Advances in Formal Slavic Linguistics 2022

Edited by

Berit Gehrke

Denisa Lenertová

Roland Meyer

Daria Seres

Luka Szucsich

Joanna Zaleska

Open Slavic Linguistics

Open Slavic Linguistics

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

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 2022

Edited by

Berit Gehrke

Denisa Lenertová

Roland Meyer

Daria Seres

Luka Szucsich

Joanna Zaleska

Contents

Contents

Preface

Advances in Formal Slavic Linguistics 2022 brings together a collection of 22 articles presenting the latest developments in formal approaches to Slavic linguistics. Covering a broad spectrum of topics across all branches of linguistics, the contributions explore various phenomena in BCMS, Slovenian, Russian, Ukrainian, Polish, Czech, Upper Sorbian, Bulgarian, and Macedonian, along with their regional dialects.

The authors use a wide range of methodologies, including fieldwork, acoustic analysis, corpus studies, psycholinguistic experiments, judgment studies, and theoretical modelling. The studies examine topics such as clitics, nominalizations, l-participles, the dual, verbal prefixes, assibilation, verbal and adjectival morphology, lexical stress, vowel reduction, focus particles, aspect, multiple wh-fronting, definiteness, polar questions, negation words, and argument structure. Analytical frameworks include Metrical Phonology, Distributed Morphology, Nanosyntax, Minimalist syntax, and formal semantics.

Early versions of these articles were presented at the conference on Formal Description of Slavic Languages 15, held in Berlin on 5–7 October 2022, with a special session on formal approaches to Ukrainian. Each article underwent a thorough, double-blind peer review process. We would like to extend our sincere thanks to the reviewers, as well as to the community proofreaders. We are also grateful to Mihaela Chirpanlieva for her assistance with Bulgarian transliteration and to the Language Science Press editorial team, Sebastian Nordhoff and Felix Kopecky, as well as Radek Šimík, for their invaluable help in producing this volume.

We hope this book provides a valuable resource for linguists interested in the complexities of Slavic languages and their implications for linguistic theory.

<div align="right">

Berlin, December 12, 2024
Berit Gehrke
Denisa Lenertová
Roland Meyer
Daria Seres

</div>

Preface

Luka Szucsich
Joanna Zaleska

0

´

Chapter 1

From scope freezing to, well, everything: Investigations into the syntax of Instrumentals in Ukrainian

Svitlana Antonyuk

University of Graz

I present arguments in favor of a particular view of argument hierarchies in alternations that results from taking a deductive approach to scope freezing (Antonyuk 2015, 2020, 2023, under review and especially from relying on the Scope Freezing Diagnostic (SFD) as a source of insights in this domain. The primary focus of the paper is on several causative constructions in Ukrainian that feature an Instrumental case-marked argument. It is argued that in all constructions under consideration Instrumentals are merged in a structurally more prominent position than either the Accusative or the Dative argument, resulting in the V Instrumental >> Accusative >> Dative relative argument ordering at Merge. Next, I follow Lavine (2022, 2023) in arguing against the Undifferentiated Initiator view (Ramchand 2008; Bruening 2013; Legate 2014; Wood 2017, i.a.) and provide a new argument in support of Lavine's structural differentiation between the Instrumental Agents of passives and Instrumental Inanimate Causers.

1 Introduction

While research into argument structure alternations (ASAs) has made great advances, yielding important results over more than six decades, it has also been characterized by the vastly different, indeed, often diametrically opposing conclusions that have been reached based on largely the same types of evidence. Limiting our attention to the Slavic languages for the purposes of this discussion, we notice that the debate around the Ditransitive Alternation (exemplified in (1) and

Svitlana Antonyuk

(2) for English and Ukrainian) has witnessed accounts that posit a derivational relation between the two frames as well as those that posit independent projection of the two structures. Within these two large groups, further differences are observed based on the posited Merge position of the internal arguments, with accounts differing on, for example, whether the Dative argument is merged in the Specifier of V (Greenberg & Franks 1991, Franks 1995, Richardson 2007), the complement of V (Bailyn 1995, 2010, 2012; Antonyuk 2015, 2020; Titov 2017), in the Specifier of an Applicative head (Dyakonova 2009), as well as accounts that posit Datives base-merged at different heights (Boneh & Nash 2017).

(1)   The Ditransitive Alternation
   a. Mike gave a toy to his cat.   Prepositional Dative (PPD)
   b. Mike gave his cat a toy.   Double Object Construction (DOC)

(2)   a. Myxajlo   po-daruvav (jakus') igrašku svojij kišci.   PPD
          Myxajlo.NOM *po*-gift.PST   some   toy.ACC his   cat.DAT
          'Myxajlo gifted (some)/a toy to his cat.'
      b. Myxajlo   po-daruvav svojij kišci   (jakus') igrašku.   DOC
          Myxajlo.NOM *po*-gift.PST his   cat.DAT some   toy.ACC
          'Myxajlo gifted his cat (some)/a toy.'   (Ukrainian)

A rather underinvestigated property of the Ditransitive Alternation, familiar since Larson (1990), is the "frozen" surface scope relation between the internal arguments in the DOC, which contrasts with scope fluidity/scope ambiguity of the PPD frame, cf. (3a)–(3b).

(3)   a. Mike gave some toy to every cat.
          Surface scope: For some toy x, for every cat y, Mike gave x to y (e.g., this week).
          Inverse Scope: For every cat x, for some toy y, Mike gave x y (different cat-toy pairings possible).
      b. Mike gave some cat every toy.   (frozen surface scope)
          Surface scope: For some cat x, for every toy y, x received y from Mike (i.e., 'one happy cat' scenario).
          Unavailable inverse scope: For every toy x, for some cat y, x was given to y (different toy-cat pairings possible).

As noted in Larson (1990), Schneider-Zioga (1988) observes the arguably identical pattern of scope freezing in the 'with'-variant of the *Spray-Load* Alternation

4

(4), thus the scope ambiguity-scope freezing patterns found in the two alternations have come to be known as their characteristic, albeit fairly understudied, property.[1]

(4)   The *Spray-Load* Alternation
   a.   Mike planted the flowers in the garden.   locative frame
   b.   Mike planted the garden with flowers.   the 'with'-frame

(5)   a.   Mike planted some sort of flowers in every garden.
           Surface scope: for some sort of flowers *x*, for every garden *y*, Mike planted *x* in *y*.
           Inverse scope: for every garden *x*, for some sort of flowers *y*, Mike planted *x* with *y*.
      b.   Mike planted some garden with every sort of flowers.
                                                              (frozen surface scope)
           Surface scope: for some garden *x*, for every sort of flowers *y*, Mike planted *x* with *y*.
           Unavailable inverse scope: for every sort of flowers *x*, for some garden *y*, Mike planted *x* in *y*.

While the above scope patterns and the scope freezing in particular have come into focus in Bruening (2001), the phenomenon has effectively been explained away: Bruening takes scope freezing to be epiphenomenal, that is, a mere consequence of the purportedly Superiority-obeying nature of the covert movement operation Quantifier Raising (QR), which, combined with Richards's (1997) TUCK-ING IN, is taken to result in situations where the structurally more prominent QP takes obligatory wide scope upon QR due to the lower one obligatorily tucking in and thus scoping below it.[2] Assuming the treatment in Antonyuk (2015, 2020, under review, 2023), where scope freezing is viewed as an empirical phenomenon in its own right and thus in need of a principled theoretical explanation,

---

[1]Scope freezing is the only property of the ditransitive alternation (DA) where the mirror-image behavior of the two frames breaks down with respect to the Barss & Lasnik's (1986) diagnostics. Yet, as noted in Harley & Miyagawa (2017), it continues to be underinvestigated and poorly understood.

[2]Bruening's (2001) treatment of scope freezing is critically evaluated and ultimately rejected in Larson et al. (2019) for English and Antonyuk & Mykhaylyk (2022) for Ukrainian. See Abels & Grabska (2022) for a recent account that adopts (a modified version of) Bruening (2001) to model scope relations in Polish ditransitives and Hallman (2018, 2024) for the causative ditransitive alternation in Syrian Arabic.

Svitlana Antonyuk

the present paper will take without justification the need for the deductive approach to scope freezing underlying the above works and will proceed to rely on the Scope Freezing Diagnostic (Antonyuk 2015) yielded by this approach in order to probe underlying argument structure relations.[3] As will hopefully become apparent, this approach to scope freezing and especially the diagnostic contributed by it is highly valuable as far as diagnostics developed for Slavic languages go, able to provide non-trivial insights into the syntax of argument structure alternations and remaining remarkably consistent where other diagnostics fail to be so. Among the findings reported here is that across a number of causative constructions involving an Instrumental NP, the latter is consistently implicated by the Scope Freezing Diagnostic (SFD) to be merged in a structurally superior position relative to its VP-internal co-argument(s). Furthermore, I show that an Instrumental Inanimate Causer NP is distinguished from the (morphologically indistinguishable) Instrumental Agent NP in passives in terms of scope behavior, which supports Lavine's (2022, 2023) broader argument for the need to distinguish between agent/animate causers, which are arguments of Voice, and a variety of inanimate causers, argued to merge lower in the structure.

The remainder of the paper is structured as follows. §2 presents empirical data from Ukrainian that demonstrate the Scope Freezing Generalization and introduces the Scope Freezing Diagnostic that is based on it. §3 is concerned with mapping out the verb phrase by applying the SFD to a variety of constructions involving Accusative, Dative and Instrumental arguments. It shows, i.a., that the relative ordering of internal arguments at Merge differs significantly from that often considered "standard" in the field (§3.1). The rest of §3 then goes deeper into the syntax of Instrumentals by contrasting their scope behavior with that of the Instrumentally case-marked subjects of passives, providing new evidence against the Undifferentiated Initiator view and thus offering support to Lavine's (2022, 2023) position (§3.2.1) according to which Inanimate Causer Instrumentals are

---

[3]An anonymous reviewer challenges the terminology, pressing about what makes this approach deductive rather than inductive. Of course, the reviewer is correct to press on it, as the present approach does in fact combine elements of both. Inductively, this approach follows the data to form generalizations to form broader theory of scope freezing. The approach is however deductive in that, relying on the general minimalist set of assumptions and the newly formed theory (or, barring that, a set of strong generalizations about scope freezing), it allows us to test a range of properties under discussion, from properties of scope freezing itself, to argument hierarchy in argument structure alternations and syntactic properties of particular constructions to, e.g., information structural phenomena manifested in such constructions, such as specificity, discourse neutrality, etc. (see esp. Antonyuk 2023). To the extent that any such insights are derived from the theory of scope freezing advocated for here, they are all derived by deductive reasoning. I am grateful to the reviewer for the chance to clarify this point.

merged lower than Agents. §3.2.2 argues that the Merge position of Inanimate Causer Instrumentals is at the same time much higher than usually assumed, specifically, Spec,*v*Cause (rather than the often assumed Complement of V position). §4 offers my conclusions.

2 Scope freezing in Ukrainian: The empirical domain

While taking such an approach to scope freezing (i.e., one deducing what scope freezing is from patterns of alternating word order-QP scope distribution) makes good scientific sense in general, it is arguably especially important as far as Slavic languages are concerned, for, as shown in Antonyuk (2015) and elaborated in Antonyuk (2023), Antonyuk (under review), unlike English, where there appear to be only two constructions that exhibit frozen surface scope, the phenomenon of frozen surface scope is not only found in (East) Slavic as well, but it is found in a considerably broader range of constructions.[4] What all these constructions have in common is the pattern, schematized in (6), whereby one linearization of internal arguments is scopally ambiguous while the opposite order of arguments in the postverbal field is surface scope frozen, allowing only the QP scope interpretation that corresponds to overt c-command relations.

(6)   a. $[_{TP}$ EA [<EA> V+v $[_{VP}$ QP2 <V> QP1]]]   scope ambiguous

   b. $[_{TP}$ EA [<EA> V+v $[_{XP}$ QP1 <V> $[_{VP}$ QP2 <V> QP1]]]]

                                                      scope frozen[*,5]

         *where QP1 $\neq$ PP$_{\text{LOCATIVE/DIRECTIONAL}}$

According to (6), the conditions under which scope freezing arises are quite specific and limited, that is, scope freezing arises whenever the structurally lower internal argument raises overtly across its structurally superordinate co-argument to a c-commanding position in the postverbal field. The word order alternations below provide some of the empirical basis for (6).[6]

---

[4]The phenomena discussed here are observed in many Slavic languages across the subgroups. See Antonyuk (2015, 2020) Antonyuk (2023) and Boneh & Nash (2017) on Russian. See Antonyuk & Mykhaylyk (2022) on the interaction of quantification and Object Shift in Ukrainian. See also Marvin & Stegovec (2012) for a brief discussion of scope freezing in Slovenian DOCs and Abels & Grabska (2022) for a detailed discussion of scope distribution in Polish ditransitives.

[5]When QP1 = PP$_{\text{LOCATIVE/DIRECTIONAL}}$ the result of argument permutation is surface scope bias, not surface scope freezing. See Antonyuk (2020, under review) for a relevant discussion.

[6]See Dyakonova (2009) and Bailyn (2010, 2012) on Russian ditransitives and the discussion of the relation between the two linearizations. Both authors agree they instantiate the two frames of the Ditransitive Alternation.

Svitlana Antonyuk

(7) The Ditransitive Alternation

  a. Myxajlo   po-daruvav jakus'   igrašku kožnij   kišci.
      Myxajlo.NOM *po*-gift.PST some.ACC toy.ACC every.DAT cat.DAT
      'Myxajlo gifted some toy to every cat.'
      Surface scope: For some toy x, for every cat y, Myxajlo gifted x to y
      (i.e., the same toy).
      Inverse scope: For every cat x, for some toy y, Myxajlo gifted x y (i.e.,
      different cat-toy pairs).

  b. Myxajlo   po-daruvav jakijs'   kišci   kožnu   igrašku.
      Myxajlo.NOM *po*-gift.PST some.DAT cat.DAT every.ACC toy.ACC
      'Myxajlo gifted some cat every toy.'
      Surface scope: For some cat x, for every toy y, Myxajlo gifted x y.
      Unavailable inverse scope: For every toy x, for some cat y, Myxajlo
      gifted x to y.

(8) The *Spray-Load* Alternation

  a. Myxajlo   za-lyv   [jakyjs'   vyd   pal'noho] [v
      Myxajlo.NOM *za*-fill.PST some.ACC type.ACC gas.GEN   into
      kožen   bak].
      every.ACC tank.ACC
      'Myxajlo filled some type of gas into every tank.'
      Surface scope: For some type of gas x, for every tank y, Myxajlo filled
      x into y.
      Inverse scope: For every tank x, for some type of gas y, Myxajlo filled
      x with y (i.e., possibly different type of gas for each tank).

  b. Myxajlo   za-lyv   [jakyjs'   bak]   [kožnym vydom
      Myxajlo.NOM *za*-fill.PST some.ACC tank.ACC every.INS type.INS
      pal'noho].
      gas.GEN
      'Myxajlo filled some tank with every type of gas.'
      Surface scope: For some tank x, for every type of gas y, Myxajlo filled
      x with y.
      Unavailable inverse scope: For every type of gas x, for some tank y,
      Myxajlo filled x into y.

(9) Reflexive Causatives

   a. Likar   infikuvav-sja jakojus' xvoroboju vid kožnoho
      doctor.NOM infect.PST.REFL some.INS illness.INS from every.GEN
      pacijenta.
      patient.GEN
      'The doctor got infected from every patient with some illness.'
      Surface scope: For some illness *x*, for every patient *y*, the doctor got infected with *x* from *y*.
      Inverse scope: For every patient *x*, for some illness *y*, the doctor got inflected by *x* with *y*.

   b. Likar   infikuvav-sja vid jakohos' pacijenta kožnoju
      doctor.NOM infect.PST.REFL from some.GEN patient.GEN every.INS
      xvoroboju.
      illness.INS
      'The doctor got infected with some illness from every patient.'
      Surface scope: For some patient *x*, for every illness *y*, the doctor got infected by *x* with *y*.
      Unavailable inverse scope: For every illness *x*, for some patient *y*, the doctor got infected with *x* by *y*.

(10)   Causative verbs

   a. Likar   infikuvav jakojus' xvoroboju kožnoho pacijenta.
      doctor.NOM infect.PST some.INS illness.INS every.ACC patient.ACC
      'The doctor infected every patient with some illness.'
      Surface scope: For some illness *x*, for every patient *y*, the doctor infected *y* with *x*.
      Inverse scope: For every patient *x*, for some illness *y*, the doctor infected *x* with *y*.

   b. Likar   infikuvav jakohos' pacijenta kožnoju xvoroboju.
      doctor.NOM infect.PST some.ACC patient.ACC every.INS illness.INS
      'The doctor infected some patient with every illness.'
      Surface scope: For some patient *x*, for every illness *y*, the doctor infected *x* with *y*.
      Unavailable inverse scope: For every illness *x*, for some patient *y*, the doctor infected *y* with *x*.

The key observation to note here is, of course, that the change in scope interpretations in each pair (i.e., the appearance of scope freezing in the (b) examples) tracks the change in overt word order, which constitutes the foundation for the

Svitlana Antonyuk

Scope Freezing Generalization in (11), according to which scope freezing is an inherently derivational, in fact, *a derivation-by-movement* phenomenon.[7]

(11) The Scope Freezing Generalization (SFG) (Antonyuk 2015)
Scope freezing obtains when one QP raises overtly across another to a c-commanding position as a result of a single instance of movement within the *v*P/VP.

That scope freezing must be limited to the domain of *v*P/VP (i.e., below VoiceP in more modern terminology) becomes obvious once the above examples are compared to those involving a subject QP: Here we observe that the interaction between an object QP and a subject QP never amounts to a frozen scope relation.[8]

(12) Simple SVO transitives: No scope freezing
Jakas'   divčynka na-hoduvala kožnu   kišku.
some.NOM girl.NOM NA-feed.PST.F every.ACC cat.ACC
'Some girl fed every cat.'
Surface scope: For some girl *x*, for every cat *y*, *x* fed *y*.
Inverse scope: For every cat *x*, for some girl *y*, *y* was fed by *x*.

(13) OSV (locally scrambled) clauses: No scope freezing
Jakus'   kišku kožna   divčynka nahoduvala.
some.ACC cat.ACC every.NOM.F girl.NOM.F NA-feed.PST.F
'Some cat, every girl fed.'
Surface scope: For some cat *x*, for every girl *y*, *x* was fed by *y*.
Inverse scope: For every girl *x*, for some cat *y*, *x* fed *y*.

(14) OVS clauses: no scope freezing
Jakus'   kišku nahoduvala kožna   divčynka.
some.ACC cat.ACC NA-feed.PST.F every.NOM.F girl.NOM.F
'Some cat was fed by every girl.'

---

[7]The importance of this conclusion should be immediately clear: It implicates a much greater degree of derivationality inherent in the derivation of numerous structures and argument structure alternations than is currently assumed.

[8]Building on the original observation in Larson (1990), Bruening (2001) provides independent evidence in favor of scope freezing being limited to object QPs, never an object and a subject QP. Thus, conclusions based on the richer East Slavic empirical data are fully aligned with the conclusions based on the English data. As argued especially in Antonyuk (under review), since the scope freezing found in Ukrainian and Russian is arguably identical to the English case, the conclusions reached on the basis of the former should be viewed as naturally extending to the latter.

> *Surface scope*: For some cat *x*, for every girl *y*, *x* was fed by *y*.
> *Inverse scope*: For every girl *x*, for some cat *y*, *x* fed *y*.

While the absence of scope freezing in OVS is important, showing that object QP–subject QP pairs do not result in scope freezing under any circumstances, much will depend on one's analysis of OVS, which is a hotly contested issue in Slavic linguistics.[9] Whatever derivation of OVS one may assume, however, the basic empirical observation is that subject QP >> object QP pairs as well as object QP >> subject QP pairs do not result in scope freezing. Coupled with the observation that direct objects readily participate in frozen surface scope configurations, the conclusion this invites is that VoiceP, the layer that on most assumptions introduces the external argument, is outside the domain within which scope freezing can be established. On the assumption that OVS clauses are derived by overtly raising the object phrase to a position preceding and c-commanding the subject, the SVO–OVS examples are particularly telling in this respect when compared with, e.g., the ditransitive alternation in (7) or, e.g., (10), where the two orders are arguably also derived by an overt instance of movement raising the lower object to a position preceding and c-commanding the higher one. Whereas in the former cases the overt instance of movement still yields scope ambiguity, in the latter case what results in surface scope freezing.

Another argument in favor of the domain of scope freezing to exclude the subject layer comes from nominalizations (see esp. Alexiadou 2009, 2017; Borer 2014; see also Chomsky 1970; Grimshaw 1990; Harley 2009; Marantz 1997, i.a.). Specifically, the nominalizations in (15) and (16), which exclude the external argument, still preserve the argument structure of the corresponding verbal layer of full sentences, with the scope relations between the internal arguments being preserved as well. Thus, once again, we see scope freezing reappear on one of the two possible linearizations of internal arguments in the absence of (an overt) subject:

(15)   a.   zalyv-annja   jakohos' vydu   pal'noho v   kožen
            za-pour-*annja*.NOM some.GEN type.GEN gas.GEN   into every.ACC
            bak
            tank.ACC
            'the pouring of some type of gas into every tank'   (cf. (8a))
            Surface scope: For some type of gas *x*, for every tank *y*, there is a
            filling of *x* into *y*.

---

[9]For the sake of clarity, I assume the analysis of OVS proposed in Antonyuk (2021).

Inverse scope: For every tank *x*, for some type of gas *y*, there is a filling of *x* with *y* (i.e., possibly different type of gas for each tank).

b. zalyvannja   jakohos' baku   kožnym vydom pal'noho
   za-pour-*annja*.NOM some.GEN tank.GEN every.INS type.INS gas.GEN
   'the pouring of some tank with every type of gas'   (cf. (8b))
   Surface scope: For some tank *x*, for every type of gas *y*, there is a filling of *x* with *y*.
   Unavailable inverse scope: For every type of gas *x*, for some tank *y*, there is a filling of *x* into *y*.

(16) a. infikuvannja   jakojus' xvoroboju kožnoho   pacijenta
      infect-*annja*.NOM some.INS illness.INS every.ACC patient.ACC
      'infecting of every patient with some illness'   (cf. (10a))
      (Lit.: the infecting with some illness of every patient)
      Surface scope: For some illness *x*, for every patient *y*, there is infecting of *y* with *x*.
      Inverse scope: For every patient *x*, for some illness *y*, there is infecting of *x* with *y*.

   b. infikuvannja   jakohos' pacijenta   kožnoju xvoroboju.
      infect-*annja*.NOM some.ACC patient.ACC every.INS illness.INS
      'the infecting of some patient with every illness'   (cf. (10b))
      Surface scope: For some patient *x*, for every illness *y*, there is infecting of *x* with *y*.
      Unavailable inverse scope: For every illness *x*, for some patient *y*, there is infecting of *y* with *x*.

Of course, Complex Event Nominals (CENs), such as the examples in (15) and (16), probably do not really "exclude" the subject layer, as the data above may superficially suggest: the subject is implicit, as evidenced by the ability of such nominalizations to be modified by agentive modifiers such as *navmysne* 'deliberate', as in *navmysne infikuvannja* 'deliberate infecting'. Thus, especially within DM, on various instantiations of the "Phrasal Layering" analysis of Complex Event Nominals, the verbal structure, including the layer that introduces the external argument, is included in nominalizations, with the nominalizer head little *n* attaching on top of fully projected verbal structure (Alexiadou 2001, 2017; Bruening 2013; Borer 2014; Iordăchioaia 2020; McGinnis 2020, i.a.).[10]

---

[10]See also Wood (2023) for a detailed theoretical overview as well as an alternative proposal based on Icelandic data that CENs can be formed in syntax without nominalizing full verbal structure.

