodniki   kličejo strica   po več
          when are.AUX.PL lonely.PL   relatives.PL call.PL uncle.SG at several
          krat   na dan.
          times a   day
          'When they are lonely, the relatives call the uncle several times a day.'
                                          [+congr, −sg]

For each condition, 6 sentences were formed, giving a total of 24 target item sets. They were counterbalanced so that each participant saw only one lexical version of a given item per condition. In addition, 48 filler sentences (32 of which represented conditions from an unrelated experimental manipulation) were added. The total number of stimulus sentences was thus 72. Each sentence (including the filler sentences) was followed by a *yes-no* comprehension question that tested the understanding of the event described in the stimulus sentence. For instance, a sentence from a set like the one in (8) could be followed by a question such as 'Does the uncle visit the relatives?' expected to be answered with a 'no' (for all the sentences in a set). The proportion of correct 'yes' and 'no' responses to the comprehension questions was evenly distributed across conditions. Per word reading times was the only dependent variable in this manipulation.

4.2 Participants

Thirty-three self-reported adult native speakers of Slovenian (21 female, mean age = 36.69, SD = 14.27; median age = 31) participated in the experiment voluntarily (providing online informed consent), anonymously, and without material compensation. All participants had normal or corrected-to-normal vision and reported no neurological disorders. One participant was excluded because they did not meet the 66.6% (two-thirds) accuracy threshold for *yes-no* comprehension questions, pre-set in advance. This left the data from 32 participants for further analysis.

4.3 Procedure

Participants were instructed to read the sentences at a natural pace and to make sure that they understood what they were reading. If an incorrect answer was given to a comprehension question, they received feedback. If a correct answer was given, they received no feedback. No answer within 7 seconds was counted as an incorrect answer. Concentration and correct comprehension were checked

with a *yes-no* question that followed each sentence and referred to its content. Before the main experiment, subjects read 4 practice sentences to familiarise themselves with the task. The experiment was programmed on the web-based Ibex Farm platform (by Alex Drummond; https://adrummond.net/ibexfarm). The order of stimulus presentation was pseudo-randomised for each participant by the experimental software, and it was ensured that at least 1 filler sentence was between two target items. The entire experimental session lasted 20–25 minutes. Participants performed the task at a location of their choice without coming to the lab. They were specifically instructed to ensure that external disturbances were kept to a minimum while performing the task.

4.4 Data analysis

Only the sentences followed by a correctly answered comprehension question were selected for analysis, which constituted 84.2 % of the total data. For all analyses, the last two regions (second part of time adverbial phrase) of the sentence were removed. Reading times shorter than 90 ms or longer than 3000 ms were trimmed as unlikely to have been generated by relevant linguistic processes. This affected approximately 0.2 % of the total data. Outliers were then identified and excluded from further analysis. The criterion was 3 standard deviations from the mean RT for a given condition and region, for each participant (excluding 79 measurements or additional 1.5 % of the total data).

To analyse the reading time data, we constructed linear mixed-effects models (Bates et al. 2015). This allowed us to model individual RTs based on manipulated fixed factors, namely Congruency and Number, while accounting for random variance in the form of participant and item. We used a maximal or near maximal random effects structure adding random slopes for Congruency and Number up to model convergence (Matuschek et al. 2017). Analyses were conducted using the *lme4* package in R version 4.0.2 (R Core Team 2020). We report χ^2 and *p*-values for main effects based on the likelihood ratio test, which compares a model containing the fixed effect of interest to a model that is identical in all respects except the fixed effect of interest, using the χ^2 distribution. *P*-values for pairwise comparisons with Tukey adjustment were obtained using the *multcomp* package in R.

4.5 Results

The time course of reading sentences in all four conditions is shown in Figure 1. Overall, reading times were higher in the non-congruent conditions (a subordinate null subject in singular followed by a plural main-clause subject (Npl) and

Matic Pavlič & Arthur Stepanov

a subordinate null subject in plural followed by a singular main-clause subject (Nsg)) than in the congruent conditions. Total reading times with standard errors per condition are shown in Table 1.
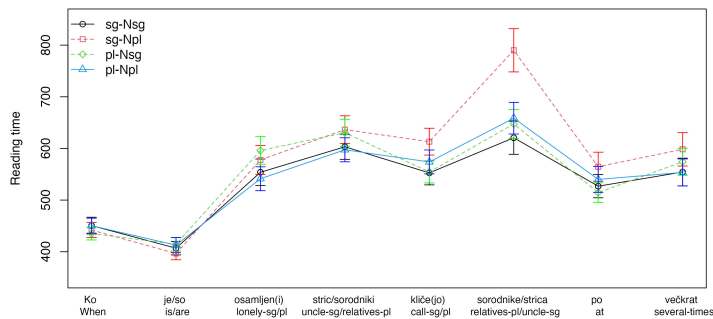
Figure 1: Time course of self-paced reading (the last two regions not shown)

Table 1: Contrasts across the Congruence and Number factors, total reading times

Main effect (CONGR)		Conditions	Total RT (ms)	SE	Main effect (NUM)	
$\chi^2(2) = 11.102$ $p = 0.0008{***}$	+congr	pl-Npl	4751	220	PL	$\chi^2(2) = 6.461$ $p = 0.01{**}$
		sg-Nsg	4849	286	SG	
	−congr	sg-Npl	5254	299		
		pl-Nsg	4976	242	PL	
Interaction CONGR*NUM: $\chi^2(1) = 5.0685$, $p = 0.024{*}$						

As Table 1 shows, there are overall main effects of Congruency as well as Number. Moreover, Congruency interacted with Number: there was no difference in reading times between sentences with singular and plural subordinate null subjects in the congruent conditions but sentences with singular subordinate null subjects were read more slowly (about 40 ms per word) than those with plural subordinate null subjects in the incongruent conditions. The main sites of slow-down were primarily the post-antecedent regions, i.e., the verb phrase following the second (i.e., main) clause subject.

Per region analyses revealed that there were no main effects or interactions up to the main verb region ($ps > 0.10$). At the main verb region (cf. *kliče*), there is no main effect of congruence, but there is a marginal main effect of Num- ber ($\chi^2(2) = 4.85, p = 0.08$) indicating that conditions with the singular sub- ordinate null subjects are read about 50 ms slower at this region, and there is a marginal interaction of the two factors ($\chi^2(1) = 3.39, p = 0.065$); pairwise com- parisons indicate that the sg-Npl condition stands out in terms of higher reading times compared to the other conditions, although the contrasts do not quite reach significance ($ps > 0.10$). Furthermore, at the direct object region there is again no main effect of congruence, but there is a significant main effect of Number ($\chi^2(2) = 17.32, p < 0.001$) and there is an interaction between the two factors ($\chi^2(1) = 11.52, p < 0.001$). Pairwise comparisons show that this region takes longer to read in the incongruent sg-Npl condition than in the congruent condi- tion sg-Nsg ($\beta = 161.84, \text{SE} = 33.4, t = 4.844, p < 0.001$), as well as in the other in- congruent condition, namely, pl-Nsg ($\beta = 138.67, \text{SE} = 33.6, t = 4.125, p < 0.001$). In contrast, there is no difference in reading this region in the other incongru- ent condition, pl-Nsg, in comparison to the corresponding congruent condition pl-Npl ($\beta = 1.08, \text{SE} = 33.5, t = 0.032, p > 0.10$). The main effect of Number marginally persists up to the next region (the preposition in Figure 1, ($\chi^2(2) = 4.52, p = 0.10$); no other effects were observed in this and the final regions.

5 Discussion and conclusions

The congruency or mismatch effect observed in our experiment suggests that Slovenian speakers are sensitive to the interpretational properties of the pronoun, despite its silent character: the parser initiates an active search mechanism in the case of the null subject *pro*, just as it would in the case of the overt pronoun. Our results are largely consistent with those of the eye-tracking experiment in which Van Gompel & Liversedge (2003, Experiment 3) tested the number-mismatch ef- fect in overt cataphora contexts (cf. (2) above), but there are also some important differences. Van Gompel and Liversedge reported significantly prolonged first- pass reading times (i.e., the sum of all fixation durations from the first fixation within a region to a fixation outside the region) in cases of incongruence or num- ber mismatch (cf. (2b)), compared to congruent cases (cf. (2a)) in (i) the region immediately following the main subject NP (which in their case is an adverb that does not occur in (2)), (ii) on the main verb (cf. 'visited' in (2b), difference only by item), and they also reported significant first-pass regressions (i.e., the percentage of leftward eye movements crossing the left boundary of the region

initiated immediately after a first-pass fixation in the region) on the direct object; the effects decay after this region. In our study, the per region dynamics is very similar for online reading: the divergence starts at the verb (there was no preverbal element in our stimuli, such as an adverb in the above study), goes steeply up to the direct object, and decays in the final regions. But there are also at least two important differences between our results and those of Van Gompel and Liversedge.

First, in our study, the main effect of congruency was observed only for total reading times across critical regions, but not for per-region measurements, whereas the authors cited above report this main effect for three critical post-subject regions. This suggests that our congruence effect is less "pronounced" than that in Van Gompel and Liversedge's study, as the difference per region is sufficient to sum up to a global-level effect, but insufficient to independently mark individual regions. If this difference proves robust, it may indeed point to an important aspect in which the processing of a *pro*-cataphora differs from that of overt pronominal cataphors in previous studies. An obvious caveat is that the experimental methodology of the two studies is different. Whether the contrast remains when the methodology is made consistent needs to be investigated further.

Another important difference between the two studies is that in our study congruency interacted with number consistently across the (post-subject) regions of interest, whereas in Van Gompel and Liversedge's study no interaction between congruency and number was observed in any post-subject region; each of the two factors affected eye movement measures independently. Our study also revealed the main source of this interaction, namely the sg-Npl condition. The fact that the two factors interacted consistently, with no main effect of congruency in specific regions, may suggest that the parser has an increased sensitivity to the number feature in the context of the active search mechanism activated by the silent *pro*. Recall that there is no overt pronominal cue to the number feature, so the parser must infer the number feature based only on the inflection of the subordinate copula plus adjective. Van Gompel and Liversedge (2003) argue, based on their results, that the use of morphological information occurs only after coreference relations have been computed (see also Cowart & Cairns 1987; Kazanina et al. 2007; Kazanina & Phillips 2010). Our results in the sg-Npl condition are broadly consistent with this conjecture. However, the performance of our speakers in the other incongruent condition, namely pl-Nsg, casts doubt on it: In this condition, neither a congruency nor a mismatch effect was observed, a priori suggesting that speakers use the morphological information about number early enough. This divergent pattern calls for an explanation. Here we offer

some initial thoughts on a possible line.[3]

One possibility is that the parser behaves differently when it tries to establish a coreference in incongruent contexts with a singular and a plural *pro* in Slovenian: While the active search mechanism accesses the singular value of *pro* from the beginning (and before the coreference is established), an alternative, more global parsing strategy could work with the plural *pro* in accordance with the schedule a la Van Gompel and Liversedge. This alternative is clearly unattractive, as it seems to overstate the relationship between parsing strategies and lexically encoded information about *pro* nominals, such as morphological features. Another alternative, which we consider more feasible and promising, is that an additional factor plays a role in modulating the active parsing scenario, which is sensitive to specific number features.

We hypothesise that this additional factor is grammatical in nature and has to do with the way various number features are semantically encoded in the grammar module, the latter playing an active role in driving sentence processing. Informally speaking, this encoding has to do with markedness of certain feature values. In contrast to the commonly held view in theoretical and experimental research that the singular is unmarked while the plural is a marked form (cf. Bock & Eberhard 1993), there is a growing consensus in the recent semantics literature on the opposite view according to which the singular is endowed with an additional property in its lexical entry, namely the singularity presupposition (the presupposition that the cardinality of the set in question is exactly 1). In this sense, the singular is semantically more "loaded" than the plural and can therefore be regarded as having a more marked value (Sauerland 2003, Sauerland et al. 2005; Spector 2007).[4]

Establishing a coreference in real time involves matching a previously activated feature value between *pro* and its antecedent (the latter term is of course not very appropriate in the cataphora context). In the case of an incongruent sg-Npl condition, this matching needs to include the singularity presupposition of *pro*: Since the (plural) main clause subject lacks this property, the matching

---

[3]An anonymous reviewer suggests that the mismatch in the pl-Nsg condition could be tolerated because of an additional available parse compatible with a split antecedent reading, as in, e.g. *When pro$_{i+j}$ are lonely, the uncle$_i$ calls the relatives$_j$ several times a day.* While a reasonable possibility for Slovenian, it does not sit easily with the results of Van Gompel & Liversedge's (2003) original English experiment (with an overt cataphora) whereby no contrast was reported between the two non-matching conditions (cf. (2c) vs. (2d)).

[4]The arguments for this view come from the domains of using the plural under the scope of negation, downward-entailing operators and the like. See the references in the text for more discussion.

cannot be complete and the mismatch effect occurs. In contrast, in the incongruent pl-Nsg condition, there is no element of presupposition checking in the process of plural *pro* establishing coreference. A possible mismatch effect is thus excluded. Note that this scenario rests on the assumption that the feature matching procedure is asymmetric. This naturally follows from the "forward-looking" character of cataphoric dependency formation: the singularity presupposition is triggered by the element initiating the dependency, that is, *pro*. The antecedent just needs to match this property, not the other way around.

   This of course raises a question as to why a similar avoidance effect does not occur in the corresponding English constructions of Van Gompel and Liversedge (cf. (2)), where the incongruence or mismatch effect occurs in both directions. We believe the answer has to do with the morphological realization of the corresponding pronoun (overt vs. null). Null *pro* is generally considered underspecified compared to overt pronominal and may instantiate less morphosyntactic structure than the latter (Cardinaletti & Starke 1999). It is possible that in English checking the morphological plural feature is additionally required as part of establishing coreference given that the parser activates it by reading the overt cue (the pronoun itself), whereas in Slovenian this additional process is not necessarily initiated due to the phonologically silent character of *pro*. In other words, besides the semantic part of establishing coreference, the English coreference formation includes a morphological part, whereas the Slovenian dependency processing does not. This would explain the divergent way the incongruent conditions are processed in Slovenian in the general context of the active search mechanism and highlight another difference between *pro*-drop and non-*pro*-drop languages in terms of cataphora resolution. Moreover, this provides an interesting starting point for further research, possibly involving other *pro*-drop languages and/ or feature continua.

   The above line of argument underscores the role of morphological component in establishing coreference. Within a model of syntactic parsing of the weak interactive type (e.g. Altmann & Steedman 1988) the processor tries to compute a coreference relation between the cataphoric pronoun and the first available antecedent. In the pl-Nsg incongruent condition, the unmarked character of *pro* does not prevent establishing this coreference relation in either English or Slovenian but the additional morphological processing routine results in a mismatch that blocks this relation in the former, but not the latter. According to this model, processing difficulty in this condition occurs because the syntactic component of the processor allows for the coreference but the morphological information on the overt pronoun in English is inconsistent with it. In the sg-Npl incongruent condition, the marked character of singular *pro* triggers the mismatch effect in

both languages regardless of the morphological realization. Alternatively, within the modular model of processing, the coreference relation happens during the first state of analysis on the basis of only syntactic information. Disruption due to an additional morphological process in English, but not in Slovenian, happens at the second, post-syntactic stage when the processor recognizes the initial analysis as inconsistent with morphological information on the pronoun and therefore has to revise the initially postulated coreference relation.

Overall, the patterns of results observed in this study demonstrated that the null *pro*, postulated on the basis of agreement information in the auxiliary+adjective complex, initiates an active search for an incoming NP as a target antecedent. Establishing a cataphoric coreference with null *pro* proceeds similarly in many respects to the corresponding process with overt pronoun, with some important differences in terms of the construction of an online representation of the coreference that bears on the overt/ null morphological distinction.

Acknowledgments

References

Alonso-Ovalle, Luis, Susana Fernández-Solera, Lyn Frazier & Charles Clifton. 2002. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Italian Journal of Linguistics* 14(2). 151–170.