This means that the examples in (15) and (16) do not really provide bullet-proof evidence that the subject is "excluded" is any real sense, merely that it appears to be inert and not participating in whatever processes result in the establishment of the surface scope freezing relation. Minimally, the account of scope freezing adopted here (spelled out in more detail in later sections) is committed to the following conclusion: the preservation of the scope freezing relation means that the layer of structure arguably (and crucially) implicated in scope freezing, ApplP (located between vP and VoiceP), must be included in the verbal structure nominalized by n (Cf. Wood 2023).

To summarize, we see that the SFG points to a derivational nature of scope freezing, resulting from an overt instance of movement I will henceforth refer to as ARGUMENT INVERSION (following Antonyuk & Mykhaylyk 2022). Taking place in the postverbal field, Argument Inversion (AI) constitutes a local instance of overt syntactic movement, and, as examples such as the anaphor binding data below suggest, AI involves A-movement:[11]

(17)   a.   Dolja   po-daruvala nas$_i$ odyn odnomu$_i$.
            Fate.NOM *po*-gift.PST   us.ACC each other.DAT
            'Fate gifted us to each other.'

       b.   Dolja   po-daruvala nam$_i$ odyn odnoho$_i$.
            fate.NOM *po*-gift.PST   us.DAT each other.ACC
            'Fate gifted us each other.'

The anaphor binding data provides clear evidence that A-movement is involved. Considering A-movement leads to new binding relations, however, it can be argued that the directionality of the derivation in fact goes in the opposite direction. That is, rather than assuming that (17a) represents the base order and (17b) is derived from it by overt movement, it has been argued that (17b) represents the base order, and (17a) obtains from overtly raising the direct object across the Dative-marked anaphor, thus establishing a new binding relation.

While this type of argument has indeed been made many times, let me point out that it works much better with cases involving binding than it does with cases involving scope relations. This is so since the binding data in (17) is symmetric, i.e., both (17a) and (17b) can be explained on either type of account. With QP scope, things are very different. On the widely held assumption that the Dative argument in ditransitives is merged in a structurally higher position than the Accusative argument, the scope freezing familiar from the DAT >> ACC ditransitives must be seen as an inherent property of the construction or, e.g., on

---

[11]The original examples are due to Asarina (2005), cited in Bailyn (2012).

Svitlana Antonyuk

Bruening's (2001) Superiority account, as an entirely epiphenomenal property. As pointed out in Larson et al. (2019), Bruening's account thus predicts that scope freezing should be found in English in many more constructions, namely all of those where the base relation between the two QPs can be plausibly analyzed (as in Bruening 2001) as involving asymmetric c-command in the base structure, – contrary to fact. Since this prediction is thoroughly falsified, we are left with a conclusion that scope freezing is, perhaps, just an inherent property of certain constructions, at least, as far as English is concerned. That is, we are left with no way of predicting where else we might find scope freezing. And perhaps one could accept it on the grounds that this property is somehow "exceptional", i.e., found in these two ditransitive constructions only.

Drawing now on our insights from Ukrainian/East Slavic, it becomes clear that our inability to model scope freezing, to predict where else in the language we may find it, is, in fact, a problem, as the phenomenon is found in a significant number of constructions, thus posing a challenge for syntactic theory. On the view adopted here, on the other hand, not only do we have an extensive list of environments where scope freezing has already been found to obtain, the Scope Freezing Generalization allows us to predict potential further environments where scope can be expected to "freeze" in this way as well. While this result is significant, the SFG also provides us with a powerful diagnostic tool for probing argument structure relations and allows for other non-trivial insights into the derivation of the extended verbal domain in East Slavic. Thus, I conclude that the below schematization accurately describes the conditions on scope freezing for the language under discussion, Ukrainian:

(18) SURFACE SCOPE FREEZING, schematized (the domain of scope freezing highlighted):
      $[_{TP}$ ExtA $[_{VoiceP}$ <ExtA> V+v $[_{XP}$ QP1 $[_{VP}$ QP2 <V> <QP1>]]]]

In what follows then, I will be relying on the SCOPE FREEZING DIAGNOSTIC (SFD), which allows us to probe relative argument structure relations at Merge (Antonyuk 2015):

(19) THE SCOPE FREEZING DIAGNOSTIC
      Frozen surface scope implicates a derived structure resulting from Argument Inversion.

The methodological goal of this paper is to demonstrate the remarkable internal consistency of the diagnostic and the non-trivial insights into the syntax of argument structure alternations that can be gained by applying it. I will not attempt

to demonstrate all the insights already derived from the application of the SFD to the East Slavic (Ukrainian and Russian) data for reasons of space, directing the reader to the original research papers where these results are presented. Here I will only summarize and briefly exemplify some of the findings so we can build on them in this paper.[12]

3 Mapping out the Verb Phrase

3.1 The Relative Ordering of Internal Arguments at Merge

As should already be clear from our preliminary discussion, the SFD is a simple diagnostic, one which can (at most) point to a derived structure among the alternating frames under investigation (if such exists), as well as point to the relative ordering of arguments at Merge. The SFD cannot tell us much about the exact Merge position of arguments; what it does give us is a heuristic according to which the linearization that is surface-scope frozen must be the derived one, and moreover that it must be derived by raising the structurally lower of the two internal arguments overtly across the structurally higher one in the postverbal field, thus gaining c-command over it.[13,14] In other words, what the SFD gives

---

[12]See Antonyuk (2015, 2020, under review), Antonyuk (2023) and Antonyuk & Mykhaylyk (2022) for the original data and findings. See Abels & Grabska (2022) for an experimental confirmation of the empirical claims in Antonyuk's work based on Polish QP scope data as well as for a critical engagement with the account of scope freezing assumed here.

[13]Antonyuk (under review) argues that, according to the SFD (and contra Antonyuk 2015), the *Spray-Load* Alternation is, in fact, *not* a case of a derivational relation between the two alternating frames. Crucially, she argues that the 'with'-variant (i.e., the Instrumental case-marked frame in Slavic) is, nevertheless, derived, as suggested by the fact that it exhibits scope freezing, just not from the locative frame, but from the Instrumental frame via Argument Inversion of the Accusative argument across NP_{INS}. That is, scope freezing in the 'with'-variant is indeed a marker of its derived status, but what it is derived from is not what is traditionally viewed as its "alternating" frame. This, of course, is good news, as the two frames differ not only in the morphosyntactic marking on their two internal arguments, but also in their theta roles, making a derivational account problematic in this case.

[14]Significantly, argument structure alternations, in Slavic and elsewhere, appear to never exhibit scope freezing on both possible orders/alternating frames. On accounts which posit independent projection of the two alternating frames, this should in principle be a logical possibility. On Antonyuk's (2015) treatment of scope freezing, assumed here, such a situation is impossible in principle, for obvious reasons. The only conceivable exception to this would be if what were (mistakenly) considered to be an alternation consisted of two structures, each of which would be derivationally related to another structure that is not viewed as part of the alternation. The closest case to this hypothetical situation would in fact be the *Spray-Load* Alternation, where (as described in the previous footnote) the frozen frame (V NP_{ACC} >> NP_{INS}) is derivationally

Svitlana Antonyuk

us is a mere ordering of arguments at Merge. For the constructions discussed earlier, these relative ordering statements are as follows:

Table 1: Relative ordering statements

	Frozen order	Base
The Ditransitive Alternation:	DAT >> ACC	ACC >> DAT
The *Spray-Load* Alternation:	ACC >> INS	INS >> ACC
Reflexive Causatives:	PP_{FROM} >> INS	INS >> PP_{FROM}
Causative verbs:	ACC >> INS	INS >> ACC

Now, to some working on argument structure alternations and verbal argument structure more generally these insights may be surprising, as they go against a lot of what has arguably been assumed to be settled in the literature. My goal in this paper is not to provide conclusive evidence in favor of a particular structure for a particular alternation, but rather to demonstrate the general insights afforded by the SFD, especially in what concerns the relative Merge position of Instrumental arguments. A larger point, hinted at here and developed in detail elsewhere (Antonyuk (n.d.)) is that, taken together, these insights suggest a rather interesting alternative view of how the derivation of the verb phrase in Discourse Configurational Slavic languages may proceed and what confounding factors have thus far prevented us from seeing this picture.

With respect to the Instrumentals specifically, the suggestion that they might be merged higher than the Accusative argument is certainly surprising, both because it is common to assume a low Merge/complement of V position for the Instrumental NP (as is standard for the Oblique case-marked arguments), as well as because the alternative would also suggest the Accusative NP must be merged low (in any case, lower than NP_{INS}). As already suggested in Table 1, the indication that the Instrumental NP is merged higher than either the Accusative-marked co-argument or the PP co-argument remains consistent even once we significantly expand the range of the constructions under investigation.

Consider the following data (which include examples presented earlier, for convenience):

(20) Causative verbs

---

related to another structure, (i.e., V NP_{INS} >> NP_{ACC}), rather than to the locative frame. Apart from this scenario, I will venture a prediction that such a situation should be impossible.

    a. Likar   infikuvav jakojus' xvoroboju kožnoho   pacijenta.
       doctor.NOM infect.PST some.INS illness.INS every.ACC patient.ACC
       'The doctor got infected from every patient with some illness'
       Surface scope: For some illness x, for every patient y, the doctor infected y with x.
       Inverse scope: For every patient x, for some illness y, the doctor infected x with y.

    b. Likar   infikuvav jakohos' pacijenta   kožnoju xvoroboju.
       doctor.NOM infect.PST some.ACC patient.ACC every.INS illness.INS
       'The doctor infected some patient with every illness'
       Surface scope: For some patient x, for every illness y, the doctor infected x with y.
       Unavailable inverse scope: For every illness x, for some patient y, the doctor infected y with x.

⇒ BASE ORDER: V NP$_{\text{INS}}$ >> NP$_{\text{ACC}}$

According to the SFD, the V NP$_{\text{INS}}$ >> NP$_{\text{ACC}}$ linearization represents the relative order of arguments at Merge. The same conclusion is reached for the 'with'-variant of the *Spray-Load* Alternation:

(21)   Instrumental/'with'-frame of the Spray-Load Alternation

    a. Myxailo   za-lyv   jakyjs' bak   kožnym vydom
       Myxailo.NOM *za*-fill.PST some.ACC tank.ACC every.INS type.INS
       pal'noho.
       gas.GEN
       Lit: 'Myxailo filled some tank with every type of gas.'
       Surface scope (∃>∀): for some tank x, for every type of gas y, Myxailo filled x with y.
       Unavailable inverse scope (∀>∃): for every type of gas x, for some tank y, Myxailo filled x into y.

    b. Myxailo   za-lyv   jakymos' vydom   pal'noho kožen
       Myxailo.NOM *za*-fill.PST some.INS type.INS gas.GEN   every.ACC
       bak.
       tank.ACC
       Lit: 'Myxailo filled with some type of gas every tank.'
       Surface scope (∃>∀): For some type of gas x, for every tank y, Myxailo filled x into y.
       Inverse scope (∀>∃): For every tank x, for some type of gas y, Myxailo filled x with y.

Svitlana Antonyuk

$\Rightarrow$ NP$_{\text{INS}}$ >> NP$_{\text{ACC}}$ is the order at Merge within the Instrumental/'with' frame.

The passive-like *-no/-to* construction, taken up in detail in Lavine (2022, 2023) and illustrated in (22), is useful for our purposes in that it involves two internal arguments, hence it can be subject to the SFD as well.

(22) The *-no/-to* construction (Lavine & Freidin 2002)
Cerkvu   bulo spaleno   blyskavkoju.
church.ACC was burnt.NON-AGR lightning.INS
'The church was burnt down by a lightning.'

What we see is that, as before, the two arguments are permutable, and overtly permuting/crossing the arguments in the postverbal field results in scope freezing of the NP$_{\text{ACC}}$ >> NP$_{\text{INS}}$ order.[15]

(23) a. Jakus'   cerkvu   bulo spaleno   kožnoju blyskavkoju/
         some.ACC church.ACC was burnt.NON-AGR every.INS lightning.INS
         z   blyskavok.
         from lightning.GEN.PL
         'Some church was burnt down by every lighting.'
         Surface scope: For some church *x*, for every lightning *y*, *x* was burnt down by *y*.
         Unavailable inverse scope: For every lightning *x*, for some church *y*, *x* burnt down *y*.
      b. Jakojus' blyskavkoju bulo spaleno   kožnu   cerkvu.
         some.INS lightning.INS was burnt.NON-AGR every.ACC church.ACC
         'Some church was burnt down by every lightning.'
         Surface scope: For some lightning *x*, for every church *y*, *x* burnt down *y*.
         Inverse scope: For every church *x*, for some lightning *y*, *x* was burnt down by *y*.
      $\Rightarrow$ BASE ORDER: V NP$_{\text{INS}}$ >> NP$_{\text{ACC}}$

---

[15]I assume that the surface word order in such structures is derived by ultimately raising either of the arguments to Spec,TP. In the scopally frozen sentence, Argument Inversion first raises the lower argument, NP$_{\text{ACC}}$, to a position to the left of its co-argument, NP$_{\text{INS}}$, still in the postverbal field, which results in scope freezing. The thus inverted NP$_{\text{ACC}}$ then undergoes raising into Spec,TP. See Antonyuk & Mykhaylyk (2022) for evidence that scope freezing, once established by AI in the postverbal field, cannot be disturbed/"unfrozen" by further syntactic movement.

Something worth pointing out is that the surface scope interpretation for (23a) is infelicitous in that it entails a situation where one and the same church was burned down by every one of the lightning strikes under consideration. Thus, if there were five lightning strikes during the night, every single one must have hit the church and contributed to its burning down for the sentence to be true on its surface scope interpretation. Yet, despite its real world-incongruent surface scope interpretation, which should have really facilitated the inverse scope reading, if it were available, to come through, the latter is nevertheless excluded for this sentence. This helps demonstrate one of the core insights about scope freezing: it is an all-or-nothing, categorical phenomenon, as opposed to surface scope bias, familiar from other contexts, which can be manipulated by the choice of lexical items, syntactic contexts, information-structural properties of the sentence, etc. This property is what arguably makes the Scope Freezing Diagnostic an incredibly reliable, internally consistent diagnostic (see esp. Antonyuk 2023).

The exact same conclusion about the higher Merge position of the Instrumental argument (relative to the Accusative) can be reached by applying the SFD to the so-called NON-AGREEING ACCUSATIVES (Lavine & Freidin 2002; Lavine 2022, 2023, i.a.). For clarity, the derivational path for the derived ((24a) and (25a)) is assumed to be largely identical to that described for the *-no/-to* constructions in footnote 14:

(24)   a. Soldata   po-ranylo   kuleju.
           soldier.ACC *po*-wound.NON-AGR bullet.INS
           'A soldier was wounded by a bullet.'

                                                        (modeled on Lavine & Freidin 2002)

       b. Kuleju   po-ranylo   soldata.
           bullet.INS *po*-wound.NON-AGR soldier.ACC
           'A soldier was wounded by a bullet.'

(25)   a. Jakohos' soldata   po-ranylo   kožnoju kuleju.
           some.ACC soldier.ACC *po*-wound.NON-AGR every.INS bullet.INS
           'Some soldier was wounded by every bullet.' *Surface scope*: For some soldier x, for every bullet y, x was wounded with y. *Unavailable inverse scope*: For every bullet x, for some soldier y, x wounded y.

       b. Jakojus' kuleju   poranylo   kožnoho soldata.
           some.INS bullet.INS *po*-wound.NON-AGR every.ACC soldier.ACC
           'Every soldier was wounded by a bullet.'
           Surface scope: For some bullet x, for every soldier y, x wounded y.

Svitlana Antonyuk

> *Inverse scope*: For every soldier x, for some bullet y, x was wounded
> with y.
> ⇒ Base order: V NP$_{INS}$ >> NP$_{ACC}$

The non-agreeing accusatives involve two internal arguments, marked for Dative
and Accusative case, allowing us to test whether the relative ordering V NP$_{ACC}$
>> NP$_{DAT}$ obtained for the DOC will be replicated. As shown in (27), this is
indeed the case:

(26) a. Xlopcevi vidrizalo   palec'   na ruci.
         boy.DAT severed.NON-AGR finger.ACC on hand
         'The boy's finger was severed.'

      b. Palec'   na ruci vidrizalo   xlopcevi.
         finger.ACC on hand severed.NON-AGR boy.DAT
         'A finger on hand was severed from a guy's hand.'

(27) a. Jakomus' xlopcevi vidrizalo   kožen   palec'   (na ruci).
         some.DAT guy.DAT severed.NON-AGR every.ACC finger.ACC on hand
         'Some boy got every one of his fingers severed.'
         Surface scope: for some boy x, for every finger y, x had y severed from
         x's hand.
         Unavailable inverse scope: for every finger x, for some boy y, x was
         severed from y's hand.

      b. Jakyjs' palec'   (na ruci) vidrizalo   kožnomu xlopcevi.
         some.ACC finger.ACC on hand severed.NON-AGR every.DAT boy.DAT
         'Some finger was severed from every boy's hand.'
         Unavailable surface scope: For some finger x, for every boy y, x was
         severed from y's hand.
          Inverse scope: For every boy x, for some finger y, x's y was severed.
      ⇒ Base order: V NP$_{ACC}$ >> NP$_{DAT}$

Now, it is interesting that up to now we have seen two of the relevant "Merge
orderings", that is, V NP$_{ACC}$ >> NP$_{DAT}$ and V NP$_{INS}$ >> NP$_{ACC,}$ established and
replicated over several types of constructions here and elsewhere. By transitivity,
NP$_{INS}$ should precede NP$_{DAT}$ as well: V NP$_{INS}$ >> NP$_{DAT}$. Since the example in
(26) can be spelled out more fully, to include an implicit Instrumental Inanimate
Causer argument, this new relative ordering, due to transitivity, can be put to
the test by applying the SFD to the Instrumental and the Dative-marked quan-
tificational arguments. Doing so confirms that V NP$_{INS}$ >> NP$_{DAT}$ is, indeed, the
correct base structure ordering:

(28) Xlopcevi vidrizalo   palec'   elektryčnoju pyloju.
      boy.DAT   severed.NON-AGR finger.ACC electric.INS   saw.INS
      'The boy's finger got severed by an electric saw.'

(29) a. Jakomus' xlopcevi vidrizalo   palec'   kožnym
          some.DAT guy.DAT severed.NON-AGR finger.ACC every.INS
          instrumentom.
          instrument.INS
          'Some boy got a finger severed by every instrument.'   (frozen)

      b. Jakymos' instrumentom vidrizalo   palec'   kožnomu
          some.INS instrument.INS severed.NON-AGR finger.ACC every.DAT
          xlopcevi.
          boy.DAT
          'With some instrument (or other), every boy's finger was severed.'
                                                                   (ambiguous)

      ⇒ BASE ORDER: V NP$_{INS}$ >> NP$_{DAT}$

We have thus arrived at the relative "Merge ordering" for 3-argument NPs, sum-marized in (30). The SFD makes abundantly clear that the NP$_{INS}$ argument cannot be equated with other obliques in occupying the lowermost, complement posi-tion (cf. Pesetsky 1995, i.a.). In fact, according to the SFD, NP$_{INS}$ is merged the highest of the three internal arguments.

(30) V NP$_{ACC}$ >> NP$_{DAT}$
      V NP$_{INS}$ >> NP$_{ACC}$
      V NP$_{INS}$ >> NP$_{DAT}$

Thus, whatever the actual Merge positions (something the SFD does not and cannot provide an answer to), the relative structure representation we arrive at looks like the following:

(31) V NP$_{INS}$ >> NP$_{ACC}$ >> NP$_{DAT}$

3.2 The Syntax of Instrumentals in Ukrainian

While the above relative ordering is all the SFD can give us, it is plenty, of course. Supplementing with other types of evidence (see Antonyuk & Mykhaylyk 2022; Antonyuk under review for arguments based on Ukrainian data; see also Bailyn 2010, 2012; Antonyuk 2015, 2020; Antonyuk 2023; Titov 2017 for related data from Russian; Kovačević 2020 for Serbo-Croatian, i.a.), we get a fairly clear idea

of what the actual Merge positions must be. Most relevantly for our purposes, the SFD evidence provided here complements the argumentation and the analysis of Inanimate Causers and the Split Voice structure (seePylkkänen 2002, 2008). In fact, we can adduce additional supporting evidence, also due to QP scope data, that strongly supports Lavine's (2022, 2023) argumentation for the differentiated structural representation of the Instrumental case-marked Agents of passives and the Inanimate Causer Instrumentals.

3.2.1 On passives vs passive-like: Against the "Undifferentiated Initiator" view

Lavine (2022, 2023) argues against the "Undifferentiated Initiator" idea advanced in Ramchand (2008); Bruening (2013); Legate (2014) and Wood (2017). Examining a range of causative constructions in the crosslinguistic perspective, Lavine argues that the "Initiator" argument in the constructions under investigation is an Oblique Causer (Natural Force) argument that originates in the VP. This Natural Force Instrumental crucially licenses the presence of vCauseP in the structure, which licenses Accusative case on the direct object in the absence of an external argument. While I will come back to the question of the precise Merge position for the Instrumental Inanimate Causer, as this is the question where the analysis proposed here differs from Lavine's, let us for now focus on the latter point, namely the argument against the Undifferentiated Initiator position of Ramchand (2008) and subsequent work. Here, Lavine's conclusions, which he reaches based on independent types of evidence, receive strong support from the quantifier scope data presented here.[16] Perhaps the strongest argument for limiting the domain of application of the Scope Freezing Generalization to vP (i.e., crucially the layer below VoiceP) has been the absence of scope freezing in doubly quantified transitive SVO sentences (i.e., subject QP >> V >> object QP structures) as well as any other constructions involving a subject QP. Crucially, passive sentences do not exhibit scope freezing either (cf. (32) and the related OVS structure in (14)).

---

[16]Lavine (2023) takes on the theoretical question of whether Agents, Instruments and Natural Forces are all realizations of the "macro-Initiator role" and provides evidence that crosslinguistically, causer arguments are differentially realized in the syntax. Some of the strongest evidence to this effect coming from the differences in the distribution of the Inanimate (Instrumental) Causer arguments in Ukrainian, Polish, and Icelandic that sets them apart from the agents of passives in these languages.

(32) Jakas'   kiška   bula nahodovana kožnoju divčynkoju.
      some.NOM.F cat.NOM.F was fed.PST.PASS.F every.INS girl.INS
      'Some cat was fed by every girl.'
      Surface scope: For some cat x, for every girl y, x was fed by y.
      Inverse scope: For every girl x, for some cat y, x fed y.

Lavine's analysis of Transitive Impersonals (which in Ukrainian happen to be morphologically indistinguishable from the Instrumental Agent of passives) makes a strong prediction regarding QP scope. It predicts that a passive structure and a superficially similar impersonal passive construction will differ with respect to the availability of scope freezing. This is so since on Lavine's analysis the NP_{INS} in an impersonal construction is merged lower than the merge position of an Instrumental agent NP (assumed to be an argument of Voice). Hence, the former, but not the latter, is predicted to show scope freezing. Moreover, based on the data we have already seen, we can predict that transitive impersonals will show scope freezing on the ACC >> INS order of arguments. Both predictions are correct (cf. (34a) and (35a)):[17]

(33) a. Cerkv-a   bul-a   spalen-a   {#blyskavk-oju /
          church.NOM was.AGR.F burned.down.PASS-AGR.F lightning.INS
          okupant-amy}.   Passive
          invaders.INS
          (*Intended*:) 'The church was burned down by {the lightning / the invaders}.'

      b. Cerkv-u   bul-o spalen-o   {blyskavk-oju / #
          church.NOM was burned.down-NON-AGR lightning.INS
          okupant-amy}.   Trans. impers.
          invaders.INS
          (*Intended*:) 'The church was burned down by {the lightning / the invaders}.'

(34) a. Jakas'   cerkv-a   bul-a spalen-a   kožnym
          some.NOM church.NOM was burned.down.PASS-AGR.F every.INS
          okupantom.   Passive
          invaders.INS
          'Some church was burned down by every invader.' (scope ambiguous)

---

[17]As detailed in Lavine (2022, 2023), a passive and a transitive impersonal differ with respect to the type of Instrumental NP they require: thus, a passive structure is infelicitous with an Inanimate Causer argument while a transitive impersonal is significantly degraded with an animate/Agent argument. This fact provides one of the arguments in favor of a differentiated approach to Instrumental NPs in Ukrainian.

Svitlana Antonyuk

    b. Jakymos' okupantom bul-a spalen-a   kožn-a
       some.INS invader.INS was   burned.down-PASS.AGR.F every.NOM
       cerkv-a.
       church.NOM
       'Every church was burned down by some invader.' (scope ambiguous)

(35) a. Jak-us'   cerkv-u   bul-o spalen-o   kožnoju
       some.ACC church.ACC was   burned.down-NON-AGR every.INS
       blyskavkoju.   Trans. impers.
       lightning.INS
       'Some church was burned by every lightning.'   (scope frozen)

    b. Jak-ojus' blyskavkoju bul-o spalen-o   kožn-u
       some.INS lightning.INS was   burned-down-NON-AGR every.ACC
       cerkv-u.
       church.ACC
       'Every church was burned down by some lightning.'

                                          (scope ambiguous)

Of course, we want to ask whether differences regarding scope freezing avail-
ability in the two structures are indeed due to different Merge positions for
the Instrumental Inanimate Causer argument and the Agent argument of a pas-
sive sentence. Needless to say, if we could attribute the lack of scope freezing
in passives to some other factor, the above argument would disappear or be
significantly weakened. Consider the facts again. Both the subjects of active
SVO sentences and the subjects of passives exhibit lack of scope freezing, as do
all other constructions in which one of the two interacting QPs is the subject
QP. In other words, while there are indeed non-trivial structural, semantic, and
information-structural differences between, e.g., OSV, OVS, and passives, con-
spicuously, the absence of scope freezing is a property they all share. Further-
more, all internal arguments irrespective of case marking and structural height
(i.e., ACC/DAT/INSTR) invariably participate in the scope ambiguity-scope freez-
ing distribution patterns under AI, which is to say that the lack of scope freezing
cannot be attributed to some confound that is due to the presence of any of these
arguments.[18] Thus, in a passive structure (which, again, differs in terms of scope

---

[18]As mentioned earlier, the only exception here is the locative/directional PPs, which consis-
tently show surface scope bias, but not the categorical scope freezing. See Antonyuk (2020,
under review) for detailed discussions and the argument that the existence of this seeming
exception to the Scope Freezing Generalization is not a problem for the SFG, but a source of
additional insight in the search for an adequate, sufficiently restrictive account of scope freez-
ing as a grammatical phenomenon.

freezing from the superficially similar transitive impersonal), the difference can hardly be due to anything other than the status of the Agent QP.

One may also wonder whether subjects fail to participate in scope freezing relations due to any of the semantic properties that a subject QP may be associated with by virtue of its syntactic and semantic prominence (e.g., givenness, specificity, topicality, etc). I believe this is unlikely: research on the interaction of specificity and quantification in Ukrainian shows that specificity and QP scope diverge, i.e., a(n object) QP can be specific and take either high or low scope; more generally, further syntactic movement and specificity of either of the two or both objects can neither perturb existing scope freezing nor establish a new scope freezing relation once a QP raises above low temporal/manner adverbs, generally taken to mark the *v*P edge (in current terms: VoiceP edge).[19] This strongly suggests that the status of a subject QP as specific would likewise not interfere with its ability to participate in a scope freezing relation. From what I can tell, the same concerns givenness and topicality, though of course the status of the structurally higher QP as topical/given/specific can lead to some surface scope bias, meaning that the wide scope for the topical/given/specific subject QP may be the preferred interpretation in such a case.

On the Undifferentiated Initiator view all causer arguments, including the non-volitional ones, are arguments of Voice (Kallulli 2006, Ramchand 2008, Bruening 2013, Legate 2014, Wood 2017, i.a.). Thus, on these accounts, we expect that the Instrumental case-marked agents of passives and the Instrumental inanimate/ non-volitional causers of Transitive Impersonal constructions would be syntactically identical, precisely because of their being generated in the same position, Spec,VoiceP. As we have just observed, their behavior is certainly not identical as far as QP scope relations are concerned. Recall that on the account adopted here, the scope freezing relation obtains as a result of a single instance of movement of the structurally lower internal argument QP to a position above the structurally higher one, but, crucially, below the Merge position of the external argument. Thus, in the impersonal *-no/-to* constructions, scope freezing obtains when the lower NP_{ACC} raises overtly above NP_{INS}. If we assume the generalization regarding scope freezing is correct, then on the Undifferentiated Initiator view, which has all initiators merged as arguments of Voice, *all* structures involving any type of initiator argument are then predicted to lack scope freezing, precisely because, as we have shown, scope freezing emerges in a layer of structure that crucially excludes VoiceP.

Needless to say, of course, much will ultimately depend on the account of passives one adopts, and this is certainly a research area characterized by continued

---

[19]See Antonyuk & Mykhaylyk (2022) for details.

debate as well as significant differences in the general approach of individual researchers. Let us focus here on the structural, configurational accounts in which the agent phrase of a passive is base-generated in the same position as in the active, in accordance with Baker's (1988) UTAH (see esp. Collins 2005; Cf. Bruening 2013; Hallman 2021, i.a.). On Collins' (2005) influential account, a participial phase containing the VP fronts around the agent in its Merge position, thus inverting the hierarchical relation between the agent and the object phrase. On an account of passives roughly such as this one, the lack of scope freezing in passives follows naturally. This is so since the agent is merged in its usual position, Spec,VoiceP (assumed to be Spec,vP in Collins 2005), and we have seen extensive evidence that external arguments never participate in the surface scope freezing relation, thus we expect the same in passives. On the present account, which follows Lavine (2022, 2023), what sets the non-volitional/inanimate causer Instrumentals apart from the agent Instrumentals of passives is precisely the structural difference in Merge positions, with the former being merged in a structurally lower position than the latter, hence the correctly predicted differences in scope behavior.[20]

Finally, note that I have provided no independent evidence that the Agent of passives is generated in the same position as the Agent of active transitives, though, arguably, the scope data (i.e., the same scope-taking ability of passives as other structures involving the Agent QP) do provide indirect evidence to this effect. Crucially, however, to argue against the Undifferentiated Initiator view, it is enough to show that the Instrumentals in Transitive Impersonals are generated in a position other than Spec,VoiceP. I conclude that the scope difference between a passive (34) and a transitive impersonal (35) is indeed due to a difference in the structural position of the Instrumental argument in the two types of constructions under consideration. Thus, we gain a novel argument against the Undifferentiated Initiator view (Ramchand 2008; Legate 2014; Wood 2017, i.a.) and in favor of making a distinction between a higher verbal layer, VoiceP,

---

[20]A question remains, of course, why the nature of a QP's Merge position should be relevant for establishing a scope freezing relation. While I cannot do this question justice here, in my related work (Antonyuk under review; Antonyuk n.d.) I argue that scope freezing is a first-phase syntax phenomenon, which arises when the lower QP raises into the specifier of ApplP (which I argue is located in Slavic between VoiceP and vP). The overt movement of the lower argument into the specifier of ApplP is argued to define and lead to the Spell out of the first phase, the domain in which the scope freezing relation is set. Note that irrespective of whether this account is correct, what we are crucially relying on here is the empirically grounded generalization that the domain of scope freezing excludes the subject QP. Thus, the external argument, taken to be generated in Spec,VoiceP, is always going to be literally too high, being outside the domain in which this scope relation can be established.

which introduces the external argument and the lower verbal layer, vP, which introduces causative semantics and an inanimate causer argument in the absence of Voice (Lavine 2022, 2023).

3.2.2 On the VP-external Merge position of argument Instrumentals

Going back to the previous point, i.e., the exact nature of the Inanimate Causer's Merge position, as mentioned earlier, and as should be clear from the SFD insights already reviewed, I argue that the Merge position of Inanimate Causer Instrumentals cannot be VP-internal. Lavine (2022, 2023), on the other hand, posits a low VP-internal position in the complement of V, which is traditionally reserved for Oblique arguments.

As a reminder, the argument hierarchy deduced from applying the SFD to a variety of constructions is the following: V NP_{INS} >> NP_{ACC} >> NP_{DAT}. Thus, while NP_{DAT} is shown to pattern with the Obliques (i.e., arguably sharing the same Merge position with directional and locative PPs), the Inanimate Causer Instrumentals are in fact shown to have the highest Merge position of the three arguments. Hence, to the extent the SFD is accepted as being correct, the complement of V position for NP_{INS} becomes ruled out. Spec,V, on the other hand, might be somewhat more likely. Specifically, in dyadic constructions involving two internal arguments, NP_{ACC} and NP_{INS}, the former is consistently shown by the SFD to be merged lower than NP_{INS}, hence it is conceivable that the Merge position for the two arguments could simply be the reverse of what is commonly assumed.[21] This conclusion appears equally unavailable, however, once we consider again examples involving three internal arguments (ex. (28) and (29) repeated here for convenience as (36) and (37)).[22]

(36) Xlopcevi vidrizalo   palec'   elektryčnoju pyloju.
       boy.DAT   severed.NON-AGR finger.ACC electric.INS   saw.INS
       'The boy's finger got severed by an electric saw.'

---

[21]This is precisely what Landau (2010) proposes as far as the position of the Accusatives in object experiencer constructions (i.e., an Accusative as a concealed low Oblique), which in Ukrainian involve Accusative and Instrumental argument NPs.

[22]The # sign next to (38a) is meant to indicate that this example (just like (23a)) is infelicitous on its surface scope interpretation in that it describes a situation where the same finger is severed with every instrument in some contextually salient set of instruments. Our world knowledge tells us this situation is impossible or improbable, but in terms of logico-semantic properties of the sentence, this interpretation indeed obtains and is in fact the only interpretation available for this sentence.

Svitlana Antonyuk

(37) a. Jakomus' xlopcevi vidrizalo   palec'   kožnym
        some.DAT guy.DAT severed.NON-AGR finger.ACC every.INS
        instrumentom.
        instrument.INS
        'Some boy got a finger severed by every instrument.'   (frozen)

    b. Jakymos' instrumentom vidrizalo   palec'   kožnomu
        some.INS instrument.INS severed.NON-AGR finger.ACC every.DAT
        xlopcevi.
        boy.DAT
        'With some instrument (or other), every boy's finger was severed.'
                                                               (ambiguous)

    ⇒ BASE ORDER: V NP$_{INS}$ >> NP$_{DAT}$

(38) a. # Xlopcevi vidrizalo   jakyjs' palec'   kožnym
          boy.DAT severed.NON-AGR some   finger.ACC every.INS
          instrumentom.
          instrument.INS
          'Some boy got a finger severed by every instrument.'   (frozen)

    b. Xlopcevi vidrizalo   jakymos' instrumentom kožen
        boy.DAT severed.NON-AGR some.INS finger.INS   every.ACC
        palec'.
        finger.ACC
        'Some boy got every finger severed by some instrument.' (ambiguous)
    ⇒ BASE ORDER: V NP$_{INS}$ >> NP$_{ACC}$

What we see here is that when NP$_{DAT}$, NP$_{ACC}$ and NP$_{INS}$ are all arguments of the same verb, the SFD points to the same relative ordering of arguments, i.e., V NP$_{INS}$ >> NP$_{ACC}$ >> NP$_{DAT}$. The above examples, while not producing any new evidence beyond that discussed earlier in the paper, nevertheless drive home the point that there is simply no place left in the VP for the Inanimate Causer argument. Hence, my proposal (in accordance with Lavine 2022, 2023 and Pylkkänen 2002, 2008), developed in detail in related work, is that the data considered in this paper is to be interpreted as evidence in favor of the Unbundled Voice projection in Ukrainian, the lower of which, call it *v*CauseP, following Lavine (2023), hosts the Inanimate Causer Argument while the higher one, VoiceP (Kratzer 1996), hosts the Agent external argument. Without going into the specifics of the larger proposal here, I will point out that semantically, it makes good sense for a Causer argument to be merged higher than the other VP-internal arguments (NP$_{ACC}$, NP$_{DAT}$, PP); syntactically, apart from the purely theoretical reasons for

28

favoring NP$_{INS}$ as being merged in Spec,vCauseP, we now have solid empirical evidence that supports precisely this view.[23]

4 Conclusions

In this paper I have provided a novel argument against the Undifferentiated Initiator view (Ramchand 2008; Bruening 2013; Legate 2014 and Wood 2017), which posits distinct structural Merge positions for the Instrumental Agent of passives and the morphologically indistinguishable Instrumental Inanimate Causer argument of a range of "Transitive Impersonal" constructions in Ukrainian, thus providing independent support for the main theoretical claim in Lavine (2022, 2023). The novel evidence presented in this paper shows that the Instrumental agents of passives and the Instrumental Inanimate Causers (e.g., in Ukrainian non-agreeing accusative constructions) behave differently with respect to QP scope, with the latter, but not the former, being able to participate in establishing a surface scope freezing relation. In this regard passives behave very much in line with a range of transitive constructions involving a subject QP: as shown in the paper, subject QPs categorically resist participating in scope freezing (Larson 1990; Bruening 2001). Thus, while no novel evidence bearing on the Merge position of the subject of passives is provided here, the paper does provide indirect support for configurational, structural accounts of passives such as Collins (2005), by showing strong similarities with respect to non-participation in scope freezing on the part of subject QPs in SVO, OSV, SOV, OVS and passive sentences. In other words, subject QPs are outside the domain in which scope freezing can be established, and the subjects of passives in this respect behave similarly to all other external arguments of transitive sentences.

I have also provided evidence in favor of a higher Merge position than is commonly assumed for a variety of Instrumental case-marked arguments. I suggest that their Inanimate Causer semantics ensures their being merged above all other internal arguments. Specifically, I argue that the Inanimate Causer Instrumentals must be merged outside the core VP (cf. Lavine 2022, 2023 for a low complement of V position for all inanimate cause Instruments), the most likely candidate for the Merge position being vCauseP (i.e., the lower Voice projection in Pylkkänen's 2002, 2008 Unbundled Voice proposal), which introduces causative semantics in

---

[23]As mentioned earlier, Lavine (2023) treats the inanimate causer arguments in Transitive Impersonal constructions as the Natural Force argument, which is arguably well aligned with the standard view of Instrumental arguments as being merged low in the VP (cf. Szucsich 2007; Schäfer 2008; Junghanns et al. 2017; Wood 2017 i.a. for other theoretical solutions).

the absence of Voice, thus very much being in the spirit of Lavine's work (see also esp. Harley 2013).

Finally, it should be pointed out that the above arguments were all made based on the insights afforded by the Scope Freezing Diagnostic (Antonyuk 2015, 2020, under review, 2023), which is a testament to its diagnostic utility. Here as elsewhere, the SFD is shown to be a remarkably consistent diagnostic tool, one that is especially well suited for the Discourse Configurational Languages such as Ukrainian and the rest of Slavic where overt syntactic movement correlates with semantic interpretation as well as discourse-related properties of the sentence.

Abbreviations

ACC	accusative	NOM	nominative
AGR	agreement	NON-AGR	non-agreeing form
DAT	dative	PASS	passive
F	feminine	PL	plural
GEN	genitive	PST	past
INS	instrumental	REFL	reflexive

Acknowledgments

References

Abels, Klaus & Dagmara Grabska. 2022. On the distribution of scope ambiguities in Polish. *Glossa: a journal of general linguistics* 7(1). DOI: 10.16995/glossa.8170.

Alexiadou, Artemis. 2001. *Functional structure in nominals*. Philadelphia/Amsterdam: John Benjamins Publishing Company. DOI: 10.1075/la.42.

Alexiadou, Artemis. 2009. On the role of syntactic locality in morphological processes: The case of (Greek) derived nominals. In Anastasia Giannakidou & Monika Rathert (eds.), *Quantification, definiteness and nominalization*, 253–280. Oxford: Oxford University Press. DOI: 10.1093/oso/9780199541089.003.0011.

Alexiadou, Artemis. 2017. On the complex relationship between deverbal compounds and argument supporting nominals. In Maria Bloch-Trojnar & Anna Malicka-Kleparska (eds.), *Aspect and valency in nominals*, 53–82. Berlin/Boston: De Gruyter Mouton. DOI: 10.1515/9781501505430-003.

Antonyuk, Svitlana. 2015. *Quantifier scope and scope freezing in Russian*. Stony Brook, NY: Stony Brook University. (Doctoral dissertation).

Antonyuk, Svitlana. 2020. The puzzle of Russian ditransitives. In Anna Pineda & Jaume Mateu (eds.), *Dative constructions in Romance and beyond*, 43–74. Berlin: Language Science Press. DOI: 10.5281/zenodo.3776533.

Antonyuk, Svitlana. 2021. Russian OVS: towards a better understanding of the construction's properties and significance. *Journal of Slavic Linguistics* 29(FASL 28 extra issue). http://ojs.ung.si/index.php/JSL/article/view/156.

Antonyuk, Svitlana. 2023. Base-generated or derived? Here's how to tell structures apart in Russian. *Annual Review of the Faculty of Philosophy (SinFonIJA 14 special issue)*.

Antonyuk, Svitlana. under review. On methodology: The scope freezing diagnostic as a probe of argument structure relations. In Svitlana Antonyuk, Ángel Jiménez-Fernández & Isabel Oltra-Massuet (eds.), *The syntax of argument structure alternations across frameworks, Special Collection issue at Glossa: a journal of general linguistics*.

Antonyuk, Svitlana. N.d. On quantification and phase theory: A phase-based approach to the scope freezing phenomenon. Ms. University of Graz.

Antonyuk, Svitlana & Roksolana Mykhaylyk. 2022. Scope freezing and object shift in Ukrainian: Does superiority matter? Remark. *Syntax* 25. 122–146. DOI: 10.1111/synt.12229.

Asarina, Alsa. 2005. *Russian Binding Theory: Two improved movement approaches*. Ms. MIT.

Bailyn, John Frederick. 1995. *A configurational approach to Russian "free" word order*. Ithaca, NY: Cornell University. (Doctoral dissertation).

Bailyn, John Frederick. 2010. What's inside VP? New (and old) evidence from Russian. In Wayles Browne (ed.), *Formal approaches to Slavic linguistics (FASL) #18: The Cornell meeting 2009*, 21–37. Ann Arbor, MI: Michigan Slavic Publications.

Bailyn, John Frederick. 2012. *The syntax of Russian*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511984686.

Svitlana Antonyuk

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing.* Chicago, IL: Chicago University Press. https://dspace.mit.edu/handle/1721.1/15069.

Barss, Andrew & Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17(2). 347–354. https://www.jstor.org/stable/4178491.

Boneh, Nora & Léa Nash. 2017. The syntax and semantics of dative DPs in Russian ditransitives. *Natural Language & Linguistic Theory* 35. 899–953. DOI: 10.1007/s11049-017-9360-5.

Borer, Hagit. 2014. Derived nominals and the domain of content. *Lingua* 141. 71–96. DOI: 10.1016/j.lingua.2013.10.007.

Bruening, Benjamin. 2001. QR obeys superiority: Frozen scope and ACD. *Linguistic Inquiry* 32(2). 233–273. DOI: 10.1162/00243890152001762.

Bruening, Benjamin. 2013. *By*-phrases in passives and nominals. *Syntax* 16. 1–41. DOI: 10.1111/j.1467-9612.2012.00171.x.

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn & Company.

Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8. 81–120. DOI: 10.1111/j.1467-9612.2005.00076.x.

Dyakonova, Marina. 2009. *A phase-based approach to Russian free word order* (LOT Dissertation Series 230). Utrecht: LOT.

Franks, Steven. 1995. *Parameters of Slavic morphosyntax.* Oxford: Oxford Academic. DOI: 10.1093/oso/9780195089707.001.0001.

Greenberg, Gerald & Steven Franks. 1991. A parametric approach to dative subjects and the second dative in Slavic. *The Slavic and East European Journal* 35(1). 71–97. DOI: 10.2307/309034.

Grimshaw, Jane B. 1990. *Argument structure* (Linguistic Inquiry Monographs 18). Cambridge, MA: MIT Press.

Hallman, Peter. 2018. Double-object constructions in Syrian Arabic. *Syntax* 21. 238–274. DOI: 10.1111/synt.12157.

Hallman, Peter. 2021. On passive and perfect participles. In Kleanthes Grohmann, Akemi Matsuya & Eva-Maria Remberger (eds.), *Passives crosslinguistically: Theoretical and exprimental approaches*, 64–97. Amsterdam: Brill. DOI: 10.1163/9789004433427_004.

Hallman, Peter. 2024. Argument structure hierarchies and alternations in causative and double object constructions. *Glossa: a journal of general linguistics* 9(1). DOI: 10.16995/glossa.9110.

Harley, Heidi. 2009. The morphology of nominalizations and the syntax of *v*p. In Anastasia Giannakidou & Monika Rathert (eds.), *Quantification, definiteness, and nominalization*, 321–343. Oxford: Oxford University Press. DOI: 10.1093/oso/9780199541089.003.0013.

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.

Harley, Heidi & Shigeru Miyagawa. 2017. Syntax of ditransitives. In Mark Aronoff (ed.), *Oxford research encyclopedia of linguistics*. Oxford: Oxford University Press. DOI: 10.1093/acrefore/9780199384655.013.186.

Iordăchioaia, Gianina. 2020. Event structure and argument realization in English zero-derived nominals with particles. *Nordlyd* 44(1). 35–51. DOI: 10.7557/12.5205.

Junghanns, Uwe, Denisa Lenertová & Dorothee Fehrmann. 2017. Parametric variation of Slavic accusative impersonals. In Olav Muueller-Reichau & Marcel Guhl (eds.), *Aspects of Slavic linguistics: Formal grammar, lexicon and communication*, 140–165. Berlin: Walter de Gruyter. DOI: 10.1515/9783110517873-007.

Kallulli, Dalina. 2006. A unified analysis of passives, anticausatives, and reflexives. In Olivier Bonami & Patricia Cabredo Hoffherr (eds.), *Empirical issues in formal syntax and semantics*, vol. 6, 201–226. Paris: CNRS. http://www.cssp.cnrs.fr/eiss6/eiss6.pdf.

Kovačević, Predrag. 2020. Not all object experiencers are the same: The role of *se* and the argument structure of Serbian psych verbs anticausatives. *Bucharest Working Papers in Linguistics* 22(2). 77–96. https://bwpl.unibuc.ro/wp-content/uploads/2021/03/4_BWPL_2_2020_Kovacevic.pdf.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer. DOI: 10.1007/978-94-015-8617-7_5.

Landau, Idan. 2010. *The locative syntax of experiencers*. Cambridge, MA: MIT Press.

Larson, Richard. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21. 589–632. DOI: 10.4324/9780203429204-8.

Larson, Richard K., Svitlana Antonyuk & Lei Liu. 2019. Superiority and scope freezing. *Linguistic Inquiry* 50(1). 233–252. DOI: 10.1162/ling_a_00300.

Lavine, James E. 2022. Verbal layers and initiators in Ukrainian. Invited talk given at Formal Description of Slavic Languages (FDSL) 15. Berlin.

Lavine, James E. 2023. *Verbal layers and initiators in Ukrainian*. Ms. Bucknell University, Lewisburg, PA.

Svitlana Antonyuk

Lavine, James E. & Robert Freidin. 2002. The subject of defective T(ense) in Slavic. *Journal of Slavic Linguistics* 10(1/2). 251–287. http://www.jstor.org/stable/24599685.

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese.* Cambridge, MA: MIT Press. https://www.jstor.org/stable/j.ctt9qf6h9.

Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *University of Pennsylvania Working Papers in Linguistics* 4(2). 201–225. https://repository.upenn.edu/handle/20.500.14332/45440.

Marvin, Tatjana & Adrian Stegovec. 2012. On the syntax of ditransitive sentences in Slovenian. *Acta Linguistica Hungarica* 59(1–2). 177–203. https://www.jstor.org/stable/26191881.