Altmann, Gerry & Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30(3). 191–238. DOI: 10.1016/0010-0277(88)90020-0.

Aoshima, Sachiko, Masaya Yoshida & Colin Phillips. 2009. Incremental processing of coreference and binding in Japanese. *Syntax* 12(2). 93–134. DOI: 10.1111/j.1467-9612.2009.00123.x.

Barber, Horacio, Elena Salillas & Manuel Carreiras. 2004. Gender or genders agreement. In *On-line study of sentence comprehension*, 309–328. Psychology Press. DOI: 10.4324/9780203509050.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: 10.18637/jss.v067.i01.

Matic Pavlič & Arthur Stepanov

Belletti, Adriana, Elisa Bennati & Antonella Sorace. 2007. Theoretical and developmental issues in the syntax of subjects: Evidence from near-native Italian. *Natural Language & Linguistic Theory* 25(4). 657–689. DOI: 10.1007/s11049-007-9026-9.

Blackwell, Sarah E. 2003. *Implicatures in discourse: The case of Spanish NP anaphora.* Amsterdam: John Benjamins. DOI: 10.1075/pbns.105.

Bock, Kathryn & Kathleen M. Eberhard. 1993. Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes* 8(1). 57–99. DOI: 10.1080/01690969308406949.

Bosch, Peter, Tom Rozario & Yufan Zhao. 2003. Demonstrative pronouns and personal pronouns: German *der* vs *er*. In *Proceedings of the 2003 EACL workshop on the computational treatment of anaphora.* https://aclanthology.org/W03-2609.

Cardinaletti, Anna & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk van Riemsdijk (ed.), *Clitics in the languages of Europe*, 145–234. Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110804010.145.

Carminati, Maria Nella. 2002. *The processing of Italian subject pronouns.* Amherst, MA: University of Massachusetts. (Doctoral dissertation). https://scholarworks.umass.edu/dissertations/AAI3039345.

Carminati, Maria Nella. 2005. Processing reflexes of the feature hierarchy (Person > Number > Gender) and implications for linguistic theory. *Lingua* 115(3). 259–285. DOI: 10.1016/j.lingua.2003.10.006.

Chomsky, Noam. 1981. *Lectures on government and binding: The Pisa lectures.* Dordrecht: Foris.

Chow, Wing-Yee, Shevaun Lewis & Colin Phillips. 2014. Immediate sensitivity to structural constraints in pronoun resolution. *Frontiers in Psychology* 5. 630. DOI: 10.3389/fpsyg.2014.00630.

Cowart, Wayne & Helen S. Cairns. 1987. Evidence for an anaphoric mechanism within syntactic processing. *Memory and Cognition* 15. 318–331. DOI: 10.3758/BF03197034.

Crain, Stephen & Janet D. Fodor. 1985. How can grammars help parsers? In David R. Dowty, Lauri Karttunen & Arnold M. Zwicky (eds.), *Natural language parsing. Psychological, computational, and theoretical perspectives*, 94–128. Cambridge: Cambridge University Press.

DeLong, Katherine A., Laura Quante & Marta Kutas. 2014. Predictability, plausibility, and two late ERP positivities during written sentence comprehension. *Neuropsychologia* 61. 150–162. DOI: 10.1016/j.neuropsychologia.2014.06.016.

Delong, Katherine A., Thomas P. Urbach, David M. Groppe & Marta Kutas. 2011. Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology* 48(9). 1203–1207. DOI: 10.1111/j.1469-8986.2011.01199.x.

Fedele, Emily & Elsi Kaiser. 2014. Looking back and looking forward: Anaphora and cataphora in Italian. In Jennifer Arlin (ed.), *University of Pennsylvania Working Papers in Linguistics 20(1)*, 79–90. Philadelphia, PN: Penn Graduate Linguistics Society. https://repository.upenn.edu/pwpl/vol20/iss1/10.

Federmeier, Kara D., Edward W. Wlotko, Esmeralda De Ochoa-Dewald & Marta Kutas. 2007. Multiple effects of sentential constraint on word processing. *Brain Research* 1146. 75–84. DOI: 10.1016/j.brainres.2006.06.101.

Frazier, Lyn & Charles Clifton. 1989. Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4(2). 93–126. DOI: 10 . 1080 / 01690968908406359.

Frazier, Lyn & Giovanni B Flores d'Arcais. 1989. Filler driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language* 28(3). 331–344. DOI: 10.1016/0749-596X(89)90037-5.

Garnsey, Susan M., Michael K. Tanenhaus & Robert M. Chapman. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18. 51–60. DOI: 10.1007/BF01069046.

Gouvea, Ana C., Colin Phillips, Nina Kazanina & David Poeppel. 2010. The linguistic processes underlying the P600. *Language and Cognitive Processes* 25(2). 149–188. DOI: 10.1080/01690960902965951.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.

Hagoort, Peter & Colin M. Brown. 1999. Gender electrified: ERP evidence on the syntactic nature of gender processing. *Journal of Psycholinguistic Research* 28. 715–728. DOI: 10.1023/A:1023277213129.

Just, Marcel A., Patricia A. Carpenter & Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111(2). 228–238. DOI: 10.1037/0096-3445.111.2.228.

Kaan, Edith, Anthony Harris, Edward Gibson & Phillip Holcomb. 2000. The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15(2). 159–201. DOI: 10.1080/01690600386084.

Kazanina, Nina. 2005. *The acquisition and processing of backwards anaphora.* College Park, MD: University of Maryland. (Doctoral dissertation).

Kazanina, Nina, Ellen F. Lau, Moti Lieberman, Masaya Yoshida & Colin Phillips. 2007. The effect of syntactic constraints on the processing of backwards

Matic Pavlič & Arthur Stepanov

anaphora. *Journal of Memory and Language* 56(3). 384–409. DOI: 10.1016/j.jml.2006.09.003.

Kazanina, Nina & Colin Phillips. 2010. Differential effects of constraints in the processing of Russian cataphora. *Quarterly Journal of Experimental Psychology* 63(2). 371–400. DOI: 10.1080/17470210902974120.

Kennison, Shelia M., Elaine C. Fernandez & J. Michael Bowers. 2009. Processing differences for anaphoric and cataphoric pronouns: Implications for theories of discourse processing. *Discourse Processes* 46(1). 25–45. DOI: 10.1080/01638530802359145.

Lee, Ming-Wei. 2004. Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research* 33(1). 51–73. DOI: 10.1023/B:JOPR.0000010514.50468.30.

Lobo, Maria & Carolina Silva. 2016. Ambiguidade pronominal em orações adverbiais do português europeu: crianças vs. adultos. *Revista da Associação Portuguesa de Linguística* 2. 319–338. DOI: 10.2-9077/rapl2a14.

Lust, Barbara (ed.). 1986. *Studies in the acquisition of anaphora: Defining the constraints*. Dordrecht: Reidel.

Mancini, Simona, Francesca Postiglione, Alessandro Laudanna & Luigi Rizzi. 2014. On the person-number distinction: Subject-verb agreement processing in Italian. *Lingua* 146. 28–38. DOI: 10.1016/j.lingua.2014.04.014.

Matuschek, Hannes, Reinhold Kliegl, Shravan Vasishth, Harald Baayen & Douglas Bates. 2017. Balancing type I error and power in linear mixed models. *Journal of Memory and Language* 94. 305–315. DOI: 10.1016/j.jml.2017.01.001.

Osterhout, Lee & Phillip J Holcomb. 1992. Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language* 31(6). 785–806. DOI: 10.1016/0749-596X(92)90039-Z.

Osterhout, Lee & Linda A. Mobley. 1995. Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language* 34(6). 739–773. DOI: 10.1006/jmla.1995.1033.

Pablos, Leticia, Jenny Doetjes & Lisa L.-S. Cheng. 2018. Backward dependencies and in-situ wh-questions as test cases on how to approach experimental linguistics research that pursues theoretical linguistics questions. *Frontiers in Psychology* 8. DOI: 10.3389/fpsyg.2017.02237.

Pablos, Leticia, Jenny Doetjes, Bobby Ruijgrok & Lisa L.-S. Cheng. 2015. Active search for antecedents in cataphoric pronoun resolution. *Frontiers in Psychology* 6. DOI: 10.3389/fpsyg.2015.01638.

R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. London: Croom Helm.

Reinhart, Tanya. 1986. Center and periphery in the grammar of anaphora. In Barbara Lust (ed.), *Studies in the acquisition of anaphora: Defining the constraints*, 123–150. Dordrecht: Reidel. DOI: 10.1007/978-94-009-4548-7_3.

Sauerland, Uli. 2003. A new semantics for number. In Robert B. Young & Yuping Zhou (eds.), *SALT 13: Proceedings of the 13th Semantics and Linguistic Theory Conference*, 258–275. Ithaca, NY: CLC Publications. DOI: 10.3765/salt.v13i0.2898.

Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives*, 413–434. Berlin: De Gruyter. DOI: 10.1515/9783110197549.413.

Silverstein, Michael. 1985. Hierarchy of features and ergativity. In Pieter Muysken & Henk van Riemsdijk (eds.), *Features and projections*, 163–232. Dordrecht: Foris.

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, 243–281. London: Palgrave-Macmillan. DOI: 10.1057/9780230210752_9.

Stepanov, Artur & Penka Stateva. 2015. Cross-linguistic evidence for memory storage costs in filler-gap dependencies with wh-adjuncts. *Frontiers in Psychology* 6. DOI: 10.3389/fpsyg.2015.01301.

Stowe, Laurie A. 1986. Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1(3). 227–245. DOI: 10.1080/01690968608407062.

Sturt, Patrick. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48(3). 542–562. DOI: 10.1016/S0749-596X(02)00536-3.

Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock & Francesca Filiaci. 2004. First language attrition and syntactic subjects: A study of Greek and Italian near-native speakers of English. *International Journal of Bilingualism* 8(3). 257–277. DOI: 10.1177/13670069040080030601.

Van Gompel, Roger P. G. & Simon P. Liversedge. 2003. The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29(1). 128–139. DOI: 10.1037/0278-7393.29.1.128.

Matic Pavlič & Arthur Stepanov

Van Petten, Cyma & Barbara J. Luka. 2006. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language* 97(3). 279–293. DOI: 10.1016/j.bandl.2005.11.003.

Wilson, Frances, Antonella Sorace & Frank Keller. 2009. Antecedent preferences for anaphoric demonstratives in L2 German. In Jane Chandlee, Michelle Franchini, Sandy Lord & Gudrun-Marion Rheiner (eds.), *Proceedings of the 33rd annual Boston University Conference on Language Development*, vol. 2, 634–645. Somerville, MA: Cascadilla Press.

Yoshida, Masaya, Nina Kazanina, Leticia Pablos & Patrick Sturt. 2014. On the origin of islands. *Language, Cognition and Neuroscience* 29(7). 761–770. DOI: 10.1080/01690965.2013.788196.

Yulong, Xu & H. Xiaodan. 2007. Cataphora in chinese and english texts: a functional pragmatic analysis. *Foreign Language Teaching and Research* 39(6). 417–423.

Zhiyi, Zhang. 2019. Pronoun anaphor and cataphor in Chinese. *Macrolinguistics* 7(11). 30–47. DOI: 10.26478/ja2019.7.11.2.

Chapter 13

On the grammatical integration of *n/t*-participles of imperfective stems in Polish and Russian

Björn Wiemer, Joanna Wrzesień-Kwiatkowska & Alexander Rostovtsev-Popiel

Johannes Gutenberg University Mainz

The article presents a critical discussion of recent work on the semantics of lexical prefixes and of the aspect meanings of *n/t*-participles of imperfective stems in contemporary Russian and Polish, and on the role of all these formations in the voice system of both languages. On this background, a corpus-based study on the development of the aspect functions of these participles for imperfective and perfective stems in Russian and Polish from 1730 until today is discussed, including their syntactic distribution (predicative, appositive, attributive use) and the role of secondary imperfective stems. Special attention is paid to coarse measures of productivity and the changing relation between type and token frequency. This study can be considered the first usage-based investigation from a diachronic perspective in Slavic linguistics, which, to a large extent, is made possible thanks to a database of aspect triplets.

Keywords: Russian, Polish, aspect, voice, *n/t*-participles, diachronic morphology, corpora

1 Introduction

The aspect system of Slavic languages is based on a binary distinction between perfective (pfv.) and imperfective (ipfv.) stems. These stems are related not only lexically, but also, on average, morphologically on the basis of productive and commonplace derivational patterns (see §3.1). The stems gain their pfv. and ipfv.

status, respectively, from their complementary distribution over sets of contexts, or conditions, which can be defined grammatically (e.g., co-occurrence restrictions with tense markers, with phasal verbs or with modal auxiliaries) or pragmatically (e.g., triggering of presuppositions, type of illocution); cf. Wiemer (2008), Wiemer & Seržant (2017: 243–255), Wiemer et al. (2020: §§2–3). All finite and non-finite verb forms are derived from these stems, including participles marked with an *n/t*-suffix (e.g., Russ. *obrabota-n-a* 'worked out', Pol. *podję-t-y* 'taken up'). These participles are employed in every Slavic language, however Slavic languages differ as for the degree to which these participles are restricted by aspect and how integrated they are into the voice system (and its intersections with perfects).[1] Russian, in particular, demonstrates considerable restrictions on *n/t*-participles from ipfv. stems, and one wonders which role they might play in passive constructions. Polish, to the contrary, has tightly integrated *n/t*-participles of either aspect into its voice system. In fact, after the late 18th century (if not earlier) the role played by ipfv. *n/t*-participles in Russian and Polish has developed in radically different ways.

This investigation is an attempt at opening a window into this divergent development. Simultaneously, it demonstrates how usage-based accounts should complement formal semantic approaches, mainly because such accounts concentrate not on model-theoretic assumptions, but on distributional patterns. Although the focus of this study is on diachrony, namely the time from 1730 up to now, we will first survey some recent findings and claims about ipfv. participles, particularly in Russian (§2.1). This will lead us to some questions (§2.2) and provide a point of departure for a corpus-based study on the functional development and productivity of *n/t*-participles in Russian and Polish. The study is connected to a database of aspect triplets (§3). After a discussion of findings (§4) some conclusions will be drawn (§5). An Appendix accessible under https://zenodo.org/record/6602167#.YqBH0OzP2Un contains tables with more detailed information on the statistical figures referred to below. The glosses will consistently distinguish between unprefixed, or simplex, ipfv. stems (IPFV1) and ipfv. stems derived via suffixation from a prefixed pfv. stem ("secondary imperfectives": IPFV2).[2] Examples cited from corpus samples will only be provided with a general indication of the source and the year or time interval. Examples without an indication of source are constructed by an informed native speaker.

---

[1]Cf. Wiemer & Giger (2005), Wiemer (2017), Arkadiev & Wiemer (2020).

[2]The glossing also indicates zero-marked categories.

2 On the status of ipfv. participles

Our considerations set out from two recent accounts of ipfv. *n/t*-participles in Russian and their involved theoretical repercussions.

2.1 Recent accounts concerning Russian ipfv. *n/t*-participles

On the basis of a query from the Russian National Corpus (RNC; https://ruscorpora. ru/new/), Borik & Gehrke (2018) found that, in Russian, *n/t*-participles of ipfv. verbs cannot be dismissed as rare or haphazard, nor are they *in toto* to be characterized as lexicalized forms (adjectives).[3] Instead, their meanings are often compositional and they do occur in constructions that can only be analyzed as true, i.e. event-oriented passives. Compare the following examples, with the (a)-examples containing ipfv. *n/t*-participles, the (b)-examples their active equivalents (the (a)-examples are cited after Borik & Gehrke 2018: pp. 61, 66, 65, respectively):[4]

(1) a. My   oba   *by-l-i*   *striže-n-y*   nagolo.
         we.NOM both-M.NOM be-PST-PL cut.hair.IPFV1-PP-PL naked.ADV
         'We both had our hair *cut off*.'

     b. Nas   obo-ix   *strig-l-i*   nagolo.
         we.ACC both-M.ACC cut.hair.IPFV1-PST-PL naked.ADV
         'They *cut* our hair *off*.'