McGinnis, Martha. 2020. Cross-linguistic contrasts in the structure of causatives in clausal nominalizations. In Bronwyn Moore Bjorkman & Daniel Currie Hall (eds.), *Contrast and representations in syntax*, 138–178. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198817925.003.0006.

Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades.* Cambridge, MA: MIT Press.

Pylkkänen, Liina. 2002. *Introducing arguments.* Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Pylkkänen, Liina. 2008. *Introducing arguments* (Linguistic Inquiry Monographs 49). Cambridge, MA: MIT Press.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon: A first-phase syntax.* Cambridge: Cambridge University Press.

Richards, Norvin. 1997. *What moves where when in which language?* MIT. (Doctoral dissertation).

Richardson, Kylie. 2007. *Case and aspect in Slavic.* Oxford: Oxford University Press.

Schäfer, Florian. 2008. *The syntax of (anti-)causatives: External arguments in change-of-state contexts* (Linguistik Aktuell/Linguistics Today 126). Amsterdam: John Benjamins. DOI: 10.1075/la.126.

Schneider-Zioga, Patricia. 1988. Double objects and small clauses. Ms. Los Angeles, CA.

Szucsich, Luka. 2007. Nothing wrong with finite T: Non-agreeing accusative impersonal sentences. In Magda Golędzinowska, Ulyana Savchenko & Richard Compton (eds.), *Formal approaches to Slavic linguistics 15*, 417–435. Ann Arbor, MI: Michigan Slavic Publications.

Titov, Elena. 2017. The canonical order of Russian objects. *Linguistic Inquiry* 48(3). 427–457. DOI: 10.1162/ling_a_00249.

Wood, Jim. 2017. The accusative subject generalization. *Syntax* 20. 249–291. DOI: 10.1111/synt.12138.

Wood, Jim. 2023. *Icelandic nominalizations and allosemy*. Oxford: Oxford University Press.

Chapter 2

Slavic creation/consumption predicates in light of Talmy's typology

Alessandro Bigolin
Centre de Lingüística Teòrica, Universitat Autònoma de Barcelona

The chapter is concerned with the licensing of creation/consumption predicates in Slavic languages, in light of Talmy's (2000) typology. I present the results of a pilot study suggesting that Slavic languages behave as verb-framed languages in the domain of creation/consumption predicates, despite these languages being commonly regarded as a type of satellite-framed languages (Talmy 2000) referred to as "weak satellite-framed languages" (Acedo-Matellán 2010; 2016). Assuming a neo-constructionist view on argument structure, I propose a morphosyntactic account of Talmy's typology according to which the verb-framed vs. satellite-framed distinction depends on a specific Phonological Form requirement, in verb-framed languages, on the null functional head v involved in verbal predication, which by assumption must incorporate its complement as an externalization condition. I propose that so-called weak satellite-framed languages, to which Slavic languages have been argued to belong, are fundamentally verb-framed languages, and that the availability of satellite-framed resultative constructions in these languages is granted by the lexical presence of result morphemes that can incorporate into v via prefixation.

1 Introduction

This chapter is concerned with the licensing of creation/consumption predicates in Slavic languages, in light of Talmy's (2000) typology. In creation/consumption predicates, the direct object is understood as being "created" or "consumed" during the event denoted by the predicate (Hale & Keyser 2002; Volpe 2004; Harley 2005; Mateu 2012, among others). For instance, *a hole* in (1a) is formed while the

Alessandro Bigolin

digging process takes place, and *the apple* in (1b) is consumed during the eating process.

(1) a. He dug a hole in his garden. (Washio 1997: 46)

  b. John ate the apple. (Folli & Harley 2005: 103)

The typological distinction proposed by Talmy (2000) divides languages into two broad classes, depending on how the Path (or "change") core component of resultative events of change of state and location is expressed.[1] In one class of languages, Path is typically encoded in a satellite (e.g., a particle, PP or AP) distinct from the main verb, which in turn may express a co-event.[2] The co-event usually provides information about the manner in which the main resultative event unfolds, or about the cause which triggers it.[3] In the other class of languages, Path is always encoded in the main verb, so that information about a co-event is either not expressed or provided via adjuncts. Languages of the former type are thus referred to as "satellite-framed", while languages of the latter type are referred to as "verb-framed". The examples from English (a satellite-framed language) and Spanish (a verb-framed language) in (2) and (3) illustrate the two patterns, for events of change of location and events of change of state respectively.

(2) Satellite-framed pattern (English):

  a. The bottle [floated]$_{\text{CO-EVENT}}$ [into the cave]$_{\text{PATH}}$. (Talmy 2000: 227)

  b. She [shot]$_{\text{CO-EVENT}}$ him [dead]$_{\text{PATH}}$. (Goldberg 1995: 136)

(3) Verb-framed pattern (Spanish):

  a. La botella [entró]$_{\text{PATH}}$ ([flotando]$_{\text{CO-EVENT}}$) a la cueva.
    the bottle enter.PST.AGR float.GER   to the cave
    'The bottle moved into the cave (floating).' (Talmy 2000: 227)

  b. Lo   [mató]$_{\text{PATH}}$ ([de un disparo]$_{\text{CO-EVENT}}$).
    ACC.M.SG kill.PST.AGR   of a   shot
    'He/she killed him with a shot.' (CORPES XXI[4])

---

[1] I consider an event "resultative" if it involves a scalar change along a scale that denotes a property or a path (Rappaport Hovav 2014).

[2] PPs are explicitly excluded in Talmy's (2000) notion of "satellite", defined as "[...] the grammatical category of any constituent other than a noun phrase or prepositional-phrase complement that is in a sister relation to the verb root" (Talmy 2000: 120). Following Mateu (2002); Beavers et al. (2010); Acedo-Matellán & Mateu (2013); Acedo-Matellán (2016), among others, I adopt a broader definition of satellite, which includes non-adjunct result PPs like the one in (2a).

[3] See Talmy (2000) for an exhaustive classification of possible conceptual interpretations attributable to co-events.

[4] *Corpus del Español del Siglo XXI*, REAL ACADEMIA ESPAÑOLA (n.d.).

Slavic languages, along with Latin, have been classified as "weak satellite-framed" (Acedo-Matellán 2010; 2016) since, although they allow the expression of Path in a satellite, this must form a prosodic word with the verb. For instance, the object *svoju ručku* 'her pen' in (4) is understood to be brought into a state where all its ink is used up which is expressed by means of the prefixal satellite *iz-* 'out', while the verb *pis-* 'write' specifies the co-event that causes the transition undergone by the referent of the direct object (Spencer & Zaretskaya 1998; Mateu 2008).

(4) Ona   [iz]$_{\text{PATH}}$-[pis]$_{\text{CO-EVENT}}$-a-l-a svoju ručku.
      she.NOM out-write-TH-PST-AGR   POSS pen.ACC
      'Her pen has run out of ink.' (Lit. 'She has written her pen out (of ink).')
                                             (Russian; Spencer & Zaretskaya 1998: 17)

The satellite-framed/verb-framed distinction is also found in the domain of predicates denoting events of creation/consumption (Mateu 2003; 2012). In a similar way to (2), satellite-framed languages allow the expression of a co-event in the verb in creation/consumption predicates, giving rise to creation/consumption predicates of the type in (5a) (hereafter, "complex creation/consumption predicates"). The predicate in (5a) can be paraphrased as "make a hole in the coat by brushing", whereby it is clear that the main verb of the predicate is understood as specifying a co-event of the main event of creation. Verb-framed languages instead consistently express the event that leads to the creation/consumption of the direct object by means of the main verb, which may be either a light verb (e.g., *make*, as in the Spanish example in (5b)) or a verb whose meaning is likely to imply the creation/consumption of the object, which in turn is interpreted as a hyponym of the verb (as in (1); see Hale & Keyser 1997; 2002). The specification of a possible co-event, as in the verb-framed change-of-location/state examples in (3), is relegated to an optional adjunct.

(5) a. Brush a hole in one's coat.   (Levin & Rapoport 1988: 279)

      b. Hizo   un agujero en su   abrigo (al   cepillar=lo).
          make.PST.AGR a   hole   in POSS coat   at.the brush.INF=ACC.M.SG
          'She made a hole in her coat, by brushing it.'

A non trivial difference between creation/consumption predicates and change-of-location/state predicates is that the argument structure of creation/consumption predicates has been argued to lack a Path component (see Rappaport Hovav & Levin 1998; Rappaport Hovav 2008; Rappaport Hovav & Levin 2010 for works adopting a lexicalist approach; see Hale & Keyser 1993; 2002; Mateu 2002; Harley 2005; Folli & Harley 2005; 2008; 2020; Ramchand 2008; Acedo-Matellán 2016,

Alessandro Bigolin

among others, for works adopting a neo-constructionist, syntactic approach). Accordingly, in light of contrasts like (5), Mateu (2012) concludes that a proper descriptive account of the cross-linguistic variation associated with Talmy's typology should not be understood in terms of a requirement about the expression of Path (either in the main verb or in a verb's satellite), but rather in terms of whether a language allows the expression of a co-event in the main verb or not.

A prediction of this line of reasoning is that weak satellite-framed languages such as Slavic languages should allow complex creation/consumption predicates of the type in (5a), since these languages more generally display constructions where the main verb expresses a co-event (as exemplified in (4)). In this chapter, I present the results of a pilot study investigating the availability of different types of creation/consumption predicates in several Slavic languages, comparing them with data from bona fide satellite-framed languages and verb-framed languages. I provide evidence suggesting that Slavic languages behave as verb-framed languages in the domain of creation/consumption predicates, as they must resort to run-of-the-mill verb-framed strategies to express such predicates and they rule out constructions such as complex creation/consumption predicates (§2). Assuming a neo-constructionist approach to argument structure (Mateu 2002; Borer 2005; Mateu & Acedo-Matellán 2012, among others), I propose a morphophonological account of Talmy's typology, which is argued to follow from a Phonological Form (PF) requirement, in verb-framed languages, on the null syntactic head v involved in verbal predication. I suggest that Slavic languages, and weak satellite-framed languages in general, should be considered as fundamentally verb-framed languages, predicting the unavailability of complex creation/consumption predicates in this class of languages (§3). Afterward, I explore the prediction – following from the present account – that a complex creation/consumption reading is available in Slavic languages for predicates that are perfectivized via so-called "internal" verbal prefixes (Svenonius 2004; Borik 2006; Gehrke 2008, among others), which have been argued to express an abstract result in a resultative construction (Gehrke 2008; Acedo-Matellán 2016; Kwapiszewski 2022, among others) (§4). Finally, I address some potential counterexamples from Latin (another weak satellite-framed language; Acedo-Matellán 2016) to the prediction that weak satellite-framed languages lack complex creation/consumption predicates of the type found in satellite-framed languages. I argue that Latin lacked such predicates in the same way as Slavic languages do, *pace* Acedo-Matellán (2016) and consistently with the predictions of the present account (§5). I draw general conclusions in §6.

2 Creation/consumption predicates in Slavic languages

In order to investigate the availability of complex creation/consumption predicates in Slavic languages, I carried out a pilot study to check, with the help of native speakers, whether it is possible to directly translate different creation/consumption predicates that are licensed in satellite-framed English into several Slavic languages. I further examined whether it is possible to directly translate the English examples into four additional bona fide satellite-framed languages and five verb-framed languages, respectively. Effort was invested in gathering evidence from different language families, contributing to the diversity of languages represented in the collected data. For the class of satellite-framed languages, data were collected from Dutch, German, Chinese, and Hungarian. Regarding verb-framed languages, data were collected from Italian, Catalan, Spanish, Basque, and Greek. Finally, for the class of Slavic languages, data were collected from Russian and Ukrainian (East Slavic languages), Polish and Slovak (West Slavic languages), and Serbian and Croatian (South Slavic languages).[5]

2.1 The English data

The English examples range from constructions involving verbs whose meaning can be taken to imply the creation/consumption of the direct object, therefore using a verb-framed strategy, to constructions that can be taken to involve the expression of a manner co-event in the main verb, and which are expected to be ungrammatical in verb-framed languages.[6] The list of the selected examples, starting with verb-framed constructions, is provided in (6) to (24).[7]

---

[5]Serbian and Croatian are considered individually alongside the other languages examined, notwithstanding classifications that see them as distinct varieties of a single language (e.g., Serbo-Croatian, or BCMS).

[6]The selection of the data was primarily based on examples from relevant literature pertaining to hyponymous objects, effected objects, and Talmy's typology. Additionally, some examples were taken from corpora or made up and subsequently checked with native speakers. Following Mateu (2002), I have included the examples in (18) and (19) as representatives of the class of complex consumption predicates, where the consumption of the direct object constitutes the main event denoted by the predicate while the verb denotes a co-event. See Kuno (1973) and Condamines (2013) for possible examples of this type in verb-framed Japanese and French, respectively (I thank an anonymous reviewer for bringing my attention to the data analysed in these works, which deserve further investigation).

[7]The examples have been arranged in the present order based on my own intuitions, as a native speaker of one of the verb-framed languages tested, about the degree of "manner" provided by the verb in each of them. Determining the degree of manner provided by the verb in each of the sentences in (6) to (24) is a complex process that takes place at the conceptual level. To the

Alessandro Bigolin

(6)   John sang a song.   (Truswell 2007: 1361)

(7)   They danced a Sligo jig.   (Gallego 2012: 98)

(8)   Ariel ate the mango.   (Ramchand 2008: 52)

(9)   He dug a hole in the ground.   (COCA[8])

(10) She wove the tablecloth.   (adapted from Folli & Harley 2020: 452)

(11) Marco painted a sky.   (Folli & Harley 2020: 438)

(12) Maria carved a doll.   (Folli & Harley 2020: 439)

(13) She burned a hole in her coat.   (made up)

(14) He scratched a hole in the ground.   (COCA)

(15) She punctured a wound in her finger.   (made up)

(16) She cut a wound in her foot.   (made up)

(17) She bit a hole in the bag.   (COCA)

(18) The adventurer walked the trail.
                              (Mateu 2002: 297, adapted from Tenny 1994: 17)

(19) The adventurer swam the channel.
                              (Mateu 2002: 297, adapted from Tenny 1994: 17)

(20) Deanne kicked a hole in the wall.   (COCA)

(21) She magicked a cursor.   (COCA)

(22) She brushed a hole in her coat.
                        (Mateu & Rigau 2002: 213, adapted from Levin & Rapoport 1988)

(23) John smiled his thanks.
                           (Mateu 2012: 255, adapted from Levin & Rapoport 1988)

---

best of my knowledge, there is currently no objective method to quantitatively measure the degree of manner provided by the verb in a specific construction, leaving the intuition-based approach as the only viable option.

[8]Corpus of Contemporary American English (Davies 2008-).

(24) Elna frowned her discomfort. (Acedo-Matellán & Kwapiszewski 2021: 35)

All the examples in (6) to (24) are taken to lack a Path component in their argument structure. While this is proposed by much work adopting both the lexicalist approach and the neo-constructionist approach (as pointed out in §1), such work is mostly concerned with the argument structure of (verb-framed) predicates in which the meaning of the verb can be taken to imply the creation/consumption of the object. Following Mateu (2012); Acedo-Matellán (2016); Folli & Harley (2020), among others, I extend such an analysis to satellite-framed predicates of creation/consumption in which the verb is taken to express a co-event. At first sight, predicates of this type might be argued to involve the argument structure of resultative predicates since most of them typically require a locative PP which is instead omissible in predicates of creation/consumption that involve a verb-framed strategy. See, in this respect, the contrast between (1a), assumed to be verb-framed, and (5a), repeated in (25a) and (25b), respectively.[9]

(25) a. He dug a hole (in his garden).
      b. Brush a hole *(in one's coat).

Based on the contrast in (25), the current assumption that satellite-framed predicates denoting events of creation/consumption do not involve a Path component in their argument structure might be questioned. Specifically, an anonymous reviewer suggests that the PP could be expressing a null Path in English predicates of the type in (5a) in the same way as it seems to do in predicates denoting events of change such as *walk in the room*, considered by the reviewer to be ambiguous between a locative and a change-of-location reading (but see, e.g., Folli & Ramchand 2005: 83 and Gehrke 2008: 90 for a different opinion).

The remainder of this subsection is devoted to showing that satellite-framed predicates of creation/consumption should not be taken to involve a null Path element in their argument structure. I argue that several reasons support this conclusion, even though the contrast in (25), at first sight, might seem to suggest otherwise. First, the claim that the PP in *walk in the room* involves a phonologically null Path is disputable since Path, in such a predicate, has been argued in previous works to be expressed by the verb *walk* (Alexiadou 2015; further see Ramchand 2008: 112, fn. 1; Nikitina 2008; Beavers et al. 2010). This verb, given the right context, may be coerced by some speakers into an interpretation as involving directionality and hence goal of motion. This explains the existence of contrasts like the one depicted in (26). Unlike *walk*, *dance* denotes an activity

---

[9]The judgement in (25b) is by an anonymous reviewer.

that typically does not imply directionality. As a result, this verb is less likely to express Path, which must therefore be expressed independently in order for the verb to appear in the change-of-location frame.

(26)   a. John walked in the room.   (in a change-of-location reading)
       b. #John danced in the room.   (in a change-of-location reading)
                                             (Alexiadou 2015: 1093)

Additionally, if the satellite-framed predicates of creation/consumption discussed in this chapter involved a phonologically null Path, the question would arise as to why Path *must* be null in these predicates. Even by assuming that (26a) is compatible with a change-of-location reading, Path can optionally be overtly realized independently of the verb in resultative predicates of this type, as (27) shows.

(27)   John walks in(to) the room.   (in a change-of-location reading)

More strikingly, Path is mandatorily realized by a morpheme different from the verb in transitive resultatives featuring direct objects that are not semantically selected by the verb (meaning that they are not a traditional object of the verb based on what lexicalist approaches consider to be the verb's lexical argument structure, and would not be suitable objects of such a verb outside the resultative construction); see the contrast between the example in (28a) and the one in (28b), both examples displaying direct objects that are not semantically selected by their respective verb. In (28b), which involves a bona fide resultative predicates, the presence of an overt Path (*to*) is mandatory. This is not the case in (28a), in contrast to what one would expect if the predicate in (28a) was resultative.[10]

(28)   a. Brush a hole in(#to) one's coat.
       b. The children run themselves *in/(in)to exhaustion.   (Iwata 2020: 281)

A further piece of evidence against considering the locative PP in satellite-framed predicates of creation/consumption as containing a null Path comes from the observation that such a PP can also be headed by the preposition *at*, as shown in (29). Unlike *in*, *at* is only compatible with a non-directional reading and is in complementary distribution with *to*. This strongly suggests that there is no null Path in the locative PPs found in the examples considered in this study.

---

[10]Arguably, a literal interpretation of the predicate in (28a) could be considered grammatical with the presence of *to*, but pragmatically aberrant, as the predicate would be interpreted as roughly meaning 'move a hole to the inside of one's coat using a brush-like object / in a brush-like manner' (Jaume Mateu, p.c.).

(29) a. They removed the coriaceous bracteoles wrapped outside of the corolla, bit a hole at the base of the corolla where the nectarines are located, and lapped up all the nectar in each flower.   (Web)

  b. To really make it resemble a tea bag, Murphy punched a hole at the top, then added a length of twine and a "tag".   (COCA)

This said, that the locative PP can be omitted in (25a) but not in (25b) is not necessarily due to grammatical reasons. Other factors, e.g., conceptual/pragmatic ones, might be involved. Note that only (25b) involves a direct object which is not semantically selected by the verb. *Brush* is a verb of surface contact, and it typically appears with direct objects denoting the surface that is brushed. It can then be expected that the *v*P in (25b) requires additional contextual information in order to be interpreted under a creation reading. In the absence of the spatial PP *in the coat*, the default inferable reading would be the pragmatically aberrant (not ungrammatical, in my view) one in which *a hole* is a selected object of *brush* (that is to say, it is an existing entity that undergoes an event of *brushing*). Such a reading disappears when the locative PP is added, as the PP introduces the semantic argument of the verb (i.e. the surface which is brushed, e.g., *her coat*), favoring the interpretation of the direct object *a hole* as an effected object thanks to the additional context. Further notice, in this respect, that locative PPs do not always appear in predicates of this type. For instance, no locative PP appears in the complex creation predicates in (21), (23) and (24), nor in the complex consumption predicates in (18) and (19). I suggest that in these predicates, the intended creation/consumption reading arises based on world knowledge/pragmatic considerations regarding the scene denoted by the event which are clear enough without the necessity of additional contextual information.[11]

2.2 Method and results

The examples in (6) to (24) were presented to the speakers in a randomized order. Translations, glosses, and grammaticality judgments were collected by consulting one linguist native speaker per language.[12] For each of the examples tested,

---

[11]This is in contrast to resultative predicates like (28b), where the licensing of a direct object that is not semantically selected by the verb always requires the presence of a phrase (e.g., a result PP) acting as a secondary predicate. Such a contrast can be taken to reflect the different status of the PPs appearing in complex creation/consumption predicates and the result PPs appearing in resultative predicates with non-selected objects, the former being adjuncts while the latter are arguments of the predicate.

[12]One exception is the native speaker of Ukrainian, who is not a linguist but who is a proficient speaker of English.

Alessandro Bigolin

it has been ensured that the intended (creation/consumption) meaning of the predicate was clear to the speakers before soliciting a grammaticality judgment. Two caveats were further considered in gathering judgments from the native speakers of the Slavic languages selected. First, considering that, as I will discuss in §4, perfective aspect in many Slavic languages is achieved through prefixes which have been argued to play a role in the event domain and interfere with the data being analyzed, the English examples were presented in the imperfective aspect when soliciting corresponding translations from the native speakers of the Slavic languages tested. For instance, the availability of the English example in (6) was checked in Slavic languages using the imperfective construction *John was singing a song*. Additionally, the speakers were asked to provide translations involving unprefixed verbs only. As a second caveat, when possible, the availability of a transitive non-creation use of those verbs which gave rise to ungrammatical translations in the languages tested was checked for each language, in order to exclude possible cases of ungrammaticality due to unrelated lexical restrictions on the transitivity of the verbs involved.[13]

The results obtained are graphically summarized in Table 1, Table 2, and Table 3 for satellite-framed languages, verb-framed languages, and Slavic languages, respectively.[14] The data collected are provided in the Appendix.

Overall, the native speakers of the satellite-framed languages tested accepted a literal translation for the vast majority of the complex creation/consumption predicates provided from English (Table 1), consistently with Talmy's typology.[15]

---

[13]Such a non-creation use pertains to transitive predicates where the direct object is understood as a pre-existing entity which undergoes the action named by the verb, and is not created or consumed during the event. Compare, for instance, (12) with *Maria carved the wood* (Folli & Harley 2020: 439), where the direct object pre-exists the carving event and undergoes the change of state specified by the verb.

[14]In the tables, empty slots correspond to cases where a direct translation of the English verb is not available in the target language. For reasons of space, the languages examined are identified in the tables using the ISO 639-2/B standardized nomenclature (US Library of Congress).

[15]I assume that Mandarin Chinese is a standard satellite-framed language of the English type. Acedo-Matellán (2016) argues that some varieties of Chinese are weak satellite-framed because the satellite-framed constructions they display present the Path and the co-event components as univerbated in a sort of V-V compound (see also Fan 2014). However, the idea that the Path and the co-event components in Chinese resultatives form a complex head is disputed. For instance, Wang (2010) presents evidence of phrasal elements that may intervene between the two members of the V-V compound in Chinese resultatives. We can see this in (i), where the complex negation *bu tai* 'not too' disrupts the adjacency between *da* 'hit' and *si* 'die'.