(2) a. [Ne raz ja   *by-l*   *uče-n*],   molču i znaju.
          NEG once I.NOM be-PST-SG.M teach.IPFV1-PP-SG.M
         '[I was *taught* more than once], I keep silent and know.'

     b. Ne raz   menja *uči-l-i*.
         NEG once I.ACC   teach.IPFV1-PST-PL
         'They have *taught* me more than once.'

(3) a. *Pisa-n-o*   ėto by-l-o   Dostoevsk-im v 1871 god-u.
         write.IPFV1-PP-SG.N this be-PST-SG.N PN-INS   in   year-LOC
         'This *was written* by Dostoevskij in 1871.'

     b. *Pisa-l*   ėto Dostoevsk-ij v 1871 god-u.
         write.IPFV1-PST-SG.M this PN-NOM   in   year-LOC
         'Dostoevskij *wrote* this in 1871.'

---

[3]In practice, *verb* equals *stem*, if not indicated otherwise.

[4]Due to the extended length of some (corpus) examples, we do not always gloss the whole example. In case only a proper subpart of an example is glossed (typically a clause), that part is surrounded by square brackets and is matched by a bracketed part in the translation.

An event-oriented use (or eventive orientation) can hardly be denied for (2) and (3).[5] In (2) this orientation is supported by an adverbial which marks the situation as repeated; a resultant state is only implied, it is indicated in the subsequent clause ('people taught me ⤳ I know (because of that)'). In (3) the event is presupposed, while the communicative focus is on the adverbial ('in 1871') which puts this event in a larger time frame. By contrast, an eventive orientation is more difficult to get for (1), rather it refers to the state that results after the hair cut was completed.

Borik & Gehrke (2018) point out that Russian ipfv. *n/t*-participles seem to be restricted to GENERAL-FACTUAL meanings; in general, progressive readings are practically unattested (cf. also Knjazev 1989: 57f.; Knjazev 2007: 489). Since the general-factual (GF) meaning is considered dominant, one wonders how GF relates to event-oriented uses of ipfv. *n/t*-participles. We should be aware that the label "general-factual" unites at least two rather different main functions, called PRESUPPOSITIONAL and EXISTENTIAL, and that in the discussion about ipfv. *n/t*-participles the presuppositional GF clearly dominates.[6] Thus, Borik & Gehrke demonstrate that ipfv. *n/t*-participles are comparable to definite descriptions, since they anaphorically refer to known, or presupposed, situations (eventualities), and these can be events; see (4). The same can be said for GF in the active voice; see (5). The parts in curly brackets contain the notional antecedents of the "anaphoric" verb forms (in italics).

(4) čto kasaetsja {platy deneg},
   [to *plač-e-ny*   *by-l-i*   naličnymi šest' tysjač   rublej].
   PTC pay.IPFV1-PP-PL be-PST-PL cash   six   thousands rubels
   'As for the payment, [six thousand rubles *were paid* in cash].'

<div align="right">(cited from Borik & Gehrke 2018: 70)</div>

(5) V ètoj porternoj {ja napisal pervoe ljubovnoe pis'mo}.
   [*Pisa-l*   karandaš-om.]
   write.IPFV1-PST-SG.M pencil.M-INS.SG
   'In this tavern, I wrote my first love letter. [I *wrote* it with a pencil].'

---

[5]One could speak of "eventive focus" as well. However, the term "focus" occurs in two different, though related senses. It either refers to the asserted part of an event structure (as in (2)–(3)), or it refers to the comment as part of the information structure of an utterance. It is hardly possible to distinguish these two senses by short circumscriptions or different synonyms, and their relation becomes clear when we realize that time adverbials which gain informational focus can be (and often are) employed as means to test the asserted part of some tense-aspect marker: such diagnostics relies on the "harmony" between both kinds of focus in an utterance. Hopefully, in the remainder the respective context will disambiguate the intended sense.

[6]For detailed analyses of GF cf. Grønn (2004), Mehlig (2011), Dickey (2015), Mueller-Reichau (2018).

(cited from Borik & Gehrke 2018: 64)

Here, we should however take into account that anaphoric use (with event modifiers) is characteristic only of presuppositional GF (see (4)–(6)), not of the existential (or verifying) type, to which (2) comes close. Apart from that, at least in contemporary Russian, ipfv. *n/t*-participles in existential GF are very rare unless they occur under negation. Without negation they sound archaic (see (7), constructed); (8) is one of the few examples without negation found in the RNC. Note that, apart from existential readings, in such cases *n/t*-participles from ipfv. stems (as well as their finite forms) may also be interpreted as referring to a repeated event:

(6) Knig-a   pečata-n-a   pri Petre Pervom.
    book.F-NOM.SG print.IPFV1-PP-SG.F at   Peter First
    'The book was *printed* under Peter the Great.'
    → presuppositional, narrow scope: temporal location possible

(7) U vas   byl-i   peče-n-y   pirogi?
    at we.GEN be-PST-PL bake.IPFV1-PP-PL pie-NOM.PL
    'Did you bake pies?'
    (more lit.: 'Did you have pies *baked*?') (> can also refer to habits)
    → existential, verum focus (wide scope): no temporal location

(8) U odnogo korolja byl šut. V junosti pošučival na svoj strax i risk na
    ploščadjax i
    [*by-l*   *poro-t*],   vsledstvie čego poumnel.
     be-PST.SG.M flog.IPFV-PP-SG.M
    'One king had a jester. In his youth, he (the jester) joked at his own peril
    and risk in the squares and [*was flogged*], as a result of which he grew
    wiser.'   (Russian; RNC; 2000)
    → existential, wide scope: no temporal location

Contrary to the presuppositional type, in existential GF the eventuality constitutes the informational focus (indicated by stress, e.g. on *byli* in 7), i.e. the part which is unknown and which can be asked about (see Footnote 5). Jointly with this, it does not matter whether this eventuality took place once or more than once.[7] The fact that the concrete temporal location is not at stake explains the just mentioned "oscillation" with habitual readings. Therefore, what existential

---

[7] This brings the existential GF close to the experiential function of perfects known from typological research (cf. Arkadiev & Wiemer 2020).

and presuppositional GF unites is the downgrading of the eventuality denoted by the VP, although this happens for diametrically opposed reasons: with presuppositional GF, downgrading concerns the information structure (since the eventuality is already known to have taken place within a specific reference interval), whereas with existential GF, downgrading concerns time location (there is no discrete interval for which it may be claimed true that the eventuality occurred, or this is irrelevant). What follows from this is that even if the predicate refers to a distinct single event, this (or any other) actionality feature is assigned background status. However, since presuppositional GF is associated to time-located events, this location can be targeted, e.g. by temporal adverbials (see (3) and (6)).

Similarly, Borik & Gehrke (2018: 59) characterize cases like (1) as adjectival participles: "unlike with verbal passives, the underlying event in adjectival passives lacks spatiotemporal location or referential event participants, and only the state associated with the adjectival participle can be located temporally". Verbal participles, in turn, can have "spatiotemporal event modifiers, referential by-/with-phrases, and similar such expressions" which highlight the event (Borik & Gehrke 2018: 59). However, adjectival participles need not be lexicalized.

Borik & Gehrke also point out that Russian ipfv. *n/t*-participles in true passives derive from a restricted set of verbs, most of them related to speech acts or with incremental objects. That is, in comparison to compositional pfv. *n/t*-participles, their overall type and token frequency in passives seems to be low, after all. Moreover, only simplex (IPFV1) stems are used, while ipfv. stems derived via suffixation from a prefixed pfv. stem (IPFV2) are absent in modern Russian. While these claims are largely supported by our findings, we will show that some of them require qualification when we look at them from a usage-based perspective (see §4).

In turn, Tatevosov (2015: 288–292) employs the behavior of Russian *n/t*-participles of IPFV1 stems as support for his claim that lexical (or "inner") prefixes add resultative subevents, while IPFV1 stems are void of this component. Compare the following examples with their logical structures, in which the subscripts A and S indicate an action and a state, respectively:

(9) a. Vanj-a   pisa-l   (stat'j-u).
        PN.M-NOM.SG write.IPFV1-PST-SG.M   article.F-ACC.SG
        'Vanja wrote / was writing (an article).'
        $[\![\text{pisa}]\!] = \lambda y \lambda x \lambda e \, [\text{WRITE}_A(e) \wedge \text{INITIATOR}(x)(e) \wedge \text{THEME}(y)(e)]$

     b. Vanj-a   na-pisa-l   *(stat'j-u).
        PN.M-NOM.SG PVB-write.PFV-PST-SG.M   article.F-ACC.SG
        'Vanja wrote / has written an article.'

$$\llbracket\text{napisa}\rrbracket = \lambda y \lambda x \lambda e \lambda s [\text{WRITE}_A(e) \wedge \text{INITIATOR}(x)(e) \wedge \text{THEME}(y)(e) \wedge$$
$$\text{CAUSING}(s)(e) \wedge \text{WRITE}_S(s) \wedge \text{ARG}(y)(s)]$$

Notably, Tatevosov treats as lexical prefixes not only those which modify, or change, the lexical meaning of the IPFV1 stem (as in Russ. *rabotat' (*den'gi)* (intended:) 'work (money)' → *za-rabotat' *(den'gi)* 'earn money'), but also so-called natural prefixes (Janda 2007), whose function overlaps with a meaning component implied by the simplex (e.g., Russ. *varit'* → *s-varit' *(sup)* 'cook (soup)', *delit'* → *raz-delit' *(gruppu)* 'divide (group)'). Natural prefixes are a precondition for the rise of aspect triplets (see §3).

Since *n/t*-participles are derived from these stems, they should also show behavior that ensues from the presence vs. absence of a resultative subevent. In fact, IPFV1 *n/t*-participles in passives usually require modifiers that relate to the event, not a subsequent state (as confirmed by Borik & Gehrke 2018, see above); compare Russ. *Pis'mo pisano* *(na tonkoj bumage)* 'The letter is written *(on thin paper)' vs. *Pis'mo napisano* (*i ležit na stole*) 'The letter has been [lit. is] written (and is lying on the table)'. Tatevosov takes this as evidence that IPFV1 stems, and with them their *n/t*-participles, lack a resultative subevent.[8] Simultaneously, he points out that ipfv. *n/t*-participles are unable to denote not only ongoing processes, but even habitual situations. Thus, the only reading "left" for them is general-factual meanings (Tatevosov 2015: 291).

The conclusion concerning GF is congruent with the analysis by Borik & Gehrke (2018), but it raises the question why certain ipfv. *n/t*-participles in Russian prefer stative readings; see (1) and the following example:

(10) Pol   *by-l*   *mošče-n*   širok-imi
      floor.M-NOM.SG be-PST-SG.M pave.IPFV1-PP-SG.M wide-INS.PL
      serovat-ymi   kamnj-ami.
      grayish-INS.PL stone-INS.PL
      'The floor *was paved* with wide grayish stones.'   (Tatevosov 2015: 292)

Tatevosov declares *moščen* 'paved' to be an adjectivized participle, so that the explanation would be the same as by Borik & Gehrke (2018) for adjectival passives (see above), in particular we understand why this form yields the same aspectual semantics as does its pfv. counterpart (*vymoščen*), but, contrary to the latter, cannot be used with a focus on the event itself. Note that this holds independently from the distinction between existential and presuppositional GF. Thus,

---

[8]Here we need not take stance as for Tatevosov's subsequent claim that (pfv.) aspect is assigned above *v*P and not a property of the verb stem (cf. Wiemer 2019: 107–110 for discussion).

(10) could be uttered in continuation, e.g., of *Oni vošli v ogromnyj zal* 'They entered a huge hall', on which (10) would add information concerning a salient part of the newly introduced referent *zal* 'hall'; but this information only would refer to the state of the floor, not to an event of paving it. Consider also cases like *Rab byl porot*[IPFV1] / *vyporot*[PFV] *triždy* 'The slave was flogged three times'. However, there are other cases of ipfv. *n/t*-participles which cannot be explained away as adjectives (or adjectival passives); see the examples and discussion in §4. Again, the question arises as to how an eventive orientation relates to GF.

Gehrke (2023) argues that (finite or participial) forms of ipfv. stems may be used with reference to concrete single events if these events have been mentioned in, or can be inferred from, the immediately preceding discourse. Following Gehrke's suggestion, we should realize that anaphoric relations to events sometimes need support by metonymic relations (between parts of events that have been mentioned and those which have not) to be inferred. In addition to Gehrke, one wonders whether it is necessary to assume an eventive component in the semantic description of forms of ipfv. stems. However, Gehrke rightly criticizes formal accounts of GF for having put too strong an emphasis on event completion; completion is neither a necessary nor a sufficient condition for choice of pfv. aspect, nor for the exclusion of ipfv. aspect.

As a matter of fact, (non-)completion is not a constitutive property of (im)perfective aspect; instead, the crucial criterion is (non-)boundedness, or whether an eventuality is presented as limited or not (cf. Lehmann 1999, Wiemer 2017, Wiemer & Seržant 2017, Breu 2021, among many others). We therefore support both points made by Gehrke. However, again, her argument is based on the presuppositional type of GF, leaving open how it might work for other usage conditions of ipfv. aspect. The bulk of examples from our study that may be classified as GF do not represent its presuppositional type, and there is much leeway in categorizing these examples anyway (see §3–§4).

Finally, there is one issue left concerning Tatevosov's (2015) analysis. Namely, why should forms of stems for which resultative subevents are lacking be incapable of denoting ongoing processes (progressive meaning) or habitual situations, that is functions typically associated to ipfv. aspect? First, a resultative subevent presupposes a change of state, and this is entailed by pfv. stems which contribute to a telic meaning of the clause (see (9)). There is thus no inherent reason why the lack of a resultative subevent should block the denotation of an activity for which boundaries are absent, or defocused (and, thus, progressive meaning). Second, pluractional meanings, like the habitual one, are insensitive to actionality distinctions (± telic; process, event, state).[9] Compare *He used to sleep*

---

[9]Cf. Tatevosov (2016: 118) and the literature referred to in Footnote 14.

after dinner (→ habitual state), *During our discussions she used to remark that P*
(→ habitual event), *Whenever I met them in the club, they used to be discussing*
the latest soccer game (→ habitual process). This becomes particularly obvious
in Russian, where ipfv. stems – both IPFV1 and IPFV2 – are usually employed
as "placeholders" of their pfv. counterparts in event readings that focus on the
attainment of a goal (= right boundary). Compare habitual readings of events
which, each time they occurred, reached their culmination point:

(11) Po utram   direktor   *vyzvaniva-l*   vsex zamov i
      on mornings director.M-SG.M call.out.IPFV2-PST-SG.M all   deputies and
      raspredelja-l   meždu nimi zadači na den'.
      distribute.IPFV2-PST-SG.M among them tasks   on day
      'In the morning, the director *used to call* all the deputies and *distribute*
      among them tasks for the day.'

(12) Kogda mne bylo pjat' let,
      [ja   každyj den' na zavtrak *s"eda-l*   tri   jabloka].
       I.NOM each   day on breakfest eat.IPFV2-PST-SG.M three apples
      'When I was five years old, [I *ate* three apples every day for breakfast].'

(13) Na vyxodnye on   vsem   k času *gotovi-l*
      on free.days he.NOM everybody-DAT.PL to hour prepare.IPFV1-PST-SG.M
      obed.
      dinner-ACC
      'On weekends, he *cooked* dinner for everyone by one o'clock.'

This ability to function as grammatical equivalents of pfv. stems in denoting
completed events can only be explained if we assume that ipfv. stems can ac-
quire properties of their pfv. counterparts. The question is to which extent this
carries over to their participles. Therefore, even if it turned out true that ipfv.
n/t-participles are incapable of denoting habitual situations, this could hardly be
explained from model-theoretic assumptions and other premises accepted by Tat-
evosov. After all, it should be checked to which extent this claim is empirically
adequate.

2.2 Questions

The claims presented by Borik & Gehrke (2018) and Tatevosov (2015) generate
some questions. First of all, Tatevosov's morpheme-centric generative analysis
would imply that the additional subevent remains with IPFV2 stems, since these

suffixed stems are derived from prefixed (pfv.) stems. Tatevosov does not consider IPFV2, certainly because in contemporary standard Russian *n/t*-participles of IPFV2 stems practically do not exist. Thus, one wonders which consequences are to follow for *n/t*-participles of IPFV2 stems in passives if they do occur. First of all, shouldn't habitual readings be compatible, if not preferred, for the reasons indicated in §2.1?

This can be tested for Polish, where, as in other West Slavic languages, *n/t*-participles of ipfv. stems are commonplace (Wiemer 2017: 135–138). Polish has completely integrated *n/t*-participles of both IPFV1 and IPFV2 stems into the aspect system and its interface with voice (Lehmann 1992, Wiemer 1996, Górski 2008). However, we do not know much about their productivity and function range, first of all, in a diachronic perspective. Moreover, one may ask to which extent the ability of ipfv. *n/t*-participles to function as full-fledged members on the aspect-voice interface correlates with the overall frequencies (for all grammatical forms) of their stems (see §4.2).

Furthermore, GF is insensitive to actionality features as well, but if GF relates to an event this may entail a result (provided the event is telic); see examples in §2.1. The point is not whether a result has ensued, but whether it is treated as the asserted part of the message or as presupposed information. That is, in accordance with Borik & Gehrke (2018) and Gehrke (2023), GF should be evaluated not so much with respect to the internal structure of events, but in terms of information structure, and (deictic or relative) time location may, but need not, become an issue (see (3), (6)). Moreover, the lexical (and natural) prefixes which, according to Tatevosov, "bind" a resultative subevent can be taken as means that establish this subevent as an undeniable part of the verb's meaning. However, this does not imply that the respective IPFV1 stems exclude a resultative subevent. Provided they occur (as VP heads) in suitable clausal contexts, they may just be able to defocus such a subevent; that is, they are labile in this respect. Otherwise, how would we explain the employment of ipfv. stems (finite forms or *n/t*-participles) in GF which evidently refer to an event (e.g., Pol. *Już byłem o to pytany*[IPFV1] 'I have already been asked about that', Russ. *Menja ob ėtom uže sprašivali*[IPFV2] 'They already asked me about that'), and their employment as replacements of pfv. verbs in the denotation of events, e.g., in the narrative present? Apart from that, it is justified to ask for stative readings of ipfv. *n/t*-participles in the contemporary and earlier stages; keeping in mind that stative readings do not automatically indicate that participles have lexicalized as adjectives (see Borik & Gehrke 2018 in §2.1).

3 Further premises and the data used for the study

Related empirical questions are addressed in the following. We present findings concerning the aspectual behavior of *n/t*-participles from IPFV1 and IPFV2 stems, primarily on the basis of a comprehensive database containing potential aspect triplets (e.g., Pol. *tworzyć*[IPFV1] – *stworzyć*[PFV] – *stwarzać*[IPFV2] 'create', Rus. *paxat'*[IPFV1] – *vspaxat'*[PFV] – *vspaxivat'*[IPFV2] 'plough'), which covers the period 1750–2018 in Russian and Polish. Triplets have the advantage that the meanings and behavior of IPFV1 and IPFV2 can be compared directly. We first comment on triplets and our database (§3.1) before we turn to the sampling procedure (§3.2) and the annotation schema (§3.3).

3.1 Triplets

In connection with the *DiAsPol*-project, a database of aspect triplets for the period 1750–2018 has been created for Polish, Czech, and Russian.[10] Aspect triplets (or, more strictly, "bi-imperfective aspect triplets", see Zaliznjak et al. 2015: 235f. henceforth simply TRIPLETS) are built on a constellation in which two ipfv. stems lexically correspond to the same cognate pfv. stem: one ipfv. stem is derived from the pfv. stem by a suffix (= IPFV2), the other is an unprefixed ipfv. stem (= IPFV1, or simplex) and itself the morphological basis for the pfv. stem. Compare, for instance, the following illustrations for Russian and Polish.[11]

(14)   IPFV1   PFV   IPFV2
         a. *gre-t'*   → *na-gre-t'*   → *na-gre-va-t'*   'warm up'
         b. *gotov-i-t'*   → *pri-gotov-i-t'*   → *pri-gotavl-iva-t'* 'prepare (meal)'

(15)   IPFV1   PFV   IPFV2
         a. *dzieli-ć*   → *roz-dziel-i-ć*   → *roz-dziel-a-ć*   'divide; separate'
         b. *kaz-a-ć*   → *na-kaz-a-ć*   → *na-kaz-ywa-ć*   'order'

Triplets result from an overlay of the two most productive patterns by which aspect pairs are created in Slavic languages, namely (for Russian):

---

[10]See https://www.diaspol.uw.edu.pl/eng/. A detailed description of the database is underway.
[11]Suffixation may consist either in an addition (as with {va} in the case of *na-gre-va-t'*), or in a replacement (as with {i} > {a} in the case of *roz-dziel-a-ć*), or in a replacement that resulted from the coalescence of two more elementary segments (as with {iva} < {i-va} in the cases of *pri-gotavl-iva-t'* and *na-kaz-ywa-ć*). These distinctions are irrelevant for the present concern.

Wiemer, Wrzesień-Kwiatkowska & Rostovtsev-Popiel

(16)   IPFV1   PFV   IPFV2
  a. *pis-a-t'* → *na-pis-a-t'*   'write'
  b.   *pere-pis-a-t'* → *pere-pis-yva-t'*   'rewrite'

Another precondition is that the derivation IPFV1 → PFV involves a natural pre-fix, so that no lexical change (or modification) obtains (see §2.1). In triplets, both IPFV1 and IPFV2 function as lexical replacements of PFV in grammatically or pragmatically defined contexts, well-known to Slavic aspectology, and IPFV1 and IPFV2 may also replace one another. Admittedly, in many cases only IPFV2 is considered an exact "lexical copy" of its pfv. counterpart, which also shares the argument requirements of the latter, while IPFV1 stems betray less strict require-ments (usually for objects) and are lexically more diffuse than the remaining pair of prefixed PFV and IPFV2. However, in many cases there is no IPFV2 – or it is derived only occasionally and not considered part of the standard language – and IPFV1 alone "fulfills the duties" of the PFV's lexical copy, as in the case of *pisat'* – *napisat'*.

Table 1 provides the number of items of which the Russian and the Polish triplet database is composed.

Table 1: Sizes of aspect triplet database

| Polish | 1,773 triplets | – | 1,386 (IPFV1), 1,773 (PFV), 1,807 (IPFV2) |
| Russian | 1,275 triplets | – | 837 (IPFV1), 1,275 (PFV), 1,461 (IPFV2) |

There are less IPFV1 than PFV stems because many IPFV1 stems enter into more than one triplet (with different prefixes), and the number of IPFV2 stems is larger than for PFV stems since there happen to be suffix variants. In this study, the latter are neglected, and the number of triplets per period coincides with the number of PFV stems.

3.2 Sampling procedure

For the participle study we established five subperiods: 1730–1800, 1801–1850, 1890–1918, 1945–1980, 1990–2020. The size of the available corpora (see Refer-ences) and the periods differed, partially quite considerably (see Table 2). For each period, we drew random samples à 100 tokens of *n/t*-participles of ipfv. and pfv. stems. Not always was this mark reached because of the corpus size, and some of the Russian samples were slightly larger (see Appendix, Part I). A sample

Table 2: Sizes of subcorpora (in tokens of expressions) for periods

	1730–1800	1801–1850	1890–1918	1945–1980	1990–2020
Russian	3,227,827	14,716,129	46,943,219	51,398,237	114,020,580
Polish	3,478,168	1,837,844	776,277	9,100,510	284,820,841

contained no more than 15 items of the same stem; on this account, the randomizing procedure could be violated.

The sampling procedure did not distinguish between the syntactic function, thus predicative participles, e.g. (17), had the same chance to get into a sample as did participles used as NP-modifiers ("attributive") or in appositions ("semi-predicative"). The percentage of predicative use per sample varies a lot, but it rarely approaches, or exceeds, 50 percent (see Appendix, Part I). However, we will not dwell on biases in syntactic use, apart from correlations between syntactic use and aspect functions (see §4.3).

Appositive use subsumes cases in which the participle constitutes a clause on its own. Of course, the distinction from, respectively, predicative and attributive use is often troublesome, but, as a rule, appositive use can be distinguished on the following criteria: (a) the participle follows its head noun; (b) it does not denote any stable (inherent) feature (as do characterizing adjectives and participles, usually lexicalized). Compare (18) as opposed to (19):

(17) Otóż t-a   reform-a   nie będzi-e
      PTC this-F.NOM.SG reform.F.-NOM.SG NEG FUT-3SG
      kojarzo-n-a   z   nik-im.
      associate.IPFV1-PP-F.NOM.SG with nobody-INS
      'Well, this reform will not be *associated* with anyone.'   → predicative

(18) Kar-ą   *stosowa-n-ą*   wobec żołnierzy
      punishment.F-INS.SG apply.IPFV1-PP-F.INS.SG towards soldiers
      jest   także areszt   wojskowy.
      be.PRS.3SG also   arrest.M-NOM.SG military-M.NOM.SG
      'Military arrest is also a penalty *applied* to soldiers.'   → appositive

(19) Ogląda-my   wpływ   *budowa-n-ego*   przez
      watch.IPFV-PRS.1PL influence.M-ACC.SG build.IPFV1-PP-N.GEN.SG over
      dziesięciolecia genius loci na now-ych   mieszkańc-ów.
      decades   genius loci on new-ACC.SG inhabitant-ACC.PL
      'We see the impact of the genius loci *built* over decades on new
      inhabitants.'   → attributive
                                                           (Polish)

Wiemer, Wrzesień-Kwiatkowska & Rostovtsev-Popiel

For Polish and Russian, samples were drawn from the respective national corpora (Polish National Corpus/PNC – http://nkjp.pl/; RNC – https://ruscorpora.ru/). The triplet database served as a means to restrict corpus queries by which sets of IPFV1 stems could be compared to sets of IPFV2 stems, i.e. two sets of ipfv. stems with basically identical lexical meaning (see §3.1). In fact, as for Russian, the triplet database turned out to be the only means to get a handle on certain frequency data from the RNC (see §4.1). We did not, however, compare specific IPFV1 stems to their individual IPFV2 counterparts (or vice versa). This would require a series of case studies that we did not intend to perform.

For Polish, the annotation schema and interface of the PNC allowed us to work with a more diversified array of samples than this was possible for the RNC. Thus, for Polish separate series of samples (à five periods) were drawn from the PNC for IPFV1, PFV and IPFV2 stems that were contained in our database, according to the following guidelines:

(i) à 25 stems with the highest token frequencies of n/t-participles (or less if the list contained less than 25 stems in the most frequent group);[12]

(ii) 25 stems selected by chance from among stems with a frequency of 1–5 n/t-participle tokens.

Below these groups are named "freq(uent)" and "infreq(uent)". In addition, for each stem type and each period we composed (iii) an entirely random sample of n/t-participles from just any possible stem (regardless of the frequency of its forms); this sample series served as control.

As for Russian, samples did not distinguish for different frequency ranges of n/t-participles from different stems, since no sufficiently reliable figures required for such a distinction could be obtained from the RNC. We therefore just created (i) random samples of n/t-participles of IPFV1 and PFV stems for all periods based on our database and (ii) analogous random samples independently from our database. In addition, we (iii) "skimmed through" all n/t-participles we could find for IPFV2 stems (46 tokens, of which 41 belong to the period 1730–1800). These were not considered for inferential statistics.

Therefore, we can compare Russian and Polish n/t-participles of IPFV1 and PFV stems. However, while the Polish data, in addition, gives a chance to estimate whether the database-driven samples show any bias in comparison to entirely random samples of n/t-participles, in Russian we can only control for possible biases of n/t-participles of IPFV1 and PFV stems from the database.

---

[12]The most frequent groups were established on the basis of salient frequency cuts (individually for each sample). In many samples there were not enough stems with a participle frequency of > 5.

3.3 Annotation: aspect functions and syntactic functions

All samples were annotated manually for aspect functions, the aforementioned syntactic functions and polarity. Russian examples were additionally annotated for nominal vs pronominal ("short" vs. "long") form. Their proportions can be inferred from the table in Part III of the Appendix; however, their distribution will not be discussed in this article. As concerns aspect functions and syntactic functions, we used the tag "_d" (= doubtful) for the closest acceptable value, since many cases turned out difficult to categorize. Overall, the Russian data contain considerably more cases of doubt in the assignment of aspect function than the Polish data, and a particularly large share of such cases (in both languages) falls on the general-factual function, which often is difficult to distinguish from a habitual or stative function, as defined below (see Parts IV–V of the Appendix). However, even on an account of doubtful cases the general-factual function in Russian begins to predominate over these two other functions only in the later periods (see §4.4.1), and there is no reason to assume that, on average, there was a bias in favor of any of these more frequent functions. In semantic annotation, such decision problems are well known, and we consider it important to mark ambiguity or problematic cases in the original data, as they supply valuable information on the "edges" of categorial distinctions. However, this issue will not be addressed here, either; instead, problematic cases were integrated into counts according to the value which we accepted as the closest one, and we comment on such problems in passing below. All annotations were thoroughly double-checked by an informed and trained native speaker and by B. Wiemer.

As concerns aspect functions of pfv. participles, we made a distinction between eventive (a.k.a. actional) and stative (i.e. resultative) use, since these are two crucial meanings distinguished for passives. For ipfv. participles, since they are more interesting in terms of inner-Slavic differentiation, a more diversified array of values was assumed. All of them are widely applied in aspectology and are briefly commented on here, with illustrations from our samples.

3.3.1 Progressive (PROG)

Situations can consist of phases. A progressive reading focuses on any internal phase(s), so that boundaries are defocused.[13] Note that, in our Russian samples, almost all cases of progressive reading raised doubts, and they need rather strong contextual support (21).

---

[13]In Klein's (1994) terms this means that Topic Time is included in Time of Situation, whereas the eventive meaning implies a limitation and, thus, amounts to the inclusion of Time of Situation in Topic Time.

(20) Poczekać kilkanaście sekund. W tym czasie pokaże się pasek postępu,
[na któr-ym   będ-ą   widoczn-e   *ładowa-n-e*
 on REL-M.LOC.SG FUT-3PL visible-M.NOM.PL load.IPFV1-PP-M.NOM.PL
element-y].
element.M-NOM.PL
'Wait a dozen or so seconds. During this time, the progress bar will show,
[on which the items being *loaded* will be visible].'

(Polish; PNC; 1990–1920)

(21) Ona volokom pritaščila na odejale drova iz saraja, namjala bumažnyx
komkov (…). Spički lomalis' i gasli, potom okazalos', čto net tjagi,
[ėto uže   kogda po   komnate popolz-l-i   plast-y
 PTC already when along room   crawl.PFV-PST-PL layer-NOM.PL
syr-ogo   dym-a   i   von'
damp-GEN.SG.M smoke.M-GEN.SG and stink.F-NOM.SG
žže-n-oj   bumag-i].
burn.IPFV1-PP-F.GEN.SG paper.F-GEN.SG
'She dragged firewood from the shed on a blanket, crumpled paper wads
(…). The matches broke and went out, then it turned out that there was
no draft, [this was already when layers of damp smoke and the stink of
burnt paper (i.e. the paper being burnt) crawled around the room].'

(Russian; RNC; 1890–1918)

3.3.2 General-factual (GF)

See the discussion in §2.1.