(i) Wo da bu tai si na   zhi zhanglang.
     I   hit NEG too die that CL cockroach
     'I can hardly hit the cockroach to death.'   (Chinese; Wang 2010: 38)

Table 1: Creation/consumption predicates in satellite-framed languages

Example	Dut	Ger	Chi	Hun
(6) John sang a song	✓	✓	✓	✓
(7) They danced a Sligo jig	✓	✓	✓	✓
(8) Ariel ate the mango	✓	✓	✓	✓
(9) He dug a hole in the ground	✓	✓	✓	✓
(10) She wove the tablecloth	✓	✓	✓	✓
(11) Marco painted a sky	✓	✓	✓	✓
(12) Maria carved a doll	✓	✓	✓	✓
(13) She burned a hole in her coat	✓	✓	✓	✓
(14) He scratched a hole in the ground	✓	✓	✓	✓
(15) She punctured a wound in her finger	✓	✓	✓	★
(16) She cut a wound in her foot	✓	✓	✓	✓
(17) She bit a hole in the bag	✓	✓	✓	✓
(18) The adventurer walked the trail	✓	✓	✓	??
(19) The adventurer swam the channel	★	★	★	★
(20) Deanne kicked a hole in the wall	✓	✓	✓	✓
(21) She magicked a cursor	??		✓	✓
(22) She brushed a hole in her coat	✓	✓	✓	✓
(23) John smiled his thanks	★	★	★	★
(24) Elna frowned her discomfort	★			★

The results obtained from the native speakers of the verb-framed languages tested are considerably different when it comes to predicates that are understood as involving the expression of a co-event by the verb (Table 2). A literal translation of the English examples gets progressively more difficult to obtain in the verb-framed languages as the predicates shift from a verb-framed strategy (the verb implying the creation/consumption of the object) to a satellite-framed strategy (the verb being understood as specifying a co-event of the main event of creation/consumption), in accordance with the typology.

As Table 3 makes clear, Slavic languages behave on a par with verb-framed languages in disallowing creation/consumption predicates where the meaning of the verb cannot be taken to imply the creation/consumption of the entity denoted by the object. The literal translations in (30) of the satellite-framed example in (22) (also in (5a)) in Russian, Ukrainian, and Polish illustrate this.

(30)   a. * Ona   čes-a-l-a   dyrku   v pal'to.   (Russian)
            she.NOM brush.IPFV-TH-PST-AGR hole.ACC in coat.LOC

Alessandro Bigolin

Table 2: Creation/consumption predicates in verb-framed languages

Example	Ita	Cat	Spa	Baq	Gre
(6) John sang a song	✓	✓	✓	✓	✓
(7) They danced a Sligo jig	✓	✓	✓	✓	✓
(8) Ariel ate the mango	✓	✓	✓	✓	✓
(9) He dug a hole in the ground	✓	✓	✓	✓	✓
(10) She wove the tablecloth	✓	✓	✓	✓	✓
(11) Marco painted a sky	✓	✓	✓	✓	✓
(12) Maria carved a doll	✓	✓	✓	✓	✓
(13) She burned a hole in her coat	★	★	★	✓	★
(14) He scratched a hole in the ground	★	★	★	✓	✓
(15) She punctured a wound in her finger	??	★	★	??	★
(16) She cut a wound in her foot	★	★	★	??	★
(17) She bit a hole in the bag	★	★	★	?	★
(18) The adventurer walked the trail	★	★	?	✓	✓
(19) The adventurer swam the channel	★	★	?		✓
(20) Deanne kicked a hole in the wall	★		★	★	★
(21) She magicked a cursor					★
(22) She brushed a hole in her coat	★	★	★	★	★
(23) John smiled his thanks	★	★	?		?
(24) Elna frowned her discomfort	★	★	★		★

b. * Vona   ter-l-a   dyrku   na kurtci.   (Ukrainian)
    she.ɴᴏᴍ brush.ɪᴘꜰᴠ-ᴘꜱᴛ-ᴀɢʀ hole.ᴀᴄᴄ in coat.ʟᴏᴄ

c. * Ona   czes-a-ł-a   dziurę   w płaszczu.   (Polish)
    she.ɴᴏᴍ brush.ɪᴘꜰᴠ-ᴛʜ-ᴘꜱᴛ-ᴀɢʀ hole.ᴀᴄᴄ in coat.ʟᴏᴄ
    Intended: 'She was brushing a hole in her coat.'

In such cases, a verb-framed construction displaying a verb whose meaning implies the creation/consumption of the direct object has to be used instead, the manner co-event being optionally expressed as an adjunct.[16]

---

[16]The results obtained further warn against making generalizations about the typological behavior of a language based on individual examples. For instance, the example in (12) seems to be generally available in the verb-framed languages examined, but it does not fare well in Slavic languages such as Russian, Ukrainian, and Slovak. Instead, the example in (14) presents a high degree of variation both in verb-framed languages and in weak satellite-framed Slavic languages, as it is accepted in half of the Slavic languages and in two of the five verb-framed languages examined. Additionally, none of the native speakers of the satellite-framed languages checked seems to accept the example in (19), even though they accept the similar example in (18) and even though (19) is accepted by the native speaker of verb-framed Greek. It is also

2 *Slavic creation/consumption predicates in light of Talmy's typology*

Table 3: Creation/consumption predicates in Slavic languages (imperfective, unprefixed predicates)

Example	Rus	Ukr	Pol	Slo	Ser	Hrv
(6) John sang a song	✓	✓	✓	✓	✓	✓
(7) They danced a Sligo jig	✓	✓	✓	✓	✓	✓
(8) Ariel ate the mango	✓	✓	✓	✓	✓	✓
(9) He dug a hole in the ground	✓	✓	✓	✓	✓	✓
(10) She wove the tablecloth	✓	✓	✓	✓	✓	✓
(11) Marco painted a sky	✓	✓	✓	✓	✓	✓
(12) Maria carved a doll	★	★	✓	?	✓	✓
(13) She burned a hole in her coat	★	✓	?	??	★	✓
(14) He scratched a hole in the ground	✓	✓	★	✓	??	★
(15) She punctured a wound in her finger	★	✓	★	★	??	★
(16) She cut a wound in her foot	★	★	★	★	★	★
(17) She bit a hole in the bag	★	★	★	★	?	★
(18) The adventurer walked the trail	★	★	★	★	??	★
(19) The adventurer swam the channel	★	★	★	★	??	★
(20) Deanne kicked a hole in the wall	★	★	★	★	★	★
(21) She magicked a cursor	★	★	★	★	★	✓
(22) She brushed a hole in her coat	★	★	★	✓	★	★
(23) John smiled his thanks						
(24) Elna frowned her discomfort						

(31)   a.   Ona   del-a-l-a   dyrku   v pal'to   ščëtkoj.   (Rus)
          she.NOM make.IPFV-TH-PST-AGR hole.ACC in coat.LOC brush.INS

     b.   Vona   rob-y-l-a   dyrku   na kurtci   ščitkoju.
          she.NOM make.IPFV-TH-PST-AGR hole.ACC in   coat.LOC brush.INS
          (Ukr)

     c.   Ona   rob-i-ł-a   dziurę   w płaszczu szczotką.   (Pol)
          she.NOM make.IPFV-TH-PST-AGR hole.ACC in coat.LOC brush.INS
     'She was making a hole in her coat with a brush.'

In the next section I propose a formal account of the patterns observed in terms of a PF requirement holding of the functional head *v* involved in verbal predicates.

---

worth noticing that the examples in (23) and (24), despite being well-formed in English, do not fare well in any of the other satellite-framed languages tested according to the native speakers consulted. Arguably, some level of idiomaticity is present in these two constructions of English, which is not shared by the speakers of the other satellite-framed languages tested. Further similar irregularities are detected, which nonetheless do not affect the emergence of clear trends consistent with the predictions following from Talmy's typology.

Alessandro Bigolin

The requirement is argued to affect both verb-framed and weak satellite-framed languages, explaining the uniformity of results observed in these languages.

3 A morphophonological account of Talmy's typology

3.1 A syntactic approach to argument structure

I adopt a neo-constructionist view of argument structure along the lines of Mateu & Acedo-Matellán (2012), according to which argument structure is conceived of as consisting of the relations established between a head and its arguments (i.e. its specifier and complement) in syntax. A fundamental distinction is drawn between functional heads, which are abstract relational elements that are necessary for the building of syntactic structures, and roots, regarded as units of conceptual content that provide real world details to syntactic predicates and are devoid of grammatically relevant information (Mateu 2002; Borer 2005; Acedo-Matellán 2010; 2016, among others).

In this approach, satellite-framed constructions are understood as involving the conflation, i.e. e(xternal)-merge (Haugen 2009), of a root with a phonologically null verbal head *v*, whose complement receives a morphological realization independently of the verb. The root conflated with *v* is understood as specifying a co-event of the main event arising from the predicate (Embick 2004; Harley 2005; Mateu & Acedo-Matellán 2012; Ausensi & Bigolin 2023, among others). In the case of resultative (change-of-state/location) predicates, such as (32) (whose syntactic structure is illustrated in Figure 1), *v* takes a small clause as complement (PredP in Figure 1), where the undergoer of the transition and the final state/location are introduced (Hoekstra 1988).[17]

(32)   The bottle floated into the cave.   (Talmy 2000: 227)

Verb-framed languages are different from satellite-framed languages in that they never show the conflation pattern depicted in Figure 1 (Mateu 2012). In verb-framed languages, the predicative complement of the small clause always forms a unit with the *v* head, whereby the only resultative predicates attested are those formed via incorporation (Mateu 2002; 2017; Mateu & Rigau 2002; Folli & Harley 2020, among others).[18] The syntactic argument structure in Figure 2, relative to the Spanish verb-framed change-of-location predicate in (33), illustrates this.

---

[17]In the structures, I represent roots with small capitals, following Acedo-Matellán (2016).

[18]Following Hale & Keyser (2002) and Mateu & Rigau (2002, 2010), I consider overt PPs expressing the final location of change-of-location events in verb-framed predicates (e.g., *a la cueva* in (33)) as hyponymous arguments that further specify the result provided by the root that
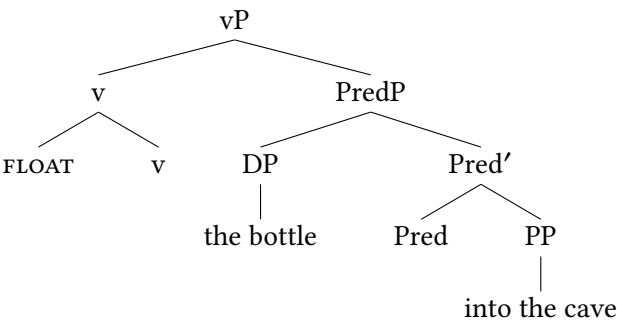
Figure 1: Syntactic structure of (32)

(33)   La botella entró (flotando) a la cueva.
        'The bottle entered the cave (floating).'   (Talmy 2000: 227)
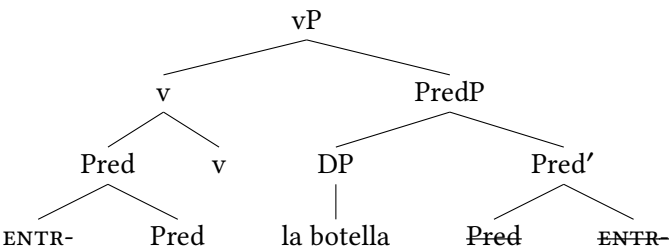

Figure 2: Syntactic structure of (33)

As for creation/consumption predicates, these are argued to involve an unergative configuration (à la Hale & Keyser 1993) consisting of a *v* head that takes as its complement either a root, which subsequently incorporates into it (the overt object emerging as a hyponym of the verb; Hale & Keyser 1997; 2002), or an independent DP. In the latter case, *v* may either appear as an overt light verb (e.g., *make*, as in the Spanish example in (5b)) or conflate with another root, giving rise to the complex creation/consumption predicates that are peculiar to the satellite-framed languages (Mateu 2012). The root incorporation pattern, corresponding to predicates of the type in (1) (see, e.g., (1a), repeated in (34)), is represented in

---

incorporates into *v*. In the syntactic structures, hyponymous arguments are omitted for ease of exposition. For discussion of possible syntactic representations of hyponymous arguments, see Hale & Keyser (1997, 2002); Mateu (2008); Haugen (2009); Gallego (2012); Real-Puigdollers (2013), among others.

Alessandro Bigolin

Figure 3. The pattern involving conflation is shown in Figure 4, which represents the syntactic structure of (5a) (repeated in (35)).[19]

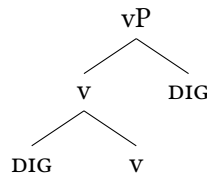(34) He dug a hole in his garden.   (Washio 1997: 46)

vP
├── v ─── DIG
│   └── DIG ─── v

Figure 3: Syntactic structure of (34)

(35) She brushed a hole in her coat.
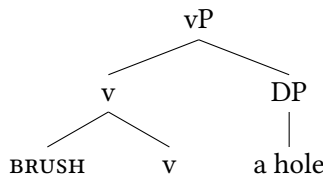   (Mateu & Rigau 2002: 213, based on Levin & Rapoport 1988)

vP
├── v ─── DP
│   ├── BRUSH ─── v
│   └── a hole

Figure 4: Syntactic structure of (35)

At first sight, the presence vs. absence of the operation conflating a root with *v* in the syntax of a given language might seem to successfully account for the language's behavior with respect to Talmy's typology. However, there are at least two reasons, one theoretical and one empirical, why the availability of this syntactic operation in a given language cannot be taken as such as an effective way of explaining the typology. On the theoretical side, as noted in Folli & Harley (2020), parameterizing the availability of a specific syntactic operation comes at the cost of giving up on the basic minimalist assumption that variation is not located in narrow syntax. On the empirical side, the results presented in §2 show that correlating Talmy's typology with the presence vs. absence of the syntactic

---

[19]The spatial PPs in (34) and (35) are treated as *v*P-external adjuncts (see also footnote 11) and are omitted from the syntactic representations for ease of exposition. For the same reason I omit the representation of the external argument, which, following considerations in Marantz (1984); Kratzer (1996); Pylkkänen (2008), among others, I assume to be introduced by a functional head Voice merged on top of the *v*P.

operation conflating a root with v leads to a wrong prediction when it comes to the possibility of licensing complex creation/consumption predicates in weak satellite-framed languages like Slavic languages (see Table 3).

In what follows, I propose an account of Talmy's typology which locates the source of the cross-linguistic variation at the PF level, understanding it in terms of differing morphophonological realization conditions of individual functional and lexical items. Not only does such an account seem to make the correct predictions with respect to the relevant patterns of cross-linguistic variation, it also provides a solution to the conundrum whereby verb-framed languages appear to consistently lack a structure building operation (*viz.* the conflation of a root with v) that is instead available in satellite-framed languages.

3.2 A PF requirement on the v head in verb-framed languages

I endorse a view of cross-linguistic variation as primarily consisting in differing morphophonological realization conditions of functional heads (Acedo-Matellán 2016; Mateu 2017, among others). In order to account for the variation observed in relation to Talmy's typology, I posit that the v head in verb-framed languages is associated with a PF requirement which imposes the incorporation of v's complement into v when v is phonologically null.[20]

(36) *Verb-framed languages' PF requirement*:

A phonologically null v must incorporate its complement.

The requirement in (36) predicts that the typological patterns noted by Talmy hold regardless of whether a result component is involved (as in the case of change-of-location/state predicates) or not (as in creation/consumption predicates). This is so because the v head is found in both resultative predicates and creation/consumption predicates, as discussed in §3.1.[21] In Figure 5 and Figure

---

[20]The requirement in (36) may ultimately be understood as an instance of Arregi & Pietraszko's (2021) "Generalized Head Movement" (GenHM) operation. This operation is captured by Arregi & Pietraszko (2021) by means of a feature [hm] on syntactic heads which, when present, requires them to form a single morphological word with the closest head of their complement. Although Arregi & Pietraszko (2021) formalize GenHM as a syntactic operation, they leave open the possibility that such an operation is carried out in the PF branch of the derivation (see Kwapiszewski 2022 for a PF implementation of GenHM). I am grateful to Víctor Acedo-Matellán for drawing my attention to the work of Arregi & Pietraszko (2021).

[21]A reviewer wonders whether the PF requirement of verb-framed languages can be argued to apply to phonologically null functional heads in general in these languages. In the remainder of this presentation I continue to focus on the functional head involved in the argument structure of verbal predicates, leaving the exploration of this hypothesis to further research.

6, illustrating the syntactic structures of the verb-framed resultative predicate in (33) and the verb-framed creation predicate in (34), respectively, I represent the PF requirement on the *v* head by means of an index *[i]* which is deleted when the requirement is satisfied.
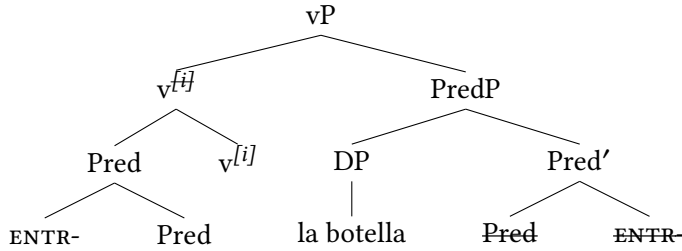

Figure 5: Syntactic structure of (33) (with a visual representation of the PF requirement (*[i]*) on *v*)
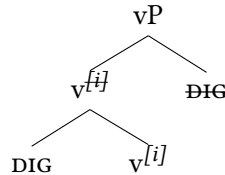

Figure 6: Syntactic structure of (34) (with a visual representation of the PF requirement (*[i]*) on *v*)

In the present account, the absence of the operation conflating a root with *v* in verb-framed languages arises as a by-product of *v*'s PF requirement. No parameterization of specific syntactic operations thus needs to be invoked. Verb-framed languages give the impression of lacking the operation conflating a root with *v*, because the syntactic configuration produced by such an operation is incompatible with the morphophonological context needed for the incorporation of *v*'s complement into *v* at PF in these languages. The syntactic representations in Figure 7 and Figure 8, corresponding to the Spanish ungrammatical satellite-framed resultative predicate in (37) and satellite-framed creation/consumption predicate in (38), respectively, illustrate this.[22] In the case of (37), ungrammaticality arises because neither the AP *limpia* 'clean' nor its root LIMP- can function as prefixes of verbs in Spanish, whereby the fulfillment of *v*'s PF requirement would give rise

---

[22](37) is grammatical in Spanish in the irrelevant readings involving a depictive or attributive interpretation of the AP (Jaume Mateu, p.c.).

to an unpronounceable sequence of morphemes.[23] Similarly, the ungrammaticality of (38) is due to the DP complement of *v* (*su agradecimiento* 'his thanks') not being able to incorporate onto *v*, which leaves the PF requirement on *v* unsatisfied.[24]

(37) * Él fregó   la   mesa limpia.
      he wipe.PST.AGR the table clean
      'He wiped the table clean.'   (Bigolin & Ausensi 2021: 519)
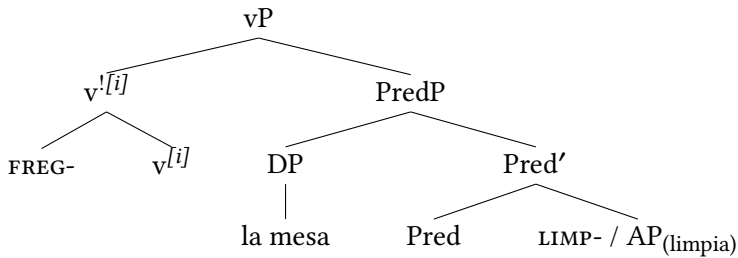

Figure 7: Syntactic structure of (37) (with a visual representation of the PF requirement (*[i]*) on *v*)

(38) * Juan sonríe   su   agradecimiento.
      Juan smile.PRS.AGR POSS gratitude
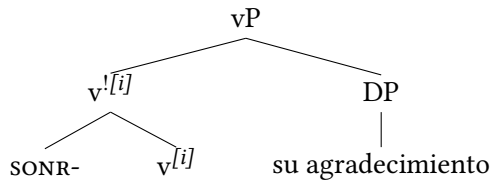   'Juan smiles his thanks.'   (Bigolin & Ausensi 2021: 527)

Figure 8: Syntactic structure of (38) (with a visual representation of the PF requirement (*[i]*) on *v*)

I propose that the requirement in (36), found in verb-framed languages, is also responsible for the pattern illustrated in (4) concerning Slavic languages (and

---

[23]In Distributed Morphology terms, one could formalize the context of insertion of the Vocabulary Item associated with LIMP- as requiring that no roots intervene between LIMP- and *v*.

[24]See Martínez Vázquez (2014) for the claim that, to a certain extent, complex creation/consumption predicates can be found in verb-framed Spanish. Further see Bigolin & Ausensi (2021) for an analysis of the examples in Martínez Vázquez (2014) as involving a verb-framed strategy.

weak satellite-framed languages in general), in which the result component of resultative predicates with manner-denoting verbs must form a prosodic word with the verb (Talmy 2000; Acedo-Matellán 2010; 2016). That is to say, I propose that weak satellite-framed languages are fundamentally verb-framed languages. The reasoning goes as follows. As the account of (37) and (38) shows, in addition to PF requirements on functional heads of the type in (36), a relevant factor in determining the availability of specific constructions in a given language is whether or not the constructions in question can be spelled out consistent with the PF restrictions on the individual items that make up the lexical inventory of the language concerned. For instance, I have argued that in Spanish (and, more generally, in verb-framed languages) there are no instances of constructions involving the conflation of a root with *v* because such a syntactic configuration prevents the fulfillment of *v*'s requirement that it incorporates its complement, as the lexical inventory of Spanish does not contain morphemes capable of expressing a Talmian Path in the form of a verbal prefix. I argue that weak satellite-framed languages differ from standard verb-framed languages in the domain of resultative predicates in that their lexicon has result-denoting morphemes which can be realized as verbal prefixes (e.g., *iz-* 'out' in (4)). The prefixal nature of such morphemes may satisfy *v*'s requirement that it incorporates its complement, by concomitantly leaving open the possibility of conflating an independent root with *v*.[25] This gives rise to a satellite-framed behavior in the domain of resultative predicates, as noted in Talmy (2000) and further discussed in Acedo-Matellán (2016). The structure in Figure (9), corresponding to the Russian predicate in (4) (repeated here as (39)), illustrates this.

(39)   Ona   iz-pis-a-l-a   svoju ručku.
       she.NOM out-write-TH-PST-AGR POSS   pen.ACC
       'Her pen has run out of ink.' (Lit. 'She has written her pen out (of ink).')
                                            (Russian; Spencer & Zaretskaya 1998: 17)

---

[25]The pattern is discussed by Mateu (2017), who however continues to consider Slavic languages (and weak satellite-framed languages in general) as fundamentally satellite-framed languages. The parallelism between prefixed resultative predicates with manner-denoting verbs of Slavic languages and English satellite-framed resultative constructions is proposed in Spencer & Zaretskaya (1998) and Mateu (2008). See Snyder (2012) for the claim that Russian patterns with verb-framed languages with respect to The Compounding Parameter of Snyder (1995; 2001). Russian and Czech have been argued to be verb-framed languages also by Gehrke (2008), who relates their verb-framedness to their realizing accomplishment structures in the verb (also by means of prefixes). I come back to Gehrke's (2008) proposal in §4.2, where I elaborate on its potential relevance for the data dealt with in the present chapter.

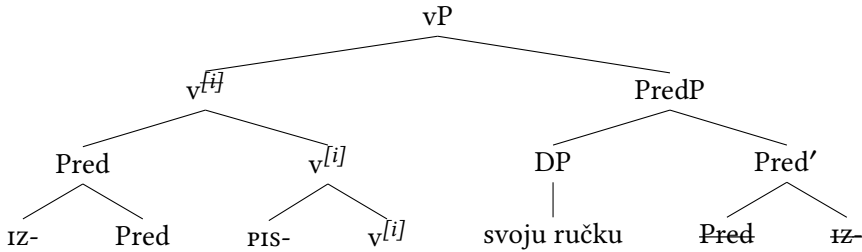2 *Slavic creation/consumption predicates in light of Talmy's typology*

Figure 9: Syntactic structure of (39) (with a visual representation of the PF requirement (*[i]*) on *v*)

From the hypothesis that weak satellite-framed languages are actually verb-framed languages (in the sense of (36)), it also follows that such languages should display a clear verb-framed behavior in the domain of creation/consumption predicates. No prefixal morpheme capable of referring to the object of creation/consumption predicates is present in the lexicon of these languages, whereby only creation/consumption predicates that involve the incorporation of a root into *v* can be licensed, in addition to predicates involving overt light verbs such as *do* or *make* (e.g., (31)). See this in Figure 10, where the syntactic representation of the ungrammatical Russian predicate in (30a) (repeated in (40)) is provided.