3.3.3 Iterative (ITER)

Here this term strictly refers to predicates that denote the repetition of an event on a single occasion (i.e. within a larger episode). This repetition can have a restricted count (e.g., *He knocked at the door five times*), the count may be unspecific (*He knocked at the door several times*), or it may be unrestricted (e.g., *He constantly knocked at the door*). In the latter case, it may become difficult to delimit iterative from progressive meaning.

Properly iterative use of participles in passives is extremely rare. In our samples we spotted only a handful of doubtful cases that might also be analyzed as progressive (22) or general-factual (23).

13 *On the grammatical integration of* n/t-*participles of imperfective stems*

(22) Wśród stukotu *obija-n-ych* garnk-ów i
among clatter   knock.on.IPFV2-PP-M.GEN.PL pot.M-GEN.PL and
talerz-y   klientk-i   dobiera-ł-y   pokrywk-i
plate.M-GEN.PL customer.F-NOM.PL choose.IPFV2-PST-PL.NVIR lid.F-ACC.PL
do rondl-i   albo talerzyk-i   do filiżanek.
to saucepant.M-GEN.PL or   plate.M-ACC.PL to cup.F-GEN.PL
'Amid the clatter of *knocked* pots and plates, customers chose pot lids or
plates for cups.'   (Polish; PNC; 1990–2020)

(23) – Ne mel'teši, Mixalyč, – proburčal Balandin.
      – [Triždy už   *govore-n-o*].   Čego opjat' nakačivaeš'?
        thrice already say.IPFV1-PP-SG.N
      '– Don't flicker, Mikhalych, Balandin muttered.
      – [It *has been said* already three times]. Why are you pumping up again?'
                                                      (Russian; RNC; 2004)

3.3.4 Habitual (HAB)

Habitual meanings occupy central stage in typologies of event-external plu-
ractionality.[14] They mark unlimited repetitions without an account of external
boundaries. What is "counted" is not subintervals within one episode, but the
episodes themselves, and this count is unspecific.[15] Habitual readings as such are
insensitive to actionality distinctions, i.e. to whatever is represented as repeated
in an unspecified number of occurrences.

(24) No nam užasno nravilos' slušat', i
     [ja   do six por   ne mog-u   ravnodušno slyša-t'
      I.NOM until this moment NEG can-PRS.1SG indifferently hear.IPFV1-INF
     igra-nn-ye   e-ju   p'esk-i].
     play.IPFV1-PP-PL 3SG.F-INS piece-NOM.PL
     'But we really enjoyed listening, and [I still cannot indifferently hear the
     pieces played by her (i.e. which she used to play, which becomes evident
     from the broader context)].'   (Russian; RNC; 1952–1971)

---

[14]We follow systematic classifications and their foundation, as in Cusic (1981), Xrakovskij (1997),
   Mattiola (2019), and, first of all, Šluinskij (2005, 2006).

[15]Notably, habitual situations need not be regular; in fact, more often than not they are irregular.
   This applies also to the meaning of ALWAYS and NEVER, which support habitual readings.

(25) Ona byla perevedena v konservatorskuju studiju –
[k načinajušč-im   ščenk-am   s   dran-ymi   nosk-ami i
 to beginning-DAT.PL puppy-DAT.PL with tattered-INS.PL sock-INS.PL and
redko *my-t-ymi*   griv-ami   do   pleč].
seldom wash.IPDV1-PP-INS.PL mane-INS.PL up.to shoulder-GEN.PL
'She was transferred to a conservatory studio – [to beginner puppies with
tattered socks and rarely washed shoulder-length manes].'

<div align="right">(Russian; RNC; 2003)</div>

(26) Oprócz tego MOP rozwinęła szeroką działalność naukowo-badawczą i
wydawniczą (…),
[ros-n-ą   też rozmiar-y   pomoc-y
 grow.IPFV1-PP-F.GEN.SG also size.M-NOM.PL help.F-GEN.SG
techniczn-ej   *świadczo-n-ej*   przez
technical-F.GEN.SG provide.IPFV1-PP-F.GEN.SG through
organizacj-ę   kraj-om   Trzeci-ego   Świat-a].
organization.F-ACC.SG country.M-DAT.PL third-GEN.SG.M world.M-GEN.SG
'In addition, the ILO has developed extensive research and publishing
activities (…), [and the size of technical assistance *provided* by the
organization to Third World countries is also growing].'

<div align="right">(Polish; PNC; 1945–1980)</div>

3.3.5 Stative (STAT)

Stative meanings capture situations without any boundaries and without any
(sub)intervals. The latter property distinguishes states from habitual situations.[16]
However, states may change. A particular case is resultative states.

(27) [List   swój   napisa-ł-a
 letter.M-ACC.SG POSSRED-M.ACC.SG writePFV-PST-F.SG
powodowa-n-a   żal-em],
cause.IPFV1-PP-F.NOM.SG regret.M-INS.SG
że w moim przewodniku po Puszczy kampinoskiej, nie znalazła
wzmianki o wsi swojego dzieciństwa i młodości.

---

[16]Examples are often ambiguous between a stative and a habitual reading for the reason that
they do not contain a clear indication of interval properties. While we cannot dwell on this
issue here, it does not much affect the statistical figures presented below, since HAB and STAT
are anyway the most frequent functions of ipfv. participles (see §4.3) and there is no reason
why, in such ambiguous cases, the annotation might have been biased toward either HAB or
STAT.

'[She wrote her letter out of regret (lit. caused by regret)] that in my guide to the Kampinoski Forest she found no mention of the village of her childhood and youth.' (Polish; PNC; 1990–2020)

(28) Ėt-ot svoeobrazn-yj gazov-yj ballon
this-M.NOM.SG peculiar-M.NOM.SG gas-M.NOM.SG cylinder.M-NOM.SG
soedinja-et-sja so stvol-om, *perekry-t-ym*
connect.IPFV2-PRS.3SG-REFL with barrel.M-INS.SG occlude.PFV-PP-INS.SG
diafragm-oj.
diaphragm.F-INS.SG
'This kind of gas cylinder is connected to the barrel, *occluded* with a diaphragm.' (Russian; RNC; 1974)

(29) [Kovrov-ye dorožk-i kazenno unyl-ogo
 carpet-NOM.PL pathway-NOM.PL state-owned dull-GEN.SG
cvet-a *ne čišče-n-y* i sbi-t-y], tam i
colour.M-GEN.SG NEG clean.IPFV1-PP-PL and knock.down.PFV-PP-PL
sjam vidny zatoptannye okurki.
'[The carpets of the official dull color are *not cleaned* and knocked down], here and there trampled cigarette butts are visible.' (Russian; RNC; 2004)

(30) Jug moskovsk-oj ojkumen-y
south.M-NOM.SG Moscow-F.GEN.SG ecumene.F-GEN.SG
dele-n na gorn-yj jugo-zapad i
divide.IPFV1-PP-SG.M on mountainous-M.ACC.SG south-west.M-ACC.SG and
ravninn-yj jugo-vostok.
flat-M.ACC.SG south-east.M-ACC.SG
'The south of the Moscow ecumene is *divided* into mountainous southwest and flat southeast.' (Russian; RNC; 2005)

(31) Tak-oj malen'k-ij, a tašč-it na buksir-e dv-e
such-M.NOM.SG small-M.NOM.SG but drag-PRS.3SG on tow.M-LOC two-F.ACC
ogromn-ye barž-i, *gruže-nn-ye* tes-om.
huge-ACC.PL barge-ACC.PL load.IPFV1-PP-ACC.PL batten.M-INS.SG
'So small, and drags in tow two huge barges, *loaded* with boards.'
(Russian; RNC; 1959)

4 Findings

We start from an account of lexical diversity (§4.1), then turn to frequency relations between *n/t*-participles and the general amount of grammatical forms

(§4.2) over to the relation between aspect and syntactic functions (§4.3), before we dwell on some more specific issues (§4.4).

4.1 Lexical diversity

In order to get an idea of how well-represented are grammatical means (or constructions) in a language (at some stage), it seems useful to assess their productivity, i.e. their spread among the stock of lexical units to which they apply. We will call this spread 'lexical diversity'. Since Slavic aspect is based on oppositions between stems (§1), we may count IPFV1, PFV and IPFV2 stems as lexical units to which participle suffixes, as grammatical means, apply. Our data allows for three, rather crude, ways to approximate lexical diversity (LD).

The first approach rests on type/token ratios, i.e. on coefficients between the number of different stems (= types) and the number of tokens in each sample. The value of type/token ratios varies between 0 and 1; the higher the value, the more diversified the number of stems which made it into the sample. Table 3 provides the figures for the Russian samples.

Table 3: Type/token ratio of stems with participles – Russian

	n/t				
	IPFV1		PFV		IPFV2
	triplets	control	triplets	control	all that could be found
1730–1780	0.21	0.40	0.24	0.69	0.67 (48 tokens)
1801–1850	0.19	0.38	0.24	0.78	only 4 tokens
1890–1918	0.17	0.41	0.23	0.81	only 1 token
1945–1980	0.18	0.42	0.20	0.86	–
1990–2020	0.19	0.50	0.21	0.81	–

Through all periods, the LDs are considerably higher in the control samples (which is natural since the choice of stems was not restricted by the triplet database), and among the control samples they are higher for PFV stems. Moreover, the figures are stable over time, except a steady increase for the PFV control samples, with a slight decrease in the last period. Moreover, we see that *n/t*-participles were derived from IPFV2 stems still in the 18th century and that, despite being infrequent, for that period their type/token ratio was comparable to the ratio of pfv. *n/t*-participles. Afterwards they drastically declined and virtually disappeared altogether (see further §4.4.2).

13 *On the grammatical integration of* n/t-*particles of imperfective stems*

Table 4 shows the figures of the Polish samples.

Table 4: Type/token ratio of stems with participles – Polish

	n/t								
	IPFV1			IPFV2			PFV		
	freq	infreq	ctrl	freq	infreq	ctrl	freq	infreq	ctrl
1730–1780	0.23	0.63	0.58	0.23	1.0	0.62	0.25	0.65	0.83
1801–1850	0.22	0.73	0.74	0.33	1.0	0.69	0.24	0.71	0.92
1890–1918	0.23	0.43	0.60	0.23	0.8	0.64	0.26	0.43	0.95
1945–1980	0.23	0.5	0.64	0.29	0.5	0.64	0.22	0.53	0.83
1990–2020	0.25	0.59	0.39	0.19	0.38	0.57	0.24	0.31	0.66

Among the freq-samples (for items restricted by the database) the type/token ratios are rather stable, and consistently so, in the range 19–33, mostly around 25–26, whereas in the infreq-samples the ratios are considerably higher, but they also vary over larger ranges, with higher values in the earlier periods particularly for PFV and IPFV2 stems. With the infreq-samples, IPFV2 stems even reach the possible maximum type/token ratio of 1, but it then drops drastically, and toward the last period the ratio of IPFV1 stems outruns the ratio of IPFV2 stems. As for the control groups, the ratios of PFV stems are consistently higher than for ipfv. stems. A decrease of the ratios can be observed for all control samples in the last period, quite drastically for IPFV1 and PFV stems, while the ratios of IPFV2 keep their level more or less over all periods.

The samples (already small as such) differ a lot as for their size. We therefore additionally calculated Herdan's Index, a type-token measure which is less sensitive to different sizes between samples, since it is based on the natural logarithms of raw figures (cf. Panas 2011: 523). However, this measure yielded the same relations between the coefficients as did the simple type/token measure presented in Table 3 and Table 4 (see Appendix, Part II).

Regardless, we should be aware that type/token ratios need not have decreased as such. Simply, the subcorpora of the later periods have a larger size, and in drawing random samples, frequent units have a better chance of making it into the sample more often. On the other hand, larger corpora also supply better chances for rare phenomena (e.g., stems with participles) to get into the sample. To draw more reliable conclusions about productivity, hapax phenomena would be more telling (cf. Baayen 2009). However, "fishing" them out would require an entirely different access to corpus data (which was not available).

Wiemer, Wrzesień-Kwiatkowska & Rostovtsev-Popiel

A way of getting an idea of productivity which comes closer to looking for hapaxes is the following. We can approach the LD of *n/t*-participles by asking for the proportion of stems which have at least one such participle form in relation to the overall amount of stems (= lexical units) in the corpus. This amounts to asking for the type frequency of *n/t*-participles in the corpus. We could not do that for the whole population of stems in the corpora, but we calculated these proportions (per period) for IPFV1, PFV and IPFV2 stems that are included in the triplet database. Table 5 provides the relevant figures.

As for Polish, there is no clear tendency, but PFV stems are most and IPFV2 are least productive. Russian PFV stems are more productive than IPFV1 stems with *n/t*-participles as well, and their figures are comparable to those of Polish PFV stems. However, in Russian both PFV and IPFV1 stems reveal a slight, but steady increase of productivity over time. Remember, however, that these observations are exclusively based on triplets.

A third approach toward LD is by calculating the mean proportions of participle token frequencies in relation to the token frequencies of the remainder of grammatical forms for each stem type (IPFV1, PFV, IPFV2). As a shortcut, these values may be dubbed "proportional frequencies". Again, this calculation was possible only for the units included in the triplet database.

Table 6 and Table 7 present the relevant figures. We start with Polish.

Table 6 testifies to a solid time stability of the mean token frequency of Polish *n/t*-participles for PFV and IPFV2 stems, only slightly less for IPFV1 stems. With PFV stems *n/t*-participles are considerably more frequent than for ipfv. stems. Moreover, IPFV2 stems employ *n/t*-participles about 2–3 times less frequently than do IPFV1 stems. However, Polish IPFV2 *n/t*-participles were by magnitudes more frequent than in Russian already in the 18th century, and their mean frequency has remained stable over time.

Table 7 supplies the figures for the Russian samples.[17]

We notice a reliable time stability for *n/t*-participles of IPFV1 stems, while such participles of IPFV2 stems were extremely rare already since 1730 and practically ceased to occur after 1918 (and probably earlier); this is why Table 5 lacks figures for the last two periods. However, as Table 3 and Table 4 show, in the first period (1730–1780) the type/token ratio of IPFV2 *n/t*-participles was quite high: with 0.67 it was comparable to the ratio of Polish IPFV2 *n/t*-participles of that period (0.62) (Table 4), simultaneously it was practically as high as of Russian PFV *n/t*-participles of that period and much higher than for IPFV1 *n/t*-participles of any period (Table 3). We will return to this issue in §4.4.2.

---

[17]Missing values indicate that corpus hits only contained citations from earlier periods.

Table 5: Type frequencies of stems (from database) with at least one *n/t*-participle

	1730–1780	1801–1850	1890–1918	1945–1980	1990–2020
POLISH					
n/t number of IPFV1 (= 100 %)	798	694	1,018	958	1,275
stems with at least one participle token	268 (33.6 %)	144 (20.7 %)	315 (30.9 %)	280 (29.2 %)	633 (49.6 %)
number of PFV (= 100 %)	851	832	1,274	1,316	1,721
stems with at least one participle token	480 (56.4 %)	405 (48.7 %)	749 (58.8 %)	648 (49.2 %)	1,138 (66.1 %)
number of IPFV2 (= 100 %)	372	391	661	565	1,009
stems with at least one participle token	46 (12.4 %)	41 (10.5 %)	132 (20.0 %)	92 (16.3 %)	395 (39.1 %)
RUSSIAN					
n/t number of IPFV1 (= 100 %)	537	692	762	804	810
stems with at least one participle token	126 (23.5 %)	185 (26.7 %)	244 (32.0 %)	261 (32.5 %)	290 (35.8 %)
number of PFV (= 100 %)	664	1,028	1,170	1,207	1,438
stems with at least one participle token	333 (50.2 %)	549 (53.4 %)	657 (56.2 %)	718 (59.5 %)	897 (62.4 %)
number of IPFV2 (= 100 %)	389	606	793	–	–
stems with at least one participle token	41 (10.5 %)	4 (0.7 %)	1 (0.1 %)	–	–

Wiemer, Wrzesień-Kwiatkowska & Rostovtsev-Popiel

Table 6: Means of proportions between *n/t*-participles and the rest of forms (Polish)

	1730−1780	1801−1850	1890−1918	1945−1980	1990−2020
ipfv1 *n/t*	0.09	0.052	0.07	0.075	0.07
pfv *n/t*	0.26	0.23	0.26	0.23	0.23
ipfv2 *n/t*	0.03	0.027	0.03	0.032	0.038

Table 7: Means of proportions between *n/t*-participles and the rest of forms (Russian)

	1730−1780	1801−1850	1890−1918	1945−1980	1990−2020
ipfv1 *n/t*	0.12	0.09	0.09	0.10	0.09
pfv *n/t*	0.31	0.39	0.40	0.45	0.26
ipfv2 *n/t*	0.009	0.006	0.007	−	−

As for PFV stems, the means fluctuate over a slightly larger range, also in comparison to the Polish sample series, even though the figures for both PFV and IPFV1 stems in Russian are consistently higher than for the equivalent Polish stems. Simultaneously, to repeat, Polish *n/t*-participles of IPFV2 stems were by magnitudes more frequent than in Russian already in the 18th century, and their mean frequency has remained stable over time.