(40) * Ona   čes-a-l-a   dyrku   v pal'to.   (Russian)
        she.NOM brush.IPFV-TH-PST-AGR hole.ACC in coat.LOC
        'She was brushing a hole in her coat.'
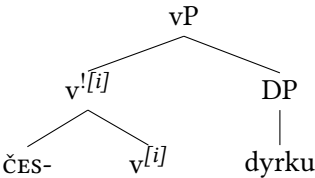
Figure 10: Syntactic structure of (40) (with a visual representation of the PF requirement (*[i]*) on *v*)

The present account provides a solution to the minimalist conundrum whereby verb-framed languages seem to lack a structure building operation (that of conflating a root with the *v* head) which is instead available in satellite-framed languages (see discussion in Folli & Harley 2020). In present terms, the resultative predicates with manner-denoting verbs and a prefixal result found in weak satellite-framed languages like Slavic languages are precisely to be regarded as constructions where a root is conflated with *v* in a verb-framed system.

Alessandro Bigolin

3.3 A comparison with some predecessors

Previous neo-constructionist approaches to Talmy's typology emphasize either that verb-framed languages always express the Path component in the main verb (Acedo-Matellán & Mateu 2013; Acedo-Matellán 2016; Folli & Harley 2020, among others) or that verb-framed languages lack predicates where, more generally, the verb expresses a manner co-event (Mateu 2012). The former approach runs into problems when considering that verb-framed languages and weak satellite-framed languages do not display complex creation/consumption predicates of the type displayed by satellite-framed languages, as nothing in this approach precludes the realization of such predicates – where no result component is involved – in these languages. Put differently, complex creation/consumption predicates are predicted to be universally available by this approach, contrary to facts.[26] The latter approach correctly predicts the unavailability of complex creation/consumption predicates in verb-framed languages, but it also predicts that weak satellite-framed languages should behave on a par with standard satellite-framed languages in allowing complex creation/consumption predicates. Furthermore, as discussed in §3.1, the generalization provided by this approach can only be taken as a descriptive one, as it cannot itself be considered explanatory without entailing a conception of syntax as a locus of parametric variation.

4 The role of perfectivizers

4.1 Internal prefixes and events of creation/consumption

In Slavic languages, the contrast between the imperfective and the perfective aspectual viewpoints is typically achieved by means of verbal prefixation and

---

[26]Aware of this prediction, Folli & Harley (2020) argue that complex creation/consumption predicates indeed do not give rise to cross-linguistic variation related to Talmy's typology and are generally available in verb-framed languages. This fact would then constitute the empirical proof that the expression of a co-event in the main verb is a universally available linguistic process. Specifically, Folli & Harley (2020) note that creation/consumption predicates such as (10), (11), and (12) are licensed both in satellite-framed English and in verb-framed Italian, and they assume that these predicates involve the expression of a manner co-event in the verb, similar to what is observed in satellite-framed resultative predicates. However, as shown in Table 2, that these specific examples do not give rise to significant cross-linguistic variation cannot be taken to conclude that no typological variation exists in the domain of creation/consumption predicates. Namely, the examples in Folli & Harley (2020) can be taken to involve verbs whose conceptual meaning implies the creation of the direct object, which in turn is interpreted as a hyponym of the verb (in the sense of Hale & Keyser 1997; 2002). As such, they can be argued to involve the verb-framed incorporation pattern exemplified in Figure 3, whereby they are allowed in verb-framed Italian.

suffixation. In a standard case, basic verbal stems have an imperfective reading, which is turned perfective via the addition of a prefix. The Russian examples in (41) illustrate this.

(41) a. My   pis-a-l-i   pis'mo.
          we.NOM write.IPFV-TH-PST-AGR letter.ACC
          'We were writing a letter.'   (Russian; Smith 1991: 302)

      b. On   na-pis-a-l   pis'mo.
          he.NOM PFV-write-TH-PST letter.ACC
          'He wrote a letter.'   (Russian; Smith 1991: 301)

Normally, the perfective prefix comes from the same inventory of morphemes which can provide the Talmian Path component in resultative predicates. Indeed, it has been argued that prefixes of this type – hereafter referred to as "internal" prefixes – denote the incorporation of a non-referential result into the verb, in a resultative structure (Ramchand & Svenonius 2002; Gehrke 2008; Acedo-Matellán 2016; Kwapiszewski 2022, among others).[27]

In the present framework, Slavic predicates perfectivized via internal prefixes (such as the Russian one in (41b)) are thus attributed the syntactic structure in Figure 11. I assume that predicates depicting events of creation/consumption made perfective via internal prefixes consistently involve the argument structure that is found in resultative (change of state/location) predicates, the direct object being interpreted as a created or consumed entity due to pragmatic factors arising from the conceptual interpretation of the construction. Insofar as these predicates involve the incorporation of *v*'s complement into *v*, as shown in Figure 11, they are predicted to be possible in Slavic languages in the same way as resultative predicates with manner-denoting verbs are, the incorporation of the prefix fulfilling the verb-framed requirement of the language as understood in (36). In what follows, I present the results of a study exploring the validity of such a prediction.

---

[27]Internal (or "lexical") prefixes are contrasted with external (or "superlexical") ones. The distinction is motivated by a series of factors which point toward the idea that internal prefixes are merged inside the vP (hence the name), while external prefixes are merged higher in the functional spine of the clause. For discussion of the distinction between internal and external verbal prefixes in Slavic languages, see Babko-Malaya (1999); Romanova (2004); Svenonius (2004); Borik (2006); Arsenijević (2006, 2007); Gehrke (2008); Žaucer (2009); Łazorczyk (2010); Tatevosov (2011); Milosavljević (2022); Kwapiszewski (2022), among many others. The classification of Slavic prefixes has also been argued to be more nuanced than the traditional bi-partite division found in the literature. For instance, Tatevosov (2008) argued that in Russian there exists a class of prefixes (e.g., *do-* and *pere-*) that exhibit an intermediate behavior between
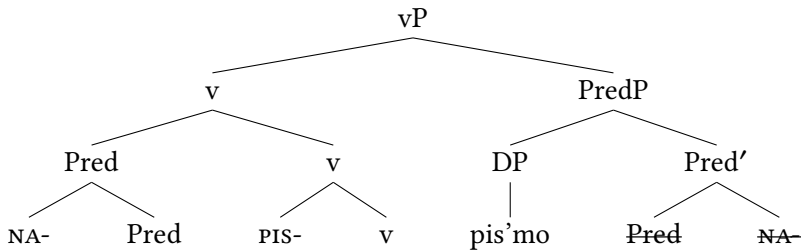
Alessandro Bigolin

Figure 11: Syntactic structure of (41b)

In order to verify the prediction, I have conducted the same test run for bona fide creation/consumption predicates which was described in §2. This time, however, the English examples have been left in their non-progressive form, to check whether the presence of the perfective prefixes in their Slavic counterparts affects the grammaticality of their literal translation in the Slavic languages. The results obtained, summarized in Table 4, show that Slavic languages clearly behave on a par with satellite-framed languages (cf. Table 1) when a perfective prefix is present, confirming the prediction.[28] The grammatical renditions of the English example in (22) in Russian, Ukrainian and Polish illustrate this ((42)). The structure for the Russian example in (42a), which is understood to hold also for the rest of the data, is provided in Figure 12.

(42)   a. Ona   pro-čes-a-l-a   dyrku   v pal'to.   (Russian)
           she.NOM PFV-brush-TH-PST-AGR hole.ACC in coat.LOC

       b. Vona   pro-ter-l-a   dyrku   na kurtci.   (Ukrainian)
           she.NOM PFV-brush-PST-AGR hole.ACC in coat.LOC

       c. Ona   wy-czes-a-ł-a   dziurę   w płaszczu.   (Polish)
           she.NOM PFV-brush-TH-PST-AGR hole.ACC in coat.LOC
       'She brushed a hole in her coat.'

The contrast in acceptability between predicates with unprefixed verbs and predicates with prefixed verbs in the expression of complex events of creation/

---

internal and external prefixes. Since the examples from Russian collected in this study do not involve such prefixes, I do not pursue this issue further here.

[28]The native speakers of Serbian and Croatian seem more conservative than the native speakers of the other Slavic languages tested in disallowing a creation/consumption reading for several of the predicates involved. At the moment, I am agnostic as to why the pattern displayed by Serbian and Croatian in this test differs in this way from that of the other Slavic languages.

Table 4: Perfective predicates with creation/consumption reading in Slavic languages (prefixed predicates)

Example	Rus	Ukr	Pol	Slo	Ser	Hrv
(6) John sang a song	✓	✓	✓	✓	✓	✓
(7) They danced a Sligo jig	✓	✓	✓	✓	✓	✓
(8) Ariel ate the mango	✓	✓	✓	✓	✓	✓
(9) He dug a hole in the ground	✓	✓	✓	✓	✓	✓
(10) She wove the tablecloth	✓	✓	✓	✓	✓	✓
(11) Marco painted a sky	✓	✓	✓	✓	✓	✓
(12) Maria carved a doll	✓	✓	✓	✓	✓	✓
(13) She burned a hole in her coat	✓	✓	✓	✓	??	✓
(14) He scratched a hole in the ground	✓	✓	✓	✓	??	★
(15) She punctured a wound in her finger	✓	✓	✓	✓	??	★
(16) She cut a wound in her foot	✓	?	✓	✓	★	★
(17) She bit a hole in the bag	✓	✓	✓	✓	✓	★
(18) The adventurer walked the trail	✓	✓	✓	✓	✓	✓
(19) The adventurer swam the channel	✓	✓	✓	✓	✓	✓
(20) Deanne kicked a hole in the wall	✓	✓	✓	✓	??	★
(21) She magicked a cursor	✓	✓	✓	✓	??	✓
(22) She brushed a hole in her coat	✓	✓	✓	✓	?	★
(23) John smiled his thanks						
(24) Elna frowned her discomfort						

Figure 12: Syntactic structure of (42a) (with a visual representation of the PF requirement ([i]) on v)

consumption in Slavic languages (compare Table 3 with Table 4) cannot be argued to depend on the aspectual shift from the imperfective reading of the former type of predicates to the perfective reading of the latter type of predicates. This is proved by the availability, for the examples that are ungrammatical in the imperfective reading provided by unprefixed verbs, of imperfective predicates obtained via secondary imperfectivization. Secondary imperfectivization is a strategy found in Slavic languages whereby a prefixed, perfective verb is turned

into an imperfective reading by means, typically (although not necessarily), of a further process of affixation (Babko-Malaya 1999; Romanova 2004; Svenonius 2004; Kwapiszewski 2022, among others). In the examples under consideration, secondary imperfectivization gives rise to grammatical predicates also in those cases where an imperfective reading involving unprefixed verbs gives rise to ungrammaticality. This is illustrated in (43) with the Ukrainian translations of (20), which is unavailable in the imperfective unprefixed version (43a) but is grammatical both in the perfective prefixed version (43b) and in the imperfective prefixed version obtained via secondary imperfectivization (43c).

(43) a. * Din byv   dyru   u stini.   (Ukrainian)
         Din kick.IPFV.PST hole.ACC in wall.LOC
         Intended: 'Din was kicking a hole in the wall.'

     b. Din pro-byv   dyru   u stini.
        Din PFV-kick.PST hole.ACC in wall.LOC
        'Din kicked a hole in the wall.'

     c. Din pro-byv-av   dyru   u stini.
        Din PRO-kick.PST-IPFV hole.ACC in wall.LOC
        'Din was kicking a hole in the wall.'

These facts suggest that the predicate's grammaticality does not rely on the perfective reading, but on the presence of the prefix, which fulfills the verb-framed requirement of the language by incorporating into v from its complement.

4.2 Incrementality in complex predicates

The meaning contribution of internal prefixes in the licensing of complex predicates in Slavic languages warrants further investigation. For instance, Gehrke (2008) posits that complex predicates, in which the main verb denotes an activity, require an accomplishment event structure, which in satellite-framed resultative constructions is licensed by an incremental structure provided by a secondary predicate. She further argues that internal prefixes of Slavic languages derive accomplishment structures, and that in these languages (specifically, she refers to Czech and Russian) accomplishment structures are realized in the verb complex, either by the verb itself or by an internal prefix. Gehrke's findings may offer an alternative explanation for why complex predicates of creation/consumption are grammatical in Slavic languages only when prefixed. Unprefixed complex predicates of creation/consumption might be infelicitous in Slavic languages due to the absence of an accomplishment structure within the verbal complex. This explanation rests on the assumption that complex predicates, cross-linguistically,

require the presence of an accomplishment event structure. However, the idea that an accomplishment event structure is needed to license satellite-framed constructions is not undisputed. For instance, Folli & Harley (2006) discuss cases of satellite-framed predicates of English, in which a PP denoting an unbounded path appears as the secondary predicate, as in (44). Since the incremental structure associated with PPs of this kind does not have a culmination point, the overall predicate lacks an accomplishment event structure.[29]

(44) a. John waltzed Matilda around and around the room for hours.

  b. John walked Mary along the river all afternoon.

(Folli & Harley 2006: 125)

  c. John walked Mary towards her car for 3 hours.

(Folli & Harley 2006: 137)

Another explanation worth considering is that unprefixed complex creation/consumption predicates are not licensed in Slavic languages due to the absence of incrementality in these constructions. I argue that this explanation is not satisfactory either. In these languages, given the right context, predicates of creation/consumption can be telic even if unprefixed (see, e.g., Gehrke 2008: 179, fn. 41; Mehlig 2012), the incremental path structure being provided by the direct object (Rappaport Hovav 2008, 2014). For instance, Mehlig (2012) argues that such a reading of the direct object in predicates denoting events of creation/consumption is possible in Russian if the extent of the entities denoted by the object has been determined in advance (e.g., from the conversational context) and these entities are referred to in the relevant imperfective predicate by means of a demonstrative (e.g., *ėtot/tot* 'this/that'). This is illustrated by Mehlig (2012) with examples like (45), where the consumption predicate *on est ėti dva banana* 'he is eating those two bananas' is successfully modified by the expression *Odin on uže s"el* 'He has already eaten one of them', which presupposes that the object has an incremental structure associated with it, because the two conditions listed above are satisfied (see the text preceding the consumption predicate in (45), where the amount of *bananas* involved in the eating event is pre-determined, and the presence of the demonstrative *ėti* 'these' in the consumption predicate).

---

[29]The temporal adverbials in the examples in (44) show that the PPs in these examples are not understood as referring to a bounded Path. I am not aware of studies concerned with the availability of such examples in Slavic languages. The unavailability of these examples in Polish (Wojciech Lewandowski, p.c.) and in Italian, however, points toward the idea that the satellite-framed/verb-framed division should not be (only) intended as a constraint in the expression of accomplishment structures in some languages and not in others.

Alessandro Bigolin

(45) Segodnja utrom ja dal[PFV] Saše dva banana. V dannyj moment on est[IPFV] ėti dva banana. Odin on uže s"el[PFV].
'This morning I gave Sasha two bananas. At the moment he is eating those two bananas. He has already eaten one of them.'

(Russian; Mehlig 2012: 216)

According to the hypothesis under discussion, the complex creation/consumption predicates that gave rise to ungrammaticality in Russian (see Table 3) should become acceptable if the contextual conditions identified in Mehlig (2012) are met, as the predicates would then be given an incremental structure by the direct object. However, the prediction is not borne out. The same results as those listed in Table 3 are obtained in Russian if the contextual conditions discussed in Mehlig (2012) are met, as illustrated in (46) with an example based on the predicate in (12) (Dària Serés, p.c.).

(46) Segodnja utrom Deanne zakazali s-delat'
today morning Deanne.DAT commission.PFV.PST.PL PFV-make.INF
dve reznye kukly. V dannyj moment *ona režet ėti
two carved dolls.ACC in this moment she.NOM carve.IPFV.PRS these
dve kukly. Skoree vsego ona uže vy-rez-a-l-a odnu
two dolls.ACC probably she.NOM already PFV-carve-TH-PST-AGR one
iz nix.
of them
'This morning Deanne was commissioned to make two carved dolls. At the moment she is carving those two dolls. She has probably already carved one of them.'

Similar considerations apply in Slovak, which also seems to license a reading of the object as having an incremental structure associated with it under the conditions in Mehlig (2012) but does not allow complex creation/consumption predicates in such contexts (Natália Kolenčíková, p.c.). In the case of Serbian, modifying expressions equivalent to the Russian *Odin on uže s"el* 'He has already eaten one of them' in (45) are compatible with predicates denoting events of consumption regardless of the contextual conditions in Mehlig (2012) (Predrag Kovačević, p.c.), yet complex creation/consumption predicates with unprefixed verbs are not licensed (see Table 3). The ungrammaticality of the predicate in (46) is accounted for by the account proposed in the present chapter; the predicate is not grammatical in Russian because it does not have a verbal prefix which fulfills the language's verb-framed requirement, by incorporating onto the *v* head.

In sum, the approach to Talmy's typology proposed in this chapter allows us to account for the kind of cross-linguistic variation related to the typology in the domain of predicates of creation/consumption regardless of whether an incremental structure is provided to the predicate by the verb, by a prefix of the verb or by a phrasal complement of the verb. In the next section, I propose that the present account of Slavic languages as verb-framed languages should be extended to Latin, which was argued to be a weak satellite-framed along with Slavic languages by Acedo-Matellán (2010; 2016).

5 Were there complex creation/consumption predicates in Latin?

As observed in Talmy (2000), and extensively explored in Acedo-Matellán (2010; 2016), Latin behaves on a par with Slavic languages in regard to Talmy's typology, in that resultative predicates where the verb denotes a co-event are allowed as long as the Path is expressed by a verbal prefix. As such, Latin is predicted to allow complex creation/consumption predicates by previous neo-constructionist accounts (see §3.3). Some examples of alleged complex creation predicates from Latin are provided in Acedo-Matellán (2016) to prove this point. This goes against the prediction of the present account of Talmy's typology, according to which neither Latin nor Slavic languages, *qua* weak satellite-framed languages, should be able to license bona fide creation/consumption predicates in the absence of an incorporation process of *v*'s complement onto *v*. Indeed, the data discussed in Acedo-Matellán (2016) are surprising in light of the pattern displayed by Slavic languages in this respect. In this section, I argue that there is no clear reason for attributing a complex creation/consumption reading to the Latin examples provided in Acedo-Matellán (2016). Afterward, I present the results of a corpus search which point toward the conclusion that complex creation/consumption predicates are absent in Latin, in line with the prediction of the present approach.

The examples discussed in Acedo-Matellán (2016) are provided in (47) to (51).

(47) Qui alteri misceat mulsum.
who.NOM another.DAT mix.SBJV.AGR honeyed_wine.ACC
'He who makes honeyed wine for someone else.'   (Latin; *Cic.* Fin. *2, 5, 17*)

(48) Vulnus [...] quod acu punctum
wound.NOM which.NOM needle.ABL puncture.PTCP.PFV.NOM
videretur.
seem.IPFV.SBJV.AGR
'A wound that seemed to have been punctured with a needle'.

Alessandro Bigolin

(Latin; *Cic.* Mil. *65*)

(49) [Serpens] volubilibus squamosos nexibus orbes
snake.NOM looping.ABL.PL scaly.ACC.PL writhing.ABL.PL coil.ACC.PL
torquet.
twist.PRS.AGR
'The snake twists his scaly coils in looping writhings.'

(Latin; *Ov.* Met. *3, 41*)

(50) Viam silice sternendam [...] locauerunt.
way.ACC flint-stone.ABL strew.PTCP.GRDV.ACC establish.PRF.AGR
'They established that the way was to be paved with flint stone.'

(Latin; *Liv. 38, 28, 3*)

(51) Aeriam truncis [...] cumulare pyram.
high.ACC log.ABL.PL gather.INF pyre.ACC
'To build a high pyre out of logs.' (Latin; *Stat.*, Teb. *6, 84*)

I suggest that most of these examples are compatible with a reading as either involving hyponymous objects or displaying resultative predicates of change of state, therefore adopting a verb-framed strategy. For instance, *pyram* 'pyre' in (51) could be interpreted as a hyponym of *cumulare* '(lit.) cumulate'. Indeed, a creation reading of this verb is also found in verb-framed Italian, as (52) shows.[30]

(52) [...] il primo dovrà aver cumulato esperienza nella
the first must.FUT.AGR have.INF gather.PTCP.PFV experience in.the

---

[30]An anonymous reviewer asks me to elaborate on the relevance of the Italian example in (52) for the conclusion that the Latin example in (51) is not a satellite-framed construction. Both the Latin and the Italian example refer to a creation event in which a 'cumulation' is formed. As Hale & Keyser (1997) noted, the conceptual content of the verb in predicates of this kind (in the cases at hand, *cumulare*, meaning 'cumulate', or 'gather') points non-referentially to the nature of the entity effected during the event (e.g., in (51) and (52), a 'cumulation' of some sort). The object, in turn, directly refers to such an effected entity. For instance, the predicate in (52) can be paraphrased as 'make a gathering that *consists of* experience'. Similarly, the predicate in (51) can be paraphrased as 'make a gathering that *consists of* a pyre'. For this reason, direct objects of this type have been referred to in the literature as 'hyponymous arguments' of the verb. As discussed in §3, predicates of this kind have been argued to involve the incorporation into *v* of a root e-merged as the complement of *v* (see Figure 3). Thus, they are expected to be well-formed in verb-framed languages. Assuming that Italian is a verb-framed language, that the construction in (51) can also be found in Italian, as (52) shows, provides additional evidence to the claim that such a construction is a verb-framed construction, whereby it does not constitute a counterexample to the proposal that Latin should be regarded as a verb-framed language.

> grande distribuzione, il   secondo sul   prodotto e   sul   contatto
> big   distribution   the second   on.the product   and on.the contact
> con i   grandi clienti.
> with the big   clients
> 'The first one must have gathered experience in large-scale distribution,
> the second one on the product and in dealing with large clients.'
>
> (CORIS[31])

As for (48), the availability of *puncture a wound* (cf. (15)) in weak satellite-framed Slavic languages and in verb-framed languages seems to be very limited, but the Ukrainian speaker fully accepts it (Table 3) and the Basque speaker considers it marginally acceptable (Table 2), suggesting that this predicate is not entirely precluded in these language types. Finally, I suggest that examples such as (49) and (50) can be compatible with a change-of-state reading of the direct object, which would imply the adoption of a verb-framed resultative structure. For instance, a snake can twist its coils also if the coils have been previously formed, e.g., by the position of the body prior to the *twisting*. Similarly, an existing road can be ordered to be covered with flint stone, supposing, for instance, that it was unpaved before. The proposed syntactic structure of the Latin example in (50), assuming a change-of-state reading of the predicate, is provided in Figure (13).
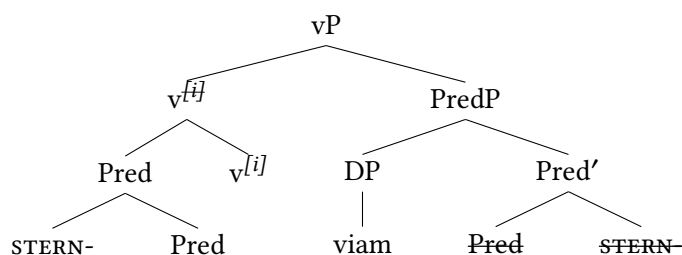
Figure 13: Syntactic structure of (50) (with a visual representation of the PF requirement ($[i]$) on v)

In order to further substantiate the prediction that complex creation/consumption predicates could not be licensed in weak satellite-framed Latin, I carried out a corpus-based investigation checking the co-occurrence, in a creation reading, of verbs that can be associated with a manner interpretation with two direct objects that seem to be particularly productive in English complex creation predicates, namely *hole* (Lat. *foramen*) and *wound* (Lat. *vulnus*). The corpus used for

---

[31]*Corpus di Riferimento dell'Italiano Scritto*, Università di Bologna.

Alessandro Bigolin

Latin, comprising texts from the Early and Classical periods (up to A.D. 200), is the *Classical Latin Texts* by The Packard Humanities Institute.[32] The verbs selected, listed in (53), were taken from Acedo-Matellán (2016).

(53) *amburo* 'burn', *caedo* 'cut, knock', *cremo* 'burn', *frico* 'rub', *rado* 'scrape', *tundo* 'beat', *uro* 'burn', *verro* 'sweep'

Importantly, the English verbs corresponding to the Latin ones in (53) can give rise to creation predicates with *hole* or *wound* as effected object, as (54) shows.

(54) a. A discharge of those energies burned a hole in his forehead and killed him.   (Ausensi & Bigolin 2023: 155)
      b. [...] his words burned a wound inside her.   (Google Books)
      c. Dad cut a hole in his chest and made me pull his heart out.   (COCA)
      d. The Devil-Is-I pulled the knife he had used to cut a wound on his thumb and lunged forward at the leader of the twelve. (Google Books)
      e. Weena knocked a hole in the wall.   (COCA)
      f. But I scraped a hole in it so I could see.   (COCA)
      g. [...] he scraped a wound on his nose that never cleared up.
                                                                       (Google Books)
      h. I erased again and again until I had rubbed a hole in the paper.
                                                                       (COCA)
      i. [...] the mooring line has rubbed a wound in the willow bark.
                                                                       (Google Books)
      j. My 'beloved' boyfriend beat a hole in my roof and now it's awefully cold in there.   (COCA)
      k. A sudden shift in the wind swept a hole in the blowing snow.
                                                                       (Google Books)

The verbs in (53) were searched for by stem, while the objects were searched for in the nominative/accusative singular and plural forms.[33] None of the verb-complement combinations investigated provided relevant results. A verb-framed construction with *vulnus* 'wound' was found instead, as shown in (55).

---

[32]https://latin.packhum.org

[33]Being neuter, both *foramen* 'hole' and *vulnus* 'wound' appear as morphologically identical in their respective nominative and accusative forms.

(55) Sed uulnera   facta   igne   dum sanescunt,
     But wound.NOM.PL make.PTCP.PFV.NOM.PL fire.ABL while heal.PRS.AGR
     defricare   bubula   urina   convenit.
     off_rub.INF bovine.ABL urine.ABL fit.AGR
     'But while the wounds made with fire are healing, it is appropriate to
     cleanse them with bovine urine.'   (*Col.*, De Re Rustica *6.7.4*)

I take this lack of evidence to tentatively suggest that Latin lacked complex cre-
ation/consumption predicates of the type found in satellite-framed languages,
and needed to resort to verb-framed strategies in the domain of creation/con-
sumption predicates in the same way as Slavic languages do. This is in line with
the prediction, following from the present account, that complex creation/con-
sumption predicates are unavailable in weak satellite-framed languages.

   Picking up the discussion in §4.2, about the possibility that unprefixed complex
predicates of creation/consumption may be disallowed in Slavic languages due to
their lack of an incremental structure, it is relevant to notice that direct objects
could be associated with an incremental structure giving rise to telicity in the
predicate in Latin. The compatibility of the consumption predicate in (56) with
the time span adverbial *intra duas horas* 'within two hours' illustrates this.

(56) [...] nitrosae   aut amarae   aquae   polenta
         nitrous.NOM.PL or bitter.NOM.PL water.NOM.PL cornmeal.ABL
     addita   mitigantur,   ut intra duas
     add.PTCP.PFV.ABL mitigate.IPFV.SBJV.PASS.AGR that within two.ACC
     horas   bibi   possint.
     hour.ACC.PL drink.INF.PASS can.IPFV.SBJV.AGR
     'Nitrous and bitter waters are softened with added cornmeal, so that they
     can be drunk within two hours.'   (*Plin.*, Nat. *24, 3, 4*)

Assuming, based on the discussion in this section, that complex creation/consum-
ption predicates were not possible in Latin, such an absence cannot be attributed
to the predicate's lack of incrementality. The morphophonological account of
Talmy's typology proposed in this chapter provides an alternative explanation
of the phenomenon that is compatible with the observation that creation/con-
sumption predicates could be telic in Latin without the presence of the prefix
(see (56)).

Alessandro Bigolin

6 Conclusions

I presented the preliminary results of a pilot study concerning the possibility of licensing complex creation/consumption predicates in Slavic languages. The results obtained were further compared with data gathered from native speakers of several satellite-framed languages and verb-framed languages. The study shows that Slavic languages, which are considered as fundamentally satellite-framed in the literature on Talmy's typology (Talmy 2000; Acedo-Matellán 2016), appear to behave on a par with verb-framed languages in disallowing creation/ consumption predicates that involve a satellite-framed strategy.

Adopting a neo-constructionist perspective on argument structure, I have put forward a morphophonological approach to the variation related to Talmy's typology, understanding verb-framedness in terms of a morphophonological realization condition imposed at PF on the null v head involved in verbal predication. A null v is required to incorporate its complement in verb-framed languages. I have further argued that Slavic languages, and weak satellite-framed languages in general, should be regarded as fundamentally verb-framed languages, capturing the mandatory prefixation of the Path component in resultative predicates and the absence of complex predicates of creation/consumption in these languages as by-products of the verb framed PF requirement on the v head.

With the present morphophonological account of Talmy's typology, I have additionally provided a solution to the minimalist conundrum whereby verb-framed languages seem to consistently lack the structure-building operation associated with the expression of a co-event in the verb, namely the operation of conflating a root with v. To the extent that the verb-framed PF requirement can be satisfied by means of prefixation, the compounding operation can indeed successfully take place in a verb-framed system, as shown by the availability of prefixed satellite-framed resultative predicates in weak satellite-framed languages.

Afterward, I have explored the prediction that a creation/consumption reading of predicates with manner denoting verbs is available in Slavic languages when the predicate is perfectivized via internal prefixes, which have been argued to involve a resultative structure that receives a reading as involving an event of creation/consumption on the basis of conceptual/pragmatic considerations. The data gathered from the native speakers of the Slavic languages tested confirmed the prediction.

Finally I have argued that Latin, as a weak satellite-framed language (Acedo-Matellán 2010; 2016), lacked complex creation/consumption predicates of the type found in bona fide satellite-framed languages in the same way as Slavic languages do. I have argued that this is the case based on an analysis of some alleged

Latin complex creation/consumption predicates provided in Acedo-Matellán (2016), which have been shown to admit a reading either as involving a hyponymous object or as involving a resultative predicate of change of state. Afterward, I have presented the results of a corpus search supporting the prediction that complex creation/consumption predicates are not licensed in Latin. The results strengthen the general hypothesis that Latin and Slavic languages behave alike with respect to Talmy's typology (Acedo-Matellán 2016), meanwhile underpinning one of the main conclusion of the present account whereby weak satellite-framed languages should be considered as fundamentally verb-framed languages.

Abbreviations

ABL	ablative	NEG	negation
ACC	accusative	NOM	nominative
AGR	agreement	PASS	passive
CL	classifier	PFV	perfective
DAT	dative	PL	plural
FUT	future	POSS	possessive
GER	gerund	PRF	perfect
GRDV	gerundive	PRS	present
INF	infinitive	PST	past
INS	instrumental	PTCP	participle
IPFV	imperfective	SBJV	subjunctive
LOC	locative	SG	singular
M	masculine	TH	theme vowel

Acknowledgments

Alessandro Bigolin

References

Acedo-Matellán, Víctor. 2010. *Argument structure and the syntax-morphology interface: A case study in Latin and other languages*. Barcelona: Universitat de Barcelona. (Doctoral dissertation).

Acedo-Matellán, Víctor. 2016. *The morphosyntax of transitions: A case study in Latin and other languages*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780198733287.001.0001.

Acedo-Matellán, Víctor & Arkadiusz Kwapiszewski. 2021. Theory and facts in Talmy's typology: A reaction to Folli & Harley 2020. Seminar at Centre de Lingüística Teòrica. Universitat Autònoma de Barcelona, January 15.

Acedo-Matellán, Víctor & Jaume Mateu. 2013. Satellite-framed Latin vs. verb-framed Romance: A syntactic approach. *Probus* 25(2). 227–265. DOI: 10.1515/probus-2013-0008.

Alexiadou, Artemis. 2015. Syntax and the lexicon. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax – theory and analysis: An international handbook*, 1088–1127. Berlin: De Gruyter Mouton.

Arregi, Karlos & Asia Pietraszko. 2021. The ups and downs of head displacement. *Linguistic Inquiry* 52(2). 241–289. DOI: 10.1162/ling_a_00377.

Arsenijević, Boban. 2006. *Inner aspect and telicity: The decompositional and the quantificational nature of eventualities at the syntax-semantics interface*. Utrecht: LOT Publications.

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

Ausensi, Josep & Alessandro Bigolin. 2023. On the argument structure realization of result verbs: A syntactic approach. *Acta Linguistica Academica* 70(1). 139–160. DOI: 10.1556/2062.2023.00567.

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

Beavers, John, Beth Levin & Shiao Wei Tham. 2010. The typology of motion expressions revisited. *Journal of Linguistics* 46. 331–377. DOI: 10.1017/S0022226709990272.

Bigolin, Alessandro & Josep Ausensi. 2021. A new resultative construction in Spanish? A reply to Rodríguez Arrizabalaga. *Folia Linguistica* 55(2). 517–546. DOI: 10.1515/flin-2021-2078.

Borer, Hagit. 2005. *Structuring sense,* vol. 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.

Condamines, Anne. 2013. Quand la passion autorise la transitivation d'un circonstanciel de lieu. *Journal of French Language Studies* 23(3). 335–355. DOI: 10.1017/S095926951300001X.

Davies, Mark. 2008-. *The Corpus of Contemporary American English (COCA).* Available online at https://www.english-corpora.org/coca/.

Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35(3). 355–392. DOI: 10.1162/0024389041402634.

Fan, Sheng-yang. 2014. *Argument structure in Mandarin Chinese: A lexical-syntactic perspective.* Madrid: Universidad Autónoma de Madrid. (Doctoral dissertation).

Folli, Raffaella & Heidi Harley. 2005. Flavors of *v.* In Paula Marie Kempchinsky & Roumyana Slabakova (eds.), *Aspectual inquiries*, 99–120. Dordrecht: Springer. DOI: 10.1007/1-4020-3033-9.

Folli, Raffaella & Heidi Harley. 2006. On the licensing of causatives of directed motion: Waltzing Matilda all over. *Folia Linguistica* 60(2). 121–155. DOI: 10.1111/j.1467-9582.2006.00135.x.

Folli, Raffaella & Heidi Harley. 2008. Teleology and animacy in external arguments. *Lingua* 118(2). 190–202. DOI: 10.1016/j.lingua.2007.02.004.

Folli, Raffaella & Heidi Harley. 2020. A head-movement approach to Talmy's typology. *Linguistic Inquiry* 51(3). 425–470. DOI: 10.1162/ling_a_00351.

Folli, Raffaella & Gillian Ramchand. 2005. Prepositions and results in Italian and English: An analysis from event decomposition. In Hank J. Verkuyl, Henriette de Swart & Angeliek van Hout (eds.), *Perspectives on aspect*, 81–105. Dordrecht: Springer. DOI: 10.1007/1-4020-3232-3.

Gallego, Ángel J. 2012. A note on cognate objects: cognation as doubling. *Nordlyd* 39(1). 95–112. DOI: 10.7557/12.2289.

Gehrke, Berit. 2008. *Ps in motion: On the semantics and syntax of P elements and motion events.* Utrecht: LOT Publications.

Alessandro Bigolin

Goldberg, Adele. 1995. *Constructions: A Construction Grammar approach to argument structure*. Chicago: The University of Chicago Press.

Hale, Kenneth & Samuel Jay Keyser. 1993. On argument structure and the lexical expression of syntactic relations. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 53–109. Cambridge, MA: MIT Press.

Hale, Kenneth & Samuel Jay Keyser. 1997. The limits of argument structure. In Amaya Mendikoetxea & Myriam Uribe-Etxebarria (eds.), *Theoretical issues at the morphology-syntax interface*, 203–230. Bilbao: Universidad del País Vasco.

Hale, Kenneth & Samuel Jay Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press.

Harley, Heidi. 2005. How do verbs get their names? Denominal verbs, manner incorporation, and the ontology of verb roots in English. In Nomi Erteschik-Shir & Tova Rapaport (eds.), *The syntax of aspect: Deriving thematic and aspectual interpretation*, 42–64. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199280445.001.0001.

Haugen, Jason D. 2009. Hyponymous objects and late insertion. *Lingua* 119(2). 242–262. DOI: 10.1016/j.lingua.2007.10.008.

Hoekstra, Teun. 1988. Small clause results. *Lingua* 74(2-3). 101–139. DOI: 10.1016/0024-3841(88)90056-3.

Iwata, Seizi. 2020. *English resultatives: A force-recipient account*. Amsterdam: John Benjamins. DOI: 10.1075/cal.26.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In Laurie Ann Zaring & Johan Rooryck (eds.), *Phrase structure and the lexicon*, 109–137. Dordrecht: Kluwer.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kwapiszewski, Arkadiusz. 2022. *Aspect and event structure: The morphosyntax of Polish verbs from a cross-linguistic perspective*. Oxford: University of Oxford. (Doctoral dissertation).

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

Levin, Beth & Tova Rapoport. 1988. Lexical subordination. In Lynn MacLeod, Gary Larson & Diane Brentari (eds.), *Papers from the 24th annual regional meeting of the Chicago Linguistic Society*, 275–289. Chicago, IL: Chicago Linguistic Society.

Marantz, Alec. 1984. *On the nature of grammatical relations*. Cambridge, MA: MIT Press.

Martínez Vázquez, Montserrat. 2014. Reaction object constructions in English and Spanish. *ES: Revista de Filología Inglesa* 35. 193–217.

Mateu, Jaume. 2002. *Argument structure: Relational construal at the syntax-semantics interface.* Barcelona: Universitat Autònoma de Barcelona. (Doctoral dissertation).

Mateu, Jaume. 2003. Parametrizing verbal polysemy: The case of *bake* revisited. In Jonathon Cihlar, Amy Franklin, David Kaiser & Irene Kimbara (eds.), *Papers from the 39th annual meeting of the Chicago Linguistic Society*, 256–269. Chicago, IL: Chicago Linguistic Society.

Mateu, Jaume. 2008. On the l-syntax of directionality/resultativity: The case of Germanic preverbs. In Anna Asbury, Jakub Dotlačil, Berit Gehrke & Rick Nouwen (eds.), *Syntax and semantics of spatial P*, 221–250. Amsterdam: John Benjamins. DOI: 10.1075/la.120.

Mateu, Jaume. 2012. Conflation and incorporation processes in resultative constructions. In Violeta Demonte & Louise McNally (eds.), *Telicity, change, and state: A cross-categorial view of event structure*, 252–278. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199693498.001.0001.

Mateu, Jaume. 2017. Two types of locative alternation. In Albert Álvarez González & Ía Navarro (eds.), *Verb valency changes: Theoretical and typological perspectives*, 51–79. Amsterdam: John Benjamins. DOI: 10.1075/tsl.120.

Mateu, Jaume & Víctor Acedo-Matellán. 2012. The manner/result complementarity revisited: A syntactic approach. In M. Cristina Cuervo & Yves Roberge (eds.), *The end of argument structure?*, 209–228. Bingley: Emerald.

Mateu, Jaume & Gemma Rigau. 2002. A minimalist account of conflation processes: Parametric variation at the lexicon-syntax interface. In Artemis Alexiadou (ed.), *Theoretical approaches to universals*, 211–236. Amsterdam: John Benjamins. DOI: 10.1075/la.49.

Mateu, Jaume & Gemma Rigau. 2010. Verb-particle constructions in Romance: A lexical-syntactic account. *Probus* 22. 241–269. DOI: 10.1515/prbs.2010.009.

Mehlig, Hans Robert. 2012. Hybrid predicates in Russian. *Journal of Slavic Linguistics* 20(2). 171–227.

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*, 17–42. Leiden: Leiden University Centre for Linguistics.

Nikitina, Tatiana. 2008. Pragmatic factors and variation in the expression of spatial goals: The case of *into* vs. *in*. In Anna Asbury, Jakub Dotlačil, Berit Gehrke & Rick Nouwen (eds.), *Syntax and semantics of spatial P*, 175–195. Amsterdam: John Benjamins. DOI: 10.1075/la.120.

Alessandro Bigolin

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/9780262162548.001.0001.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511486319.

Ramchand, Gillian & Peter Svenonius. 2002. The lexical syntax and lexical semantics of the verb-particle construction. In Line Mikkelsen & Christopher Potts (eds.), *Proceedings of WCCFL 21*, 387–400. Somerville, MA: Cascadilla Press.

Rappaport Hovav, Malka. 2008. Lexicalized meaning and the internal structure of events. In Susan Rothstein (ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*, 13–42. Amsterdam: John Benjamins. DOI: 10.1075/la.110.

Rappaport Hovav, Malka. 2014. Building scalar changes. In Artemis Alexiadou, Hagit Borer & Florian Schäfer (eds.), *The syntax of roots and the roots of syntax*, 259–281. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199665266.001.0001.

Rappaport Hovav, Malka & Beth Levin. 1998. Building verb meanings. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments: Lexical and compositional factors*, 97–134. Stanford, CA: CSLI Publications.

Rappaport Hovav, Malka & Beth Levin. 2010. Reflections on manner/result complementarity. In Malka Rappaport Hovav, Edit Doron & Ivy Sichel (eds.), *Lexical semantics, syntax, and event structure*, 21–38. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199544325.001.0001.

REAL ACADEMIA ESPAÑOLA. N.d. Banco de datos (CORPES XXI) [en línea]. *Corpus del Español del Siglo XXI (CORPES)*. <http://www.rae.es>.

Real-Puigdollers, Cristina. 2013. *Lexicalization by phase: The role of prepositions in argument structure and its cross-linguistic variation*. Barcelona: Universitat Autònoma de Barcelona. (Doctoral dissertation).

Romanova, Eugenia. 2004. Lexical vs. superlexical prefixes. *Nordlyd* 32(2). 255–278. DOI: 10.7557/12.69.

Smith, Carlota S. 1991. *The parameter of aspect*. Dordrecht: Kluwer. DOI: 10.1007/978-94-015-7911-7.

Snyder, William. 1995. *Language acquisition and language variation: The role of morphology*. Cambridge, MA: MIT. (Doctoral dissertation).

Snyder, William. 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-formation. *Language* 77(2). 324–342. DOI: 10.1353/lan.2001.0108.

Snyder, William. 2012. Parameter theory and motion predicates. In Violeta Demonte & Louise McNally (eds.), *Telicity, change, and state: A cross-categorial view of event structure*, 279–299. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199693498.001.0001.

Spencer, Andrew & Marina Zaretskaya. 1998. Verb prefixation in Russian as lexical subordination. *Linguistics* 36. 1–39. DOI: 10.1515/ling.1998.36.1.1.

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

Talmy, Leonard. 2000. *Toward a cognitive semantics,* vol. 2: *Typology and process in concept structuring*. Cambridge, MA: MIT Press.

Tatevosov, Sergei. 2008. Intermediate prefixes in Russian. In *Proceedings of Formal Approaches to Slavic Linguistics 16*, 423–445. Ann Arbor, MI: Michigan Slavic Publications.

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

Tenny, Carol. 1994. *Aspectual roles and the syntax-semantics interface*. Dordrecht: Kluwer. DOI: 10.1007/978-94-011-1150-8.

Truswell, Robert. 2007. Extraction from adjuncts and the structure of events. *Lingua* 117(8). 1355–1377. DOI: 10.1016/j.lingua.2006.06.003.

Volpe, Mark. 2004. Affected object unergatives. *Snippets* 8(4). 12–13.

Wang, Chyan-an Arthur. 2010. *The microparametric syntax of resultatives in Chinese language*. New York, NY: New York University. (Doctoral dissertation).

Washio, Ryuichi. 1997. Resultatives, compositionality, and language variation. *Journal of East Asian Linguistics* 6. 1–49. DOI: 10.1023/A:1008257704110.

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 3

Delimitatives, diminutive-iteratives and the secondary imperfective in North Slavic

Petr Biskup

University of Leipzig

This paper is concerned with diminutive-iterative verbs, delimitative verbs with the prefix *po-* and the secondary imperfective suffix. It is argued that diminutive-iterative *po*-verbs are derivationally based on delimitative predicates. Further, the secondary imperfective suffix is not an undifferentiated element. It is argued that the two instances of the imperfectivizing suffix – the iterative one and the progressive one – merge in distinct structural positions and that the delimitative prefix *po-* occurs between them. In the derivation of diminutive-iteratives, delimitative *po-* selects a predicate with a scalar structure and the Davidsonian event variable and contributes an extensive measure function to the base predicate. The iterative *-yva*, with its pluractional semantics, then iterates the eventuality denoted by the *po*-predicate.

1 Introduction

This section introduces diminutive-iterative *po*-verbs, delimitative *po*-verbs and secondary imperfective verbs and briefly overviews their morphosyntactic and semantic properties relevant to the analysis pursued in following sections.

1.1 Diminutive-iterative verbs

As to their form, diminutive-iterative verbs contain the prefix *po-* and the imperfectivizing/iterative suffix -ʏᴠᴀ, as shown in (1) for Russian, in (2) for Polish,

Petr Biskup

and in (3) for Czech (see also Švedova 1980: 600, Katny 1994: 66–70, and Petr 1986: 398). From now on, I will use -YVA as a shorthand that also stands for other allomorphs, e.g. in Russian, it stands for the allomorphs *-yva/-iva*, *-va* and *-a*.

(1) po-lëž-iva-t'
DEL-lie-ITER-INF
'to lie from time to time' (Russian)

(2) po-płak-iwa-ć
DEL-cry-ITER-INF
'to cry from time to time' (Polish)

(3) po-sed-á-va-t
DEL-sit-TH-ITER-INF
'to sit from time to time' (Czech)

Concerning morphological (grammatical/viewpoint) aspect properties, diminutive-iterative *po*-verbs are always imperfective. With respect to their meaning, diminutive-iteratives are usually described as denoting a short action (with low intensity) that is repeated several times; see e.g. Zaliznjak & Šmelev (1997: 103) for Russian, Czochralski (1975: 23) for Polish and Karlík et al. (1995: 194, 209) for Czech. The action does not have to be repeated regularly and the iterative suffix brings about an unspecified number of instances of the particular eventuality. Typically, the number of repetitions depends on the context.

1.2 Secondary imperfective verbs

Secondary imperfective verbs contain an -YVA allomorph, which derives imperfective predicates from perfective stems, as shown in examples (4), (5) and (6). Crucially, -YVA allomorphs used in formation of diminutive-iterative verbs are identical to the -YVA allomorphs used in secondary imperfective verbs.

(4) a. za-pis-a-t'[PFV]
behind-write-TH-INF
'to write down'
b. za-pis-yva-t'[IPFV]
behind-write-SI-INF
'to write down'
'to be writing down' (Russian)

(5) a. pod-pis-a-ć^{PFV}

(5) a. pod-pis-a-ć PFV
         below-write-TH-INF
         'to sign'

  b. pod-pis-ywa-ć IPFV
         below-write-SI-INF
         'to sign'
         'to be signing'   (Polish)

(6) a. vy-ps-a-t PFV
         out-write-TH-INF
         'to excerpt'

  b. vy-pis-ova-t IPFV
         out-write-SI-INF
         'to excerpt'
         to be excerpting'   (Czech)

Secondary imperfective verbs can have (at least) four meanings. 1. progressive, expressing that a certain eventuality is in progress; 2. iterative, which expresses the successive occurrence of several instances of a certain eventuality; 3. habitual (generic), which describes an eventuality that is characteristic of an extended time period; and 4. general-factual, which typically refers to a realized or even completed eventuality (in a fashion similar to perfective verbs); see e.g. Comrie (1976: 24–40), Dahl (1985: 75–102), Dickey (2000: 49–125), Grønn (2004: 22–30), Timberlake (2004: 417–424) and Petruxina (2011: 64–76). This article is concerned with the progressive and iterative meanings.