In general, a comparison of Table 6 and Table 7 reveals that Russian has been employing *n/t*-participles for passives slightly more frequently than Polish, even with IPFV1 stems. This might seem surprising, however what needs to be checked is the relation of token frequency to lexical diversity. A comparison of Table 3 and Table 4 shows that the type/token ratios for IPFV1 stems and, in particular, for PFV stems in Polish are on average higher than in Russian. To the extent that type/token ratios can be understood as a rough indicator of productivity, Russian shows lower productivity than Polish, whereas in terms of proportional frequencies (Table 6 and Table 7) the relation is inverse: *n/t*-participles of IPFV1 and PFV stems on average occurred more often in Russian than in Polish.

4.2 Correlation between general token frequency and frequency of *n/t*-participles (Polish)

Let us now look whether, on average, the token frequency of *n/t*-participles for IPFV1, PFV, IPFV2 stems depends on the overall token frequency of forms occurring for these stems in the corpus. The more linear this relation, the stronger the correlation (Pearson's ρ; cf. Levshina 2015: 116–126). We have been able to calculate this correlation only for the Polish samples based on the triplet database.[18] The results are obvious (see Table 8). First, this correlation is strongest for PFV stems, consistently over all periods; it drops after the first period, but then remains more or less at the same level. Second, for IPFV2 stems the correlation grows almost steadily, and it outruns the correlations for IPFV1 stems in the last three periods. We may take this as an indication of an increasing integration of IPFV2 stems – and jointly with them of their *n/t*-participles – into the grammatical system.

Table 8: Correlation between token frequencies of *n/t*-participles and all forms (Polish)

Pearson's ρ	1730–1780	1801–1850	1890–1918	1945–1980	1990–2020
IPFV1	0.61	0.52	0.50	0.44	0.59
PFV	0.86	0.73	0.71	0.60	0.73
IPFV2	0.35	0.43	0.52	0.50	0.65

4.3 Relation between aspect functions and syntactic functions

We tested on significance (χ^2 or Fisher's exact test) and strength of association, or effect size (Cramer's V, with values ranging between 0 and 1).[19] Here only the main results are communicated.

---

[18]The database supplies a large, but manageable amount of verb stems, and the Polish corpora provide annotations which can be used directly for getting the relevant frequency data. The RNC does not provide such annotations and search tools, so that we could not avail ourselves of the relevant frequencies for these correlations.

[19]Tests on significance inform about the likelihood that the same correlation would be obtained from other samples; this likelihood is customarily indicated by a *p*-value (the smaller this value, the lesser the probability that the given result has been obtained by chance). In turn, strength of association measures the correlation itself. The latter does not depend on the sample size, whereas significance increases with sample size (cf. Levshina 2015: 129f.)

As concerns *n/t*-participles of PFV stems, in either language a consistent dominance of the stative function regardless of syntactic status can be observed. In most samples over all periods, this dominance reaches a significance level between $p < .05$ (*) and $p < 0.001$ (***). Higher *p*-values (i.e. less significance) are due to the eventive function, which shows a bias toward predicative use. Nonetheless, coefficients of eventive:stative function are only rarely higher than 0.5, i.e. the share of eventive *n/t*-participles in a sample is only rarely larger than 33 %. This happens in the last period of the Polish freq-sample (20:25), for the last period of the Polish control sample (15:11), and in the last period of the Russian control sample (16:25). That is, the relative frequency of eventive use slightly increases over time, but, except for the last period in the Polish control samples, it never prevails over the stative one. Moreover, for the Russian samples Cramer's V only rarely rises above 0.3, whereas for most Polish samples its value is between 0.4 and 0.7; that is, the correlation is in general stronger in Polish.

As concerns ipfv. *n/t*-participles, we only compared the most frequent aspect functions, STAT and HAB, against the remainder of aspect functions. That is, GF is not the predominant function in either of the languages. This also applies to their predicative use; in Polish, the habitual function proves to be particularly dominant in the predicative use of IPFV2 *n/t*-participles. Admittedly, beside examples with an undoubtedly stative reading (see (32)), a larger number of examples turned out difficult to categorize, as e.g. (33), which might be assigned habitual function, or (34), which might be interpreted as GF. See examples for Russian:

(32) (…) našu večno bedstvovavšuju prijatel'nicu Ninu Alovert,
      [u kotor-oj   do s-ix   por   za
       by which-F.GEN.SG up.to this-GEN.PL time-GEN.PL for
      telefon   *ne plače-n-o* (…)].
      telephone.M-ACC.SG NEG pay.IPFV1-PP-SG.N
      'our ever-poor friend Nina Alovert, [who still hasn't paid (lit. on which it
      isn't paid) for her phone]'   (Russian; RNC; 1998)

(33) [Skol'ko   sil   na «Orfej-a»   *trače-n-o*],
       how.much power-GEN.PL on Orpheus.M-ACC.SG spend.IPFV1-PP-SG.N
      a čego-to ėtoj tragedii vsegda ne dostavalo!
      '[How much effort had been spent (lit. *is spent*) on "Orpheus"], but
      something was always lacking in this tragedy!'   (Russian; RNC; 2010)

(34) Pavlo vskočil, paren' molodoj, krov' svežaja,
[lagerj-ami   ešče *ne   trepa-n*],   na galuškax ukrainskix rjažka
 camp-INS.PL yet   NEG maul.IPFV1-PP-SG.M
ot"edennaja.
'Pavlo sprang to his feet, a young lad with fresh blood, [*not* yet *mauled* by
the camps], used to stuffing his face with Ukrainian dumplings.'

<div align="right">(Russian; RNC; 1961)</div>

For Polish, analogous cases could be adduced.

On the other hand, there is also a considerable number of examples for which
GF has been assigned only with doubts, as in (35)–(36):

(35) Žal' tol'ko,   GF or STAT?
[čto   ona   davno   *ne   čišče-n-a*].
 COMP she.NOM.SG long.ago NEG clean.IPFV1-PP-SG.F
'It is only a pity [that it (the scapula) has *not* been (lit. is not) *cleaned* for a
long time].'   (Russian; RNC; 2011)

(36) A naši ženščiny vse ravno byli v nas,   GF or HAB?
[i   skol'ko *by-l-o*   o   nix   *govore-n-o*].
and how.much be-PST-SG.N about 3.LOC.PL say.IPFV1-PP-SG.N
'And our women were all the same in us, [and how much *was said* about
them].'   (Russian; RNC; 1990–1996)

There is no reason to assume that such hesitations in the assignment of aspect
functions have skewed their general distribution in the samples toward any of
these functions (GF vs HAB or STAT); see also Footnote 16.

Regardless, as concerns *n/t*-participles of IPFV1 stems, results on different lev-
els of significance (from $< .05^*$ to $< .001^{***}$) for Polish are found in most periods
of the freq- and the control samples, with Cramer's V between 0.24 and 0.47 (ex-
cept in the first period of the control sample: V = 0.53). In general, STAT is more
closely associated to attributive use, but in the control samples we also observe
an extreme preference of HAB in predicative use; compare (37)–(38).

(37) W świecie już   od dawna są   *produkowa-n-e*
in world   already since.long be.PRS.3PL produce.IPFV1-PP-NOM.PL
różn-ego   rodzaj-u   ciągadł-a   ciśnieniow-e   o
various-M.GEN.SG kind.M-GEN.SG die.N-NOM.PL of.pressure-N.NOM.PL of
podobn-ej   zasadzi-e   działani-a.
similar-GEN.SG.F principle.F-GEN.SG action.N-GEN.SG
'Various types of pressure dies with a similar principle of operation have
been *produced* in the world for a long time.'   (Polish; PNC; 1945–1980)

(38) Ci zaś, którzy jak dzieci na wielkanocnych plackach odłubują tylko
      rodzynki życia i zjadają,
      [c-i   wiecznie za dziec-i   będ-ą
       this-VIR.NOM.PL eternally for child.N-ACC.PL FUT-3PL

      mia-n-i   i *sądze-n-i*   jako
      have.IPFV1-PP-VIR.NOM.PL and judge.IPFV1-PP-VIR-NOM.PL as

      dziec-i.]
      child.N-NOM.PL
      'And those who, like children on Easter cakes, pick only the raisins of life
      and eat them, [those will eternally be considered children and be *judged*
      as children].'   (Polish; PNC; 1801-1850)

The samples of Russian IPFV1 *n/t*-participles produced highly significant *p*-
values for almost all samples, except the control samples, among which there
is no clear tendency for the stative function (nor more remarkable values of
Cramer's V). For the triplet-based samples, we observe a clear association be-
tween stative function and attributive use only in the last two periods; see (39).

(39) Grubo *bele-nn-ye*   sten-y   ne kaza-l-i-s'
      roughly whiten.IPFV1-PP-PL wall-NOM.PL NEG seem-PST-PL-REFL

      sliškom   goly (…).
      too.much naked-NOM.PL
      'The roughly *whitewashed* walls did not seem too bare.'
                                                  (Russian; RNC; 1996–1997)

The samples of Polish IPFV2 *n/t*-participles yielded results on significance levels
 < 0.5 (*) to < 0.001 (***) in all freq-samples and in all control samples but the
last, i.e. this period does not show a clear distribution. In general, STAT tends
towards attributive and appositive use (see (40)–(41)), but HAB does not show
any clear preference, except for the last period in the freq-samples. The infreq-
samples yielded no clear results.

(40) Po ukończeniu obrządku,
      [z   jednej strony gospodarz,   a   z   drugiej
       from one side   housekeeper.M-NOM.SG and from other

      gospodyn-i   dom-u   traktowa-l-i
      housekeeper.F-NOM.SG house.M-GEN.SG treat.IPFV1-PST-PL

      wszystk-ich   z   kolei *pokraja-n-em*   jaj-em]   i nie
      everybody-ACC.PL from turn slice.IPFV2-PP-N.INS.SG egg.N-INS.SG and not

      opuszczali najmniejszego dziecka.

> 'After completing the rite, [the housekeeper, from one side, and his wife, from the other, treated everybody with a sliced egg], and they did not omit even the smallest child.' (Polish; PNC; 1801–1850)

(41) (…) zaraz   zatamowa-ł-a   krew   zapomocą
        immediately block.PFV-PST-SG.F blood.ACC with.aid

      bibuł-y,   *umacza-n-ej*   w jak-imś płyni-e
      blotting.paper.F-GEN wetten.IPFV2-PP-F.GEN.SG in some   liquid

      gryząc-ym.
      biting
      '(…) she immediately blocked the blood with blotting paper, dipped (lit. *wettened*) in some acrid liquid.' (Polish; PNC; 1890–1918)

4.4 Specific issues

In the remainder, we will discuss aspect functions of ipfv. *n/t*-participles in both languages (§4.4.1) and the fate of *n/t*-participles of IPFV2 stems in Russian (§4.4.2).

4.4.1 General-factual vs habitual and progressive function of ipfv. *n/t*-participles

As stated in §4.3, GF has not turned out a predominant aspect function of ipfv. *n/t*-participles for most of the samples over the periods, and this also holds for Russian. This may surprise given the prominence ascribed to this function in research dealing with ipfv. *n/t*-participles in contemporary Russian, which has concentrated on predicative use (see §2.1). Here we discuss GF together with PROG on the background of HAB, and we start with Russian.

    As Table 9 shows, the token frequency of progressive readings has always been low, and toward the present period it approaches zero. Only in one sample of the 1st and of the 2nd period did PROG prevail over GF, the predominance of the latter increases toward the current period. GF also increases in comparison to HAB, but only during the last two periods has it taken dominance over HAB. Thus, its salience mentioned in the aforementioned research appears to be recent.[20]

    As concerns PROG, even the few examples found in 1945–1980 raise doubts, as their temporal reference is not entirely clear and they can also be assigned GF, perhaps even ITER (see (42)–(43)). In other cases, one can argue for assigning STAT, also for earlier periods as in (44):

---

[20]We use "na" if none of the functions is attested. If only one of the compared functions is attested, GF, HAB or PROG, respectively.

Table 9: HAB, GF, PROG for Russian IPFV1 *n/t*-participles

| | 1730–1780 | | 1801–1850 | | 1890–1918 | | 1945–1980 | | 1990–2020 | |
	tripl	ctrl	tripl	ctrl	tripl	ctrl	tripl	ctrl	tripl	ctrl
HAB	26	48	19	51	13	32	1	15	0	24
GF	22	2	8	11	9	17	12	40	25	23
PROG	5	5	10	7	3	5	1	6	0	0
coeff. GF/PROG	4.4	0.4	0.8	1.6	3.0	3.4	12.0	6.7	GF	GF
coeff. HAB/PROG	5.2	9.6	1.9	7.3	4.3	6.4	1.0	2.5	na	HAB
$\sum$ sample	108	97	119	100	100	100	142	90	109	100

(42) Posle ėtogo professora Universiteta v vide protesta ustroili obed v čest'
      Mendeleeva,
      [vo vremja kotor-ogo   *govore-n-y*   *by-l-i*
       at time   which-M.GEN.SG say.IPFV1-PP-PL be-PST-PL
      sootvetstvujušč-ie   reč-i].
      appropriate-NOM.PL speech-NOM.PL
      'After that, as a protest, the professors of the University organized a
      dinner in honor of Mendeleev, [during which appropriate speeches *were
      pronounced*].'   (Russian; RNC; 1968)

(43) S utra do noči stol lomilsja ot edy i vina – ot lobii, sacivi, žarenoj ryby
      look, (…)
      [glinjan-yx goršočk-ov s   *tuše-nn-ym*   v ostr-yx
       clay-GEN.PL pot-GEN.PL with stew.IPFV1-PP-INS.SG in pungent-LOC.PL
      prjanostj-ax mjas-om].
      spice-LOC.PL meat.M-INS.SG
      'From morning to night, the table was full of food and wine – from lobia,
      satsivi, fried loco fish (…), [clay pots with meat *stewed* in hot spices].'
                                                        (Russian; RNC; 1963)

(44) [Siloslav,   bud-uči *trevož-en*   i-mi],   govoril volšebnice,
       PN.M-NOM.SG be-CVB alarm.IPFV1-PP.SG.M 3-INS.PL
      čto ne možet probyt' tut ni odnoj minuty.
      '[Siloslav, *alarmed* by them], told the sorceress that he could not stay
      here for a single minute.'   (Russian; RNC; 1766–1768)

13 *On the grammatical integration of* n/t-*participles of imperfective stems*

A few clear examples of progressive use can be found in the samples of the first two periods; compare (45)–(46):

(45) [Kogda ja *by-l*   *vede-n*   na kazn'],   to šestvie
       when 1SG be-PST-SG.M lead-PP-SG.M to execution.F-ACC.SG
       moe bylo takim obrazom.
       '[When I *was (being) led* to execution], my procession was like this.'
                                                    (Russian; RNC; 1766–1768)

(46) Ne mogu opisat' vamъ radosti,
       [*čuvstvova-nn-oj*   mn-oju vъ s-iju   minut-u].
        feel.IPFV1-PP-F.GEN.SG 1SG-INS in this-F.ACC.SG minute.F-ACC.SG
       'I cannot describe to you the joy [I felt (lit. *felt* by me) at this moment].'
                                                          (Russian; RNC; 1812)

Examples from the period 1890–1918 are more difficult to classify. In (47), for instance, the classification as progressive or iterative (which itself is extremely rare) depends on whether the focus is on a series of intervals within a discrete larger episode (→ iterative) or on the continuity of attempts (→ progressive); the problem is not just whether we are dealing with subintervals, but whether accentuating such intervals is the proper "point" – a question that can at best be solved on the basis of a larger discourse fragment.