1.3 Delimitative verbs

Delimitative verbs are formed with the help of the delimitative prefix *po-* and with the reduplicated form *popo-* in the case of motion verbs in Czech, as demonstrated by (7), (8), and (9). The prefix is adjoined to an unprefixed, imperfective stem and derives a perfective predicate, as shown in the examples under discussion.

(7) a. čit-a-t' IPFV
         read-TH-INF
         'to read'
         'to be reading'

Petr Biskup

    b. po-čit-a-t'$^{\text{PFV}}$
       DEL-read-TH-INF
       'to read for a while'   (Russian)

(8) a. siedzi-e-ć$^{\text{IPFV}}$
       sit-TH-INF
       'to sit'
       'to be sitting'
   b. po-siedzi-e-ć$^{\text{PFV}}$
       DEL-sit-TH-INF
       'to sit for a while'   (Polish)

(9) a. nés-t$^{\text{IPFV}}$
       carry-INF
       'to carry'
       'to be carrying'
   b. popo-nés-t$^{\text{PFV}}$
       DEL-carry-INF
       'to carry sth. a little'   (Czech)

As to the meaning, the prefix delimits the eventuality denoted by the base predicate. Typically, it is a temporal delimitation, as in (7b) and (8b), but a property scale and a path scale can be delimited, too, as shown in (9b) for the path scale.

    Delimitative *po*-verbs are standardly claimed to be perfectiva tantum, i.e. they do not form secondary imperfectives; see Isačenko (1962: 391), Kopečný (1962: 110), Zaliznjak & Šmelev (1997: 94), and Łaziński (2020: 77). Thus, e.g. the delimitative verbs in (10a), (11a), and (12a) cannot be imperfectivized and receive a progressive delimitative meaning, as shown by the second translations in examples (10b), (11b), and (12b). The *po*-verb-YVA forms in the (b) examples can only have the diminutive-iterative meaning, as demonstrated by the first translations.

(10) a. po-čit-a-t'$^{\text{PFV}}$
        DEL-read-TH-INF
        'to read for a while'
    b. po-čit-yva-t'$^{\text{IPFV}}$
        DEL-read-ITER-INF
        'to read from time to time'
        Unavailable reading: 'to be reading for a while'   (Russian)

(11) a. po-płak-a-ć[PFV]
          DEL-cry-TH-INF
          'to cry for a while'

     b. po-płak-iwa-ć[IPFV]
          DEL-cry-ITER-INF
          'to cry from time to time'
          Unavailable reading: 'to be crying for a while'   (Polish)

(12) a. popo-nés-t[PFV]
          DEL-carry-INF
          'to carry sth. a little'

     b. popo-náš-e-t[IPFV]
          DEL-carry-ITER-INF
          'to carry sth. from time to time'
          Unavailable reading: 'to be carrying a little/for a while'   (Czech)

I leave attenuative *po*-verbs like the Russian *poprideržat'* 'to hold gently' aside in this article because the attenuative *po*- and the delimitative *po*- behave as two distinct elements morphosyntactically.[1] While the attenuative *po*- selects perfective stems that are prefixed, delimitative *po*-verbs are derived from the base, imperfective stems (see e.g. Isačenko 1962: 391, 396, Zaliznjak & Šmelev 1997: 101, Petr 1986: 398. Further, in Czech, the delimitative *po*- often adds a dative reflexive argument that is licensed by the agent and is obligatory.[2] In contrast, attenuative *po*-verbs can have the dative *si* 'self' but it is never obligatory. In addition, the two types of *po*-verbs behave differently with respect to the formation of secondary imperfectives. As discussed in the preceding paragraph, delimitative *po*-verbs are considered to be perfectiva tantum. The authors mentioned there do not discuss whether or not attenuative *po*-verbs can be imperfectivized but (at least some) *po*-attenuatives form secondary imperfectives; e.g. Russian *poprideržat'* 'to hold

---

[1]This contrasts with the semantic analysis by Součková (2004a,b) and Kagan (2016), who treat the two types of *po*- prefixes as one and the same semantic element, measuring degrees on a scale. Note that an analysis with two distinct morphosyntactic *po*-s does not preclude the possibility that the two prefixes have identical or very similar semantic properties.

[2]Consider e.g. (i) and the discussion of (23b) in §2. The most straightforward analysis would introduce the dative *si* 'self' in the specifier of the delimitative *po*- projection, where the argument is c-commanded by the agent placed in the specifier of VoiceP.

   (i) Strejda *(si) po-lyžoval a   odjel domů.
        uncle   self DEL-skied and went home
        'My uncle skied for a while and went home.'   (Czech)

gently' derives the imperfective form *popriderživat'* 'to hold gently/to be hold-ing gently'. Tatevosov (2009: 96) and Klimek-Jankowska & Błaszczak (2022: 7–9, 2024: ex. (85), (92)) place the Russian and Polish, respectively, attenuative *pod-* below the imperfectivizing -ʏᴠᴀ. If the attenuative *po-* behaves in the same way, then attenuative *po-*verbs should be able to undergo secondary imperfectiviza-tion. Given that the attenuative *po-* selects a perfective stem, it must belong to positionally restricted prefixes in terms of Tatevosov (2009), which means that it merges below the imperfectivizing -ʏᴠᴀ.

The remainder of the article is structured as follows. §2 argues that diminutive-iterative *po-*verbs are derived from delimitative predicates. §3 then offers a mor-phosyntactic and semantic analysis with the relevant derivational steps. §4 con-cludes the article.

2 Diminutive-iterative verbs are derived from delimitative predicates

Recall from the preceding section that both delimitative and diminutive-iterative verbs contain the prefix *po-* and include a delimited degree scale. Given this and the imperfectivizing and iterative effects of -ʏᴠᴀ in diminutive-iterative verbs, a natural idea is that diminutive-iteratives, as in (13b), are derived by applying the imperfectivizing -*yva* to the delimitative predicate, as in (13a).

(13) a. po-lež-a-t'
          ᴅᴇʟ-lie-ᴛʜ-ɪɴꜰ
          'to lie for a while'
      b. po-lëž-iva-t'
          ᴅᴇʟ-lie-ɪᴛᴇʀ-ɪɴꜰ
          'to lie from time to time'   (Russian)

Besides the morphological, semantic and aspectual arguments, there are also phonological facts which support such an analysis. Suffixes used in the deriva-tion of diminutive-iteratives are identical with -ʏᴠᴀ allomorphs used in the "stan-dard" secondary imperfectivization. Also phonological processes involved in the formation of diminutive-iterative *po-*verbs are identical with phonological pro-cesses involved in the derivation of "standard" secondary imperfectives. Com-pare the diminutive-iterative example (14) with the secondary imperfectivization example in (15). In both examples -*yva* shifts the accent from the theme to the root vowel and in both roots, we also observe the vowel gradation (lengthening)

from /o/ to /a/.[3] Consider also (13), which displays a vowel gradation in the root, too, and which also manifests the accent shift from the theme to the root.

(14) a. po-kol-ó-t'^{PFV}
           DEL-prick-TH-INF
           'to prick a little'
      b. po-kál-yva-t'^{IPFV}
           DEL-prick-ITER-INF
           'to prick from time to time'   (Russian)

(15) a. s-pros-í-t'^{PFV}
           with-ask-TH-INF
           'to ask'
      b. s-práš-iva-t'^{IPFV}
           with-ask-SI-INF
           'to ask'
           'to be asking'   (Russian)

The Polish examples in (16) and (17) show that in formation of both diminutive-iteratives and ordinary secondary imperfectives, /j/ is inserted to block hiatus.

(16) a. po-pi-ć^{PFV}
           DEL-drink-INF
           'to drink a little'
      b. po-pi-ja-ć^{IPFV}
           DEL-drink-ITER-INF
           'to drink from time to time'   (Polish)

(17) a. wy-bi-ć^{PFV}
           out-beat-INF
           'to kill off'
      b. wy-bi-ja-ć^{IPFV}
           out-beat-SI-INF
           'to kill off'
           'to be killing off'   (Polish)

A hiatus-blocking process is also present in the Czech examples in (18) and (19). This time, /v/ is inserted between the theme vowel -a and the imperfectivizing -a

---

[3]The accent is represented with the diacritic length mark.

(according to the standard analysis, /v/ is the imperfectivizing suffix itself; see e.g. Karlík et al. 1995: 194; consider also Matushansky's 2009: 397 unifying analysis of imperfectivizing suffixes in Russian, arguing that /v/ derives from an underlying back rounded yer). In addition, both examples also display a lengthening of the theme after adding the imperfectivizing suffix.

(18) a. po-kašl-a-t$^\text{PFV}$ si
          DEL-cough-TH-INF (self)
          'to cough a little'

     b. po-kašl-á-va-t$^\text{IPFV}$
          DEL-cough-TH-ITER-INF
          'to cough from time to time'   (Czech)

(19) a. při-děl-a-t$^\text{PFV}$
          at-do-TH-INF
          'to fix'

     b. při-děl-á-va-t$^\text{IPFV}$
          at-do-TH-SI-INF
          'to fix'
          'to be fixing'   (Czech)

However, the standard literature is against the analysis in which diminutive-iterative *po*-verbs are derived from delimitative predicates (see e.g. Isačenko 1960: 279–282, 1962: 407, Švedova 1980: 600, Zaliznjak & Šmelev 1997: 94, 104). According to them, delimitatives like (13a) are perfectiva tantum, diminutive-iteratives like (13b) are imperfectiva tantum and the verbs belong to different Aktionsarten: delimitative and diminutive-iterative. The authors claim that diminutive-iteratives are derived by circumfixation of *po*- and -YVA to the imperfective simplex predicate, i.e. to *ležat'* 'to lie' in the case of (13b) (see Katny 1994: 58ff. and Petr 1986: 419 for analogous claims with respect to Polish and Czech). This means that there is a strange coincidence. Both types of verbs have *po*- and some sort of a delimitative/diminutive meaning and the perfective verbs with the delimitative *po*- do not have a secondary imperfective counterpart and the imperfective verbs with *po*- (and -YVA) do not have a perfective counterpart. Tatevosov (2009: 133–134) shows that Isačenko's (1960) arguments are not strong enough and argues that there is a clear (derivational) relation between the existence of delimitative *po*-verbs and the existence of diminutive-iterative *po*-verbs in Russian.

The generative literature places delimitative *po*- and the imperfectivizing -YVA in different structural positions. The delimitative prefix is mostly higher than -

YVA; see e.g. Romanova (2004: 271–272), Tatevosov (2008: 437–438) and Klimek-Jankowska & Błaszczak (2022: 9, 2024: ex. (85), (92)) (but see also Součková 2004a: 408–409, who assumes circumfixation, and Jabłońska 2004: 377–381, who proposes two positions for *po*-, below and above -YVA). This placement has the advantage that it can account for why delimitative verbs are not (progressively) imperfectivized. However, it brings about two scope problems. First, it makes a false prediction with respect to morphological aspect. Given the position of the prefix being higher than -YVA, one expects diminutive-iterative verbs to be perfective contrary to the facts; see §1.1 again. Second, if the delimitative prefix scoped over -YVA, one would expect an interpretation with the repetition of standard actions in a delimited/short time frame, again contrary to the facts. For instance, in example (10b), repeated here as (20), one should receive a repetition of (normal) actions of reading in a short time frame.

(20) po-čit-yva-t'
      DEL-read-ITER-INF
      'to read from time to time'   (Russian)

However, (20) is interpreted as a repetition of short actions of reading. This indicates reversed scope properties, with -YVA being higher than *po*-.

Generally, there are three possibilities for how to derive diminutive-iterative *po*-verbs from *po*- delimitatives, as schematized in (21).

(21) a. [*po*- ... [-*yva*- ... [... √root]]]
      b. [-*yva*- ... [-*po*- ... [... √root]]]
      c. [-*yva*- + *po*- ... [... √root]]]

The first possibility, as just discussed, makes wrong predictions with respect to interpretational possibilities and morphological aspect properties. Thus, since diminutive-iteratives are imperfective and since morphological aspect is determined by the last (highest) aspectual morpheme (e.g. Karcevski 1927: 96, Vinogradov 1947: 500, Dostál 1954: 482, Isačenko 1962: 416–418, Zinova & Filip 2015: 601–602, Biskup 2020: 1–4, Tatevosov 2020: 28 and Zinova 2021: 36–38), the second option, shown in (21b), is preferred over the first possibility in (21a). Note that the delimitative *po*- cannot attach to the predicate after the imperfectivizing -YVA because there are no exceptions to the perfectivizing effect of prefixation (Smith 1997: 242); there are only apparent exceptions (e.g. Schuyt 1990: 292, Zaliznjak &

Petr Biskup

Šmelev 1997: 67–68, Zaliznjak 2017: 4–6).[4]

Given that interpretational properties of diminutive-iterative verbs are correctly derived by (21b), in contrast to (21a), the second option is preferred over the derivation in (21a) more generally. As to the circumfixation derivation in (21c), this option is generally disfavored, especially if both affixes exist independently (see e.g. Marušič 2023), as is the case with diminutive-iterative *po*-verbs. What is more, there is no reason to assume the special operation of circumfixation when (21b) successfully derives the relevant facts.

Since diminutive-iterative *po*-verbs cannot receive the progressive delimitative interpretation, only the iterative interpretation, as in (10b), (11b) and (12b), I assume that the iterative -YVA differs from the progressive -YVA with respect to their structural position. The progressive -YVA is generated below the delimitative *po-* and the iterative -YVA merges above the delimitative prefix, as illustrated in (22) (see Ramchand 2004: 33, who proposes that -YVA can occur in two different projections: Asp and Cuml).

(22) [...Iter -YVA ... [... Del *po-* ... [... Prog -YVA ... [... √root]]]]

An argument against diminutive-iteratives derived from delimitatives could be based on the fact that there are derivational chains with missing links (see e.g. Součková 2004a: 409). Specifically, in contrast to the Russian *pokašljat'* 'to cough a little' and the Polish *pokaszleć/pokasłać* 'to cough a little', Czech (and Slovak) do not have the middle step of the derivational chain *pokašlat* (there is only *pokašlat si*), as shown in (23b).[5]

(23) a. kašl-a-t
          cough-TH-INF
          'to cough'

      b. *po-kašl-a-t
          DEL-cough-TH-INF

---

[4] Zinova & Filip (2015: 605–607) argue that iterative *pere-* and attenuative *po-* do not have to perfectivize in Russian. They assume that imperfectives like *perezapisyvat'* 'to (be) rerecord(ing)' can be derived by attaching *pere-* to the imperfective *zapisyvat'* 'to (be) record(ing)'. Their supporting argument is however based on borrowings, which are known to be anomalous in various respects. Specifically, when the prefixes discussed attach to a borrowed biaspectual verb, the new verb is still biaspectual like *perekvalificirovat'* 'to requalify'. A comparison with other languages suggests that in Russian the verb is not adapted enough to be able to accept the perfectivizing effect of the prefix. In contrast, the Czech *překvalifikovat* 'to requalify' is already perfective and there is also its imperfective counterpart *překvalifikovávat*.

[5] The Polish situation is somewhat controversial but *pokaszleć/pokasłać* can be found in *Wielki słownik języka polskiego* (Żmigrodzki 2022).

    c.   po-kašl-á-va-t
       DEL-cough-TH-ITER-INF
       'to cough from time to time'   (Czech)

Such an argument, however, is valid only if we assume a lexicalist framework, in which verbs like diminutive-iteratives are derived from complete word forms (see e.g. the definition of derivational chains in Zinova & Filip 2015: 601-602). In morphosyntactic approaches like the one assumed here verbs are derived incrementally, morpheme by morpheme, in a bottom-up fashion, not by attaching the imperfectivizing *-va* in (23c) to the complete verb **pokašlat*.

    The same also holds for other morphemes in derivational chains, e.g. prefixes. Hence, *vypisat'* 'to excerpt' in (24a) is not derived by prefixation of *vy-* to the complete verb *pisat'* 'to write' since in the morphosyntactic approach assumed here, prefixes typically merge before the tense morpheme (infinitival *-t'* in (24)) and in certain cases also before the theme marker.[6]

(24)   a.   vy-pis-a-t'
           out-write-TH-INF
           'to excerpt'

        b.   pis-a-t'
           write-TH-INF
           'to write'

To conclude this section, the imperfectivizing marker -YVA can appear in different structural positions and consequently, it can spell out distinct semantic properties.

---

[6]An analogous reasoning also applies to disappearing elements, e.g. *si* 'self' in (i). One might argue that in cases like (i), incrementality is violated because *si* does not have to be present in (i.c), in contrast to (i.b). Again, since the verb in (i.c) is not derived from the complete verb with *si* in (i.b), there cannot be disappearing *si* in (i.c).

(i)   a.   křič-e-t
          shout-TH-INF
          'to (be) shout(ing)'

       b.   po-křič-e-t   *(si)
          DEL-shout-TH-INF   self
          'to shout a little for oneself'

       c.   po-křik-ova-t   (si)
          DEL-shout-ITER-INF self
          'to shout from time to time'   (Czech)

Petr Biskup

3 Analysis

In what follows, I elaborate the proposal from the preceding section, first from the semantic perspective, then from the syntactic point of view.

3.1 Combining *po-* and the verb stem

According to Isačenko (1962: 392), there are several restrictions on the formation of delimitative *po*-verbs. The restrictions are found not only in Russian, as shown in (25), (26) and (27). At least partially, they can be accounted for if it is assumed that the delimitative prefix needs a scale (for scale (degree) approaches to *po-* see Filip 2000, 2003; Jabłońska 2004, Součková 2004a,b; Kagan 2016 and Zinova 2021).

(25)   * po-zyskać
       DEL-win
       Intended: 'to win for a while'   (Polish)

(26)   * po-bodnout
       DEL-stab
       Intended: 'to stab for a while'   (Czech)

(27)   * po-stoit'
       DEL-cost
       Intended: 'to cost for a while'   (Russian)

Given that achievements denote a momentaneous change of state, there is no (protracted) scale that could be delimited by the prefix and cases like (25) are ungrammatical.[7] In the same vein, since semelfactive verbs denote punctual eventualities, the delimitative *po-* cannot measure them and examples like (26) are ungrammatical as well.

An interesting case is stative predicates like the one in (27), which also cannot be prefixed with the delimitative *po-*. The ungrammatical status cannot be ascribed to the fact that a temporal extent scale is missing because the base predicate of such states can be modified by durative adverbials, as is the case e.g. with Russian *znat' kogo tri dnja* 'to know somebody for three days'. The reason also cannot be the homogeneity requirement of delimitative *po-* (Filip 2000: 61, 2003: 91; Mehlig 2006: 247) because states are homogeneous. Atelicity does not play a

---

[7]The so-called degree achievements are known for their special behavior; they provide an appropriate change-of-state scale.

role either because states are atelic and other atelic predicates – e.g. activities like the Polish *płakać* 'to cry' – can be prefixed with delimitative *po-*. So, one might argue that the problem lies in the concept of change because stative predicates do not entail a change, in contrast to dynamic predicates (activities, accomplishments and achievements). However, this reasoning is not correct, either, since there are also states that can be prefixed with the delimitative *po-*; consider the examples (28), (29) and (30).

(28) Russian
    a. po-stojat'   b. po-sidet'   c. po-spat'
       DEL-stand   DEL-sit   DEL-sleep

(29) Polish
    a. po-stać   b. po-siedzieć   c. po-spać
       DEL-stand   DEL-sit   DEL-sleep

(30) Czech
    a. po-stát   b. po-sedět   c. po-spat
       DEL-stand   DEL-sit   DEL-sleep
       'to stand for a while'   'to sit for a while'   'to sleep for a while'

The unprefixed verbs in (28)–(30) belong to the class of "interval statives" (Dowty 1979: 173–180) and differ from "static states" (Bach 1986: 6) like 'to know', 'to own' and 'to cost', which are less dynamic and do not accept the delimitative *po-* in North Slavic. Maienborn (2003, 2005) analyzes dynamic states like 'to stand' as "Davidsonian states" because they refer to eventualities in the sense of Davidson, so they introduce a Davidsonian event argument. In contrast, static states belong to "Kimian states" in her analysis. They do not have a Davidsonian event variable but introduce a specific Kimian-state referential argument. In fact, the German modifier *ein bisschen* 'a little bit' distinguishing Davidsonian states from Kimian states in her eventuality diagnostic (Maienborn 2005: 297–299) behaves like delimitative *po-* with respect to grammaticality judgements. Here I follow Maienborn's proposal and assume that delimitative *po-* selects a predicate with a Davidsonian event argument (that in addition has some scalar structure, as discussed above).

According to Filip (2000: 61–66, 2003: 89–90), delimitative *po-* applies to a homogeneous predicate and contributes an extensive measure function, which is contextually specified and meets or falls short of some contextually determined expectation value. Součková (2004a: 410, 2004b: 73) modifies Filip's proposal and

Petr Biskup

argues that the measure function only applies to events, as shown in (31).[8]

(31) $[\![po\text{-}]\!] = \lambda P \lambda e[P(e) \wedge m(e) = c_{\text{relatively.small}}]$

P is a variable over predicates, *m* stands for the extensive measure function applied to an event and *c* means that its value is contextually determined. The function *m* measures events that contain some scalar structure, concretely, it measures the degree of change on the appropriate scale and it can apply to various dimensions (types of scales). It depends on lexical properties of the particular verb, which dimension – if at all; recall the discussion of achievements and semelfactives in (25) and (26) – is accessible. For instance, in the case of motion verbs, there is a scale of progress along the path; in degree achievements, there is an increase in the degree on a property scale; but in most cases, the prefix is applied to a time scale. Note, however, that it is not possible to reduce all delimitative cases to the time scale and that particular scales do not have to coincide with respect to the degree of change. For instance, in the case of Czech *poponést* 'to carry sth. a little' in (9b), it is the progress on the appropriate path that is measured and the relatively small value of this progress (the short path) can be in contrast to the duration of the carrying eventuality, which under appropriate circumstances can even be very long.

As just mentioned, (31) states that the value of the measure function is relatively small in the context. I use the equal relation in (31) instead of the less-than-or-equal relation (in contrast to e.g. Kagan 2015: 47, 2016: 310; Klimek-Jankowska & Błaszczak 2024: ex. (68)) since it directly brings about the quantization property of *po*-delimitatives. necessary for iteration, as discussed in the following section. More concretely, if a delimited event of crying takes five minutes in the specific context, i.e. the value of the measure function is contextually determined to be equal to five minutes, then there will be no proper part of the event that also falls in the denotation of this crying *po*-predicate. In contrast, if the less-than-or-equal relation were used in (31), there could be a proper part – e.g. crying for just four minutes – falling in the denotation of the crying *po*-predicate.

Thus, building on Součková's proposal, the meaning of the Polish *popłaka(ć)* 'to cry for a while' – after applying the delimitative prefix *po-* to the predicate *płaka(ć)* 'to cry' – will look like (32).

---

[8]The proposal that the measure function applies to events can be already found in Piñón (1994: 362–363). Součková (2004a,b) also argues that the prefix can also apply to non-homogeneous predicates since she unifies delimitative *po-* and the attenuative *po-*. Given that I keep the two prefixes apart (see §1.3), I assume that the delimitative *po-* applies to homogeneous predicates. The attenuative po- then probably only applies to non-homogeneous predicates, as in the case of the Russian *poprideržat'* 'to hold gently'

(32) $