(47) K severo-vostoku ot Černovic protivnik otčajannymi kontr-atakami,
       [*vede-nn-ymi*   bol'š-imi sil-ami],   pytalsja zaderžat'
        carry.out.IPFV1-PP-INS.PL big-INS.PL power-INS.PL
       naše nastuplenie.
       'To the northeast of Chernivtsi, the enemy tried to delay our advance
       with deperate counter-attacks [*carried out* by large forces].'
                                                          (Russian; RNC; 1916)

Now let us turn to Polish, for which we first discuss the relation between GF and PROG, then between HAB and PROG.

Table 10 provides coefficients between GF and PROG; values higher than 1 testify to a predominance of GF, values between 1 and 0 indicate a predominance of PROG. If only one of the two functions is attested, the cell indicates GF or PROG, respectively. As Table 10 shows, no clear tendencies can be inferred from a pairwise comparison of IPFV1 and IPFV2 stems over the periods.[21]

---

[21]We use "na" if none of the functions is attested.

Table 10: GF / PROG coefficients for Polish ipfv. *n/t*-participles

coefficient	1730–1780	1801–1850	1890–1918	1945–1980	1990–2020
freq					
IPFV1	1.7	0.4	0.4	0.1	1.1
IPFV2	1.0	GF	PROG	0.1	3.0
infreq					
IPFV1	6.0	9.0	2.0	0.4	1.0
IPFV2	GF	na	PROG	0.2	0.4
control					
IPFV1	GF	0.7	0.4	0.8	0.3
IPFV2	3.7	1.0	0.1	0.1	0.6

Despite the lack of a clear tendency of PROG in relation to GF, and in contrast to Russian, ipfv. *n/t*-participles do not "lose" PROG, but retain it. It even seems to slightly increase by the modern period, at least in the control and the infreq-samples, in which PROG now seems a bit more prominent for *n/t*-participles of IPFV2 stems than of IPFV1 stems. Good examples of PROG are difficult to find for the earliest periods (maybe because of data scarcity). All these findings hold true for either stem type; see (48)–(51).

The following four examples show the employment of ɪᴘꜰv1 stems ((48)–(49)) and of ɪᴘꜰv2 stems ((50)–(51)) in progressive function.

(48) Puścił żagiel i, *gna-n-y* pochyleni-em się
let sail and drive.ɪᴘꜰv1-ᴘᴘ-ᴍ.ɴᴏᴍ.sɢ stoop.ɴ-ɪɴs.sɢ ʀᴇꜰʟ
statk-u, pobieg-ł mimowolnie drobn-ym
ship.ᴍ-ɢᴇɴ.sɢ run.ᴘꜰv-ᴘsᴛ-ᴍ.sɢ involuntarily small-ɪɴs.sɢ.ᴍ
kroki-em na tył.
step.ᴍ-ɪɴs.sɢ on rear-ᴀᴄᴄ
'He let go off the sail and, *driven* by the stoop of the ship, involuntarily ran in small steps to the rear.' (Polish; PNC; 1890–1918)

(49) Największ-a inwestycj-a obecnie
largest-ꜰ.ɴᴏᴍ.sɢ investment.ꜰ.ɴᴏᴍ.sɢ presently
realizowa-n-a w park-u to budow-a
perform.ɪᴘꜰv1-ᴘᴘ-ꜰ.ɴᴏᴍ.sɢ in park-ʟᴏᴄ ᴘᴛᴄ construction.ꜰ-ɴᴏᴍ.sɢ

aqua-parku „Fala".
aquapark
'The largest investment currently underway [lit. *(being) realized*] in the
park is the construction of the "Fala" aquapark.' (Polish; PNC; 1990–2020)

(50) Wesz-l-i,   i w głęb-i   stodoł-y   ujrze-l-i
enter.PFV-PST-PL and in depth.F-LOC.SG barn.F-GEN.SG spot.PFV-PST-PL
wialni-ę do czyszczeni-a zboż-a, *obsługiwa-n-ą*   przez
cleaning.plant.F.ACC   operate.IPFV2-PP-F.ACC.SG through
kilk-u   robotnik-ów.
some-ACC laborer.M-ACC.PL
'They entered and, at the far end of the barn, saw a grain cleaning plant
operated by some laborers.'   (Polish; PNC; 1890–1918)

(51) Zauważyłem go wychodząc na pokład, gdy   znika-ł
noticed   him exiting   on deck   when disappear.IPFV-PST-SG.M
za   ruf-ą,   *zawija-n-y*   już   w zwoj-e
behind stern.F-ACC.SG wrap.IPFV2-PP-M.NOM.SG already in coul.M-ACC.PL
mokr-ej   mgł-y.
wet-GEN.SG.F fog.F-GEN.SG
'I spotted it as I stepped onto the deck, when it disappeared after the
stern, *wrapped* in coils of wet fog.'   (Polish; PNC; 1990–2020)

These observations are indicative that the absence vs. presence of the prefix in
the stem does not have a conceivable impact on the average aspectual behavior
of the ipfv. *n/t*-participle. First of all, since both IPFV1 and IPFV2 *n/t*-participles
occur in progressive function, either stem type can defocus a culmination point
entailed by telicity, whether it be induced by the lexical meaning of the stem
or only by the prefix. Moreover, IPFV1 stems can convey an idea of unlimited
repetition of culmination points (i.e. HAB + telic) as well. See examples from the
samples:

(52) Nie rozumieli dlaczego znów mają płacić za elektroniczne,
[skoro   wcześniej płaci-l-i   już   za wyparkow-e,
 because earlier   pay.IPFV1-PST-PL already for evaporator-M.ACC.PL
tańsz-e   w obsłudze i   powszechnie *stosowa-n-e*
cheaper-M.SACC.PL in service   and commonly apply.IPFV1-PP-M.ACC.PL
w cał-ym   kraj-u].
in whole-LOC.SG.M country.M-LOC.SG
'They did not understand why they had to pay for electronics again,

[since they had already paid for evaporators, cheaper to operate and widely *applied* throughout the country].'   (Polish; PNC; 1990–2020)

(53) Wprowadzenie takiego systemu i dalsze udoskonalanie go
[pozwala   poprawi-ć   jakość
allow.ɪᴘғᴠ-ᴘʀs-3sɢ improve.ᴘғᴠ-ɪɴғ quality.ғ-ᴀᴄᴄ.sɢ
świadczo-n-ych   usług]   i utrzymać na odpowiednim
provide.ɪᴘғᴠ1-ᴘᴘ-ғ.ɢᴇɴ.ᴘʟ service.ғ-ɢᴇɴ.ᴘʟ
poziomie koszty eksploatacji.
'The introduction of such a system and further improvement of it [allows to improve the quality of services *provided*] and to keep operating costs at an appropriate level.'   (Polish; PNC; 1990–2020)

(54) W Poznaniu   *realizowa-n-e*   są   częściowo
in Poznań.ᴍ-ʟᴏᴄ.sɢ perform.ɪᴘғᴠ1-ᴘᴘ-ғ.ɴᴏᴍ.ᴘʟ be.ᴘʀs.3ᴘʟ partially
dostaw-y   w butelk-ach   0,5 l.
delivery.ғ-ɴᴏᴍ.ᴘʟ in bottle-ʟᴏᴄ.ᴘʟ
'In Poznań, some deliveries are made in 0.5 l bottles.'
                                        (Polish; PNC; 1945–1980)

Finally, a look at the coefficients between HAB and PROG (see Table 11) reveals that, although, again, there is no clear tendency for any of the sample groups, HAB dominates over PROG in most samples (coefficient > 1), and more consistently so than between PROG and GF (see Table 10). The degree of this dominance varies a lot, but it is often very high, particularly for IPFV2 *n/t*-participles. To the contrary, if PROG dominates (coefficient < 1), it is for IPFV1 stems.[22]

4.4.2 Russian *n/t*-participles of IPFV2 stems

As pointed out in §4.1, although we found that Russian *n/t*-participles of IPFV2 stems (altogether 48 tokens) were considerably less frequent than their equivalents in Polish already in the 18th century and obviously ceased to be in active use by the beginning 20th century, their type/token ratio was remarkably high during 1730–1780. With 0.67 it was comparable to the LD of Polish IPFV2 *n/t*-participles in the same period (0.62) (see Table 4); simultaneously, it was practically as high as of Russian PFV *n/t*-participles of that period (0.69) and much higher than for IPFV1 *n/t*-participles of any period (see Table 3). From this we might infer that, despite their rarity, *n/t*-participles of IPFV2 stems showed a

---

[22]HAB if no PROG is attested in the sample.

Table 11: HAB / PROG coefficients for Polish ipfv. *n/t*-participles

coefficient	1730–1780	1801–1850	1890–1918	1945–1980	1990–2020
freq					
IPFV1	8.0	2.5	0.8	0.8	3.1
IPFV2	13.5	HAB	1.7	1.6	76.0
infreq					
IPFV1	6.0	8.0	2.0	1.1	7.5
IPFV2	HAB	HAB	7.5	1.7	7.4
control					
IPFV1	HAB	1.5	0.5	0.9	2.2
IPFV2	13.7	4.5	3.2	2.8	4.8

broader spread in the lexicon (= amount of verb stems on which they occurred). In fact, these participles appear to have been more productive than any Russian ipfv. *n/t*-participles after the end of the 18th century, before their frequency dropped abruptly and shortly after became zero.

Moreover, among the 41 instances we found in the RNC for 1730–1780, the habitual function prevails (16 instances), while there are 10 cases with stative function, and most of the 9 GF uses are doubtful; there are also 2 debatable cases of progressive use. See (55)–(59).

(55) exemplifies the habitual function.

(55) Na sej konecъ postroeny byli vně zemljanago Kammer-Kolležskago vala,
po vsěmъ bol'šimъ dorogamъ anbary i torgovyja města,
[gdě   s"estn-ye   pripas-y   *skladyva-n-y*], i gdě by vъ slučaě
 where edible-NOM-PL supply-NOM.PL lay.IPFV2-PP-PL
nadobnosti dolžny byli priěžžajuščie na torgi krest'jane črezъ ogradu
vsemu činit' prodažu (…)
'In this end, along all the main roads, (a number of) barns and trading
places were built outside of the Kamer-Kollezhsky rampart, [where
comestibles *used to be laid*] and where peasants who would come to
trade, in case of necessity, had to carry out sales over a fence.'

(Russian; RNC; 1775)

Example (56) illustrates the stative function.

(56) Ja sobral iz naxodjasčixsja v zemle razvalin nekoliko izrascov zelenyx,
byvšix v stroenii,
[meždo kotor-ymi dv-a cel-ye šestiugol'nik-a
 between which-INS.PL two-M.NOM.PL whole-NOM.PL hexagon.M-GEN.SG
rozpisyva-n-y zolot-om, drug-ie, na kotor-yx
paint.IPFV2-PP-PL gold.N-INS.SG other-NOM.PL on which-LOC.PL
vrezyva-n-y liter-y bel-ye, in-ye
carve.into.IPFV2-PP-PL letter.F-NOM.PL white-NOM.PL other-NOM.PL
sostavliva-n-y iz kusk-ov razn-ago cvet-a
compile.IPFV2-PP-PL from piece-GEN.PL distinct-GEN.SG colour.M-GEN.SG
poliva-nn-yx] i viditsja bez uzora.
permeat.IPFV2-PP-GEN.PL
'I collected several green tiles from the underground ruins, that had been
in the building, [among which there were two intact hexagons painted
with gold, some with white letters carved into them, some compiled from
pieces permeated with distinct colours] and appear patternless.'
(Russian; RNC; 1741)

The following is a clear illustration of the existential type of the general-factual
function.

(57) No vъ čislě tovarov, šersti i xlopčatoj bumagi, kotoraja ne vъ dělě,
otkudabъ vezena ni byla, xotja by i svidětel'stvo imeli,
[čto *provětriva-n-o*], ne propuskat'.
 COMP air.IPFV2-PP-SG.N
'But of the wares, do not let pass wool and unmanufactured cotton,
regardless of whence it be brought, (and) even if (the carriers) present a
certificate (stating) [that (the item) *has been aired out*].'
(Russian; RNC; 1771)

As for (58), it is difficult to tell whether it shows the general-factual or the pro-
gressive reading.

(58) Prežde vsego vspomni, čto, kogda ty ešče byv mladencem, vyšel iz
spasitel'noj kupeli:
[togda svjaščenn-ym mir-om ušes-a tvo-i *by-l-i*
 then holy-M.INS.SG myrrh.M-INS.SG ear.N-NOM.PL thy-NOM.PL be-PST-PL
pomazyva-n-y], s proiznošeniem six slov: Vo uslyšanie very.
anel.IPFV2-PP-PL

'First of all, recall that when thou wert an infant, thou camest from the redeeming laver: [then *were* thyne ears *aneled* with holy myrrh], along with the pronunciation of the following words: For the (true) understanding of faith.' (Russian; RNC; 1777)

(59) most likely exemplifies the progressive function, referring to a telic process.

(59) [No naposledok on-ym   igumen-om   i   bojar-y
      but lastly.ADV   that-M.INS.SG hegumen.M-INS.SG and boyar-INS.PL
      tak   *by-l*   *obezpokoeva-n*],   čto, poslušav ix, kljatvu,
      that.way be-PST-SG.M discomfort.IPFV2-PP-SG.M
      dannuju Jaroslavu, prestupil i (…) do smerti o tom sožalel i nikogda bez
      plača i vozdyxanija ne vospominal.
      '[But lastly, (he) *was* so much *discomforted* by that hegumen and the
      boyars] that, having listened to them, (he) violated the oath (that he had)
      made to Jaroslav and (…) regretted that till his death and never recalled
      (that) without lamentation and sighing.' (Russian; RNC; 1750)

Among the few instances encountered in the second and third period, practically all are problematic in the assignment of aspect functions. For instance, in the following example it is hardly possible to decide whether we are dealing with GF or HAB:

(60) Xotja zemleopisatel'noj ėkspedicii i ne nadležit otnjud' vxodit' v
     meževanie zemel', no v opisanijax svoix dolžna ona označit',
     [k-em,   kak-ie   zeml-i   obrabatyvaj-u-tsja
      who-INS what.kind.of-NOM.PL land-NOM.PL cultivate.IPFV2-PRS.3PL-REFL
     ili zapušče-nn-ye k-em,   prežde   *obrabotyva-n-y*
     or neglect.PFV-PP-PL who-INS previosuly.ADV cultivate.IPFV2-PP-PL
     by-l-i].
     be-PST-PL
     'Although the geographical expedition ought not at all to conduct
     boundary surveys, even so it has to specify in its descriptions [what lands
     are cultivated and by whom, or by whom the ones that are left untilled
     had previously *been cultivated*].' (Russian; RNC; 1822)

At any rate, the history of IPFV2 *n/t*-participles in Russian provides an example of a category which showed productivity in the lexicon (high type-frequency), although it was rare on token level, before it "died out". With regard to aspect functions (as far as less than 50 examples in our samples can be indicative), the predominance of HAB seems to confirm what one would predict from Tatevosov's

Wiemer, Wrzesień-Kwiatkowska & Rostovtsev-Popiel

(2015) analysis of verb stems with lexical prefixes. However, we also find *n/t*-participles from IPFV2 stems which hardly refer to culmination points, either because of the actionality of the stem (as atelic *osmeivano* 'ridiculed' in (61)) or because it otherwise is difficult to "get" (*vozpitovany* 'raised' in (62)).[23]

(61)   (…) a drugie po zavisti ko mne, tret'i po trusosti (…) raznymi sposobami davali mne v tom prepinanija, tak čto togda počti v každom tex gospod dome to moe, čtob v polki vmesto anglijskix iz rossijskix sukon mundiry delat',

    [predprijati-e   *osmeiva-n-o*   i   xudo
     undertaking.N-NOM.SG ridicule.IPFV2-PP-SG.N and illy.ADV

    tolkova-n-o   *by-l-o*].
    interpret.IPFV1-PP-SG.N be-PST-SG.N

    '(…) and others out of envy toward me and yet others out of their cowardice (…) created impediments for me in that in almost every house of those lords, [my undertaking with regard to manufacturing full dress uniforms for the army (lit. regiments) out of Russian broadcloths instead of English ones *was* then *ridiculed* and ill-interpreted].'

                                            (Russian; RNC; 1766–1777)

(62)   O sem potrebno by vnjatnee razsmotret', ibo onye sut' dvojakie,
    [odn-i   bogougodn-omu   ustav-u   Petr-a
     one-NOM.PL charitable-M.DAT.SG charter.M-DAT.SG PN.M-GEN.SG

    Velik-ago   *by-l-i*   *zbira-n-y*,   *vozpitova-n-y*,   i
    great-M.GEN.SG be-PST-PL gather.IPFV2-PP-PL raise.IPFV2-PP-PL and

    obuča-em-i   v sirotsk-ix   dom-ex].
    teach.IPFV2-PRS.PP-PL in orphan-LOC.PL house-LOC.PL

    'This (matter) requires a more articulate consideration, for those are double-natured, [the solitary *were gathered*, *raised*, and *taught* in orphanages in accordance with Peter the Great's charitable charter].'

                                                  (Russian; RNC; 1733)

5 Conclusions and outlook

We may draw some conclusions. These have to be cautious, at least as for their empirical basis, since this was the probably first corpus-based pilot study on the

---

[23](61) also illustrates that *n/t*- and *m*-participles of IPFV2 stems could occurred in coordination. However, here the role of *m*-participles (in Russian) is not considered.

development of participles and their role in the aspect-voice system of Russian and Polish.

As concerns aspectual semantics and aspect functions, some principled remarks appear appropriate. First, the general-factual function (GF) associated to ipfv. aspect should be primarily assessed in terms of information structure (presupposed vs asserted information) rather than in temporal semantics. Second, the status of resultative subevents for IPFV1 stems in a telic setting (on clause level) is labile. A resultative subevent is asserted by suitable (lexical, including natural) prefixes, but it need not be absent (and can be "activated") if such prefixes are lacking. Actually, this is what happens with IPFV1 stems when they "replace" their PFV counterparts (with natural prefixes), e.g. in the narrative present tense or in pluractional functions.

Third, since IPFV2 stems inherit the resultative subevent from "their" PFV stem, there is no point in restricting their use from habitual readings – *pace* Tatevosov (2015), who does not seem to notice the consequences of his reasoning concerning the role of lexical (including natural) prefixes. Functions of external pluractionality are themselves insensitive to actionality distinctions, as is GF. Fourth, since, following Tatevosov's reasoning, IPFV1 stems are void of (non-cancellable) resultative subevents, there is also no reason why their *n/t*-participles should block, or avoid, habitual or progressive readings. If such readings are indeed avoided (or inacceptable) for ipfv. *n/t*-participles in contemporary Russian, the reason can exactly not be sought in their semantic structure; instead, the reason should be connected to the way these participles are integrated (or not) into the grammatical system at the interface between aspect and voice. This can be clearly seen from the largely different development of these participles in Polish, including also IPFV2 stems. The relatively infrequent occurrence of ipfv. *n/t*-participles in progressive function is nothing particular of constructions with these participles, but a general feature of aspect use in Slavic languages: many ipfv. verbs do not allow for progressive function (cf. Lehmann 1998 for Russian), i.e. their type frequency is limited, and this applies also to token frequency (Wiemer et al. 2020, and M. Łaziński, p.c., for Polish).

In addition, from among the findings of our corpus-based study we may point out the following. Although participles of IPFV2 stems show a certain preference for habitual situations, this does not entail a focus on achieved culmination points (as Tatevosov's 2015 reasoning would indeed predict). The token frequency of *n/t*-participles in habitual function was particularly high in the early periods, both for IPFV1 and for IPFV2 stems. But while in Russian IPFV2 *n/t*-participles became obsolete, their Polish equivalents have been integrated tightly into the aspect-voice interface. This can be seen from their productivity indicators and

from their more even distribution over aspect functions (in parallel to IPFV1 *n/t*-participles). Russian IPFV2 *n/t*-participles, before they disappeared, were mainly used in habitual and stative meaning, not in GF. Furthermore, type frequency (lexical diversity) does not seem to depend much on token frequency. For instance, despite restricted usage in 18th-19th c. Russian (and their subsequent disappearance), the type/token ratio of IPFV2 *n/t*-participles in the 18th century was not lower than for their Polish equivalents.

As concerns ipfv. *n/t*-participles in general-factual use, the corpus data indicate that in Russian this function has become prominent only in the last two periods (i.e. since 1945). Moreover, even in this recent time, presuppositional GF does not seem to dominate over negated ipfv. *n/t*-participles which mark existential GF. This "polarity split" for subtypes of GF needs further research. By contrast, Polish *n/t*-participles of ipfv. stems do not show any bias toward GF; in general, aspect functions are rather evenly distributed over *n/t*-participles of both IPFV1 and IPFV2 stems. This testifies to their tight integration into the aspect-voice interface in the language.

Admittedly, these conclusions rest on a restricted amount of data, albeit largely assembled via random sampling. Therefore, our findings should be tested against a larger amount of data, as well as for other Slavic languages in which ipfv. stems demonstrate productive derivation of *n/t*-participles. An analogous caveat is justified concerning productivity. We have applied different, and rather rough measures of productivity (lexical diversity) with partially different vantage points. This has, in part, led to superficially contradictory results (see §4.1). Type/token ratios give us only a bird's-eye view (see Table 3 and Table 4), while Table 5 supplies more detailed information on types and Table 6 and Table 7 are oriented toward tokens. Type/token ratios are a basic (and certainly insufficient) measure of productivity beside measurements oriented toward phenomena that come close to hapax legomena (cf. Baayen 2009). However, our study was not concerned particularly with such phenomena; it was mainly oriented toward the productivity of grammatical patterns between different types of verb stems. A database of aspect triplets has proved helpful in getting a handle on corpora whose annotation appears often insufficient for determining grammatical patterns related to aspect and voice. New approaches toward productivity in diachronic corpus studies, such as permutation testing (cf. Säily & Suomela 2017), should be checked as for their suitability in cases like the one presented here.

Acknowledgements

References

Arkadiev, Peter & Björn Wiemer. 2020. Perfects in Baltic and Slavic. In Robert Crellin & Thomas Jügel (eds.), *Perfects in Indo-European languages and beyond* (Current Issues in Linguistic Theory), 123–214. Amsterdam, Philadelphia: John Benjamins. DOI: 10.1075/cilt.352.05ark.

Baayen, Harald. 2009. Corpus linguistics in morphology: Morphological productivity. In Anke Lüdeling & Merja Kytö (eds.), *Corpus linguistics: An international handbook*, 899–919. Berlin, New York: De Gruyter Mouton. DOI: 10.1515/9783110213881.2.899.

Borik, Olga & Berit Gehrke. 2018. Imperfective past passive participles in Russian. In Denisa Lenertová, Roland Meyer, Radek Šimík & Luka Szucsich (eds.), *Advances in formal Slavic linguistics 2016* (Open Slavic Linguistics 1), 53–76. Berlin: Language Science Press. DOI: 10.5281/zenodo.2545513.

Breu, Walter. 2021. Actionality and the degree of temporal dynamics (Focus aspect, status aspect, and the role of valency). *STUF – Language Typology and Universals* 74(3–4). 435–466. DOI: 10.1515/stuf-2021-1041.

Cusic, David Dowell. 1981. *Verbal plurality and aspect.* Stanford, CA: Stanford University. (Doctoral dissertation).

Dickey, Stephen M. 2015. Outline of a comparative analysis of the development of the imperfective general-factual in Slavic. In Rosanna Benacchio (ed.), *Glagol'nyj vid: grammatičeskoe značenie i kontekst (Sbornik dokladov III Konferencii aspektologičeskoj komissii, sostojavšejsja v Padue s 30.9. po 4.10.2011)*, 179–195. München: Sagner.

Gehrke, Berit. 2023. 'True' imperfectivity in discourse. In *Advances …* Berlin: Language Science Press.

Górski, Rafał L. 2008. *Diateza nacechowana w polszczyźnie (studium korpusowe)*. Kraków: Lexis.

Grønn, Atle. 2004. *The semantics and pragmatics of the Russian factual imperfective*. Oslo: Acta Humaniora.

Janda, Laura A. 2007. What makes Russian bi-aspectual verbs special? In Dagmar Divjak & Agata Kochanska (eds.), *Cognitive paths into the Slavic domain*, 83–109. Berlin / New York: Walter de Gruyter.

Klein, Wolfgang. 1994. *Time in language*. London: Routledge.

Knjazev, Jurij P. 1989. *Akcional'nost' i statal'nost': ix sootnošenie v russkix konstrukcijax s pričastijami na -n, -t*. München: Sagner.

Knjazev, Jurij P. 2007. *Grammatičeskaja semantika: Russkij jazyk v tipologičeskoj perspektive*. Moscow: Jazyki slavjanskix kul'tur.

Lehmann, Volkmar. 1992. Le prétérit déictique et le prétérit narratif en polonais moderne. In Marguerite Guiraud-Weber & Charles Zaremba (eds.), *Linguistique et slavistique: Melanges offerts à Paul Garde*, 545–557. Aix-en-Provence, Paris: Publications de l'Université de Provence 2 vol.

Lehmann, Volkmar. 1998. Eine Kritik der progressiven Funktion als Kriterium aspektueller Verbkategorisierung. *Die Welt der Slaven: Internationale Halbjahresschrift für Slavistik* 43(2). 295–306.

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

Levshina, Natalia. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam, Philadelphia: John Benjamins.

Mattiola, Simone. 2019. *Typology of pluractional constructions in the languages of the world*. Amsterdam, Philadelphia: John Benjamins.

Mehlig, Hans Robert. 2011. General questions and verbal aspect in Russian. *Scando-Slavica* 57(2). 1–24. DOI: 10.1080/00806765.2011.631779.

Mueller-Reichau, Olav. 2018. *Das Rätsel allgemeinfaktischer Interpretationen im Aspektsystem des Russischen*. Wiesbaden: Harrassowitz.

Panas, Epaminondas E. 2011. Indexes of lexical richness can be estimated consistently with knowledge of elasticities: Some theoretical and empirical results. In Peter Grzybek & Reinhard Köhler (eds.), *Exact methods in the study of language and text: Dedicated to Gabriel Altmann on the occasion of his 75th birthday*, 521–532. Berlin, Boston: De Gruyter Mouton.

Säily, Tanja & Jukka Suomela. 2017. Types2: exploring word-frequency differences in corpora. In Turo Hiltunen, Joe McVeigh & Tanja Säily (eds.), *Big and rich data in English corpus linguistics: Methods and explorations*, vol. 19.

Helsinki: VARIENG. http://www.helsinki.fi/varieng/series/volumes/19/saily_suomela/.

Šluinskij, Andrej B. 2005. *Tipologija predikatnoj množestvennosti: Količestvennye aspektual'nye značenija*. Moskva: Moskovskij gosudarstvennyj universitet imeni M. V. Lomonosova. (Doctoral dissertation).

Šluinskij, Andrej B. 2006. K tipologii predikatnoj množestvennosti: organizacija semantičeskoj zony. *Voprosy jazykoznanija* (1). 46–75.

Tatevosov, Sergei. 2015. Severing imperfectivity from the verb. In Gerhild Zybatow, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebova (eds.), *Slavic grammar from a formal perspective: The 10th anniversary FDSL conference*, 465–494. Frankfurt am Main: Peter Lang. DOI: 10.3726/978-3-653-05335-7.

Tatevosov, Sergej G. 2016. *Glagol'nye klassy i tipologija akcional'nosti*. Moskva: JaSK.

Wiemer, Björn. 1996. Analityczne passivum w języku rosyjskim i polskim. In Violetta Koseska-Toszewa & Danuta Rytel-Kuc (eds.), *Semantyka i konfrontacja językowa*, vol. 1, 167–178. Warszawa: SOW.

Wiemer, Björn. 2017. Slavic resultatives and their extensions: Integration into the aspect system and the role of telicity. *Slavia* 86(2–3). 124–168.

Wiemer, Björn. 2008. Zur innerslavischen Variation bei der Aspektwahl und der Gewichtung ihrer Faktoren. In Sebastian Kempgen, Karl Gutschmidt, Ulrike Jekutsch & Ludger Udolph (eds.), *Deutsche Beiträge zum 14. Internationalen Slavistenkongress, Ohrid 2008*, 383–409. München: Sagner.

Wiemer, Björn. 2019. Tipologija akcional'nosti: Svojstva finitnoj klauzy, klassifikacija glagolov i edinyj podxod k slovoizmenitel'nomu i derivacionnomu vidu. *Voprosy jazykoznanija* (1). 93–129. DOI: 10.31857/S0373658X0003595-3.

Wiemer, Björn & Markus Giger. 2005. *Resultativa in den nordslavischen und baltischen Sprachen: Bestandsaufnahme unter arealen und grammatikalisierungstheoretischen Gesichtspunkten*. München, Newcastle: LINCOM Europa.

Wiemer, Björn & Ilja A. Seržant. 2017. Diachrony and typology of Slavic aspect: What does morphology tell us? In Walter Bisang & Andrej Malchukov (eds.), *Unity and diversity in grammaticalization scenarios* (Studies in Diversity Linguistics 16), 239–307. Berlin: Language Science Press. DOI: 10.5281/zenodo.823246.

Wiemer, Björn, Joanna Wrzesień-Kwiatkowska & Piotr Wyroślak. 2020. How morphologically related synonyms come to make up a paradigm. *Russian Linguistics* 44(3). 231–266. DOI: 10.1007/s11185-020-09231-0.

Xrakovskij, Viktor S. 1997. Semantic types of the plurality of situations and their natural classification. In Viktor S. Xrakovskij (ed.), *Typology of iterative constructions*, 3–64. München, Newcastle: LINCOM Europa.

Zaliznjak, Anna A., Irina L. Mikaėljan & Aleksej D. Šmelev. 2015. *Russkaja aspektologija: V zaščitu vidovoj pary*. Moskva: JaSK.

Advances in formal Slavic linguistics 2021

Advances in formal Slavic linguistics 2021 offers a selection of articles that were prepared on the basis of talks given at the conference Formal Description of Slavic Languages 14 or at the satellite Workshop on Secondary Imperfectives in Slavic, which were held on June 2–5, 2021, at the University of Leipzig. The volume covers all branches of Slavic languages and features synchronic as well as diachronic analyses. It comprises a wide array of topics, such as degree achievements, clitic climbing in Czech and Polish, typology of Slavic l-participles, aspectual markers in Russian and Czech, doubling in South Slavic relative clauses, congruence and case-agreement in close apposition in Russian, cataphora in Slovenian, Russian and Polish participles, prefixation and telicity in Serbo-Croatian, Bulgarian adjectives, negative questions in Russian and German and imperfectivity in discourse. The numerous topics addressed demonstrate the importance of Slavic data and the analyses presented in this collection make a significant contribution to Slavic linguistics as well as to linguistics in general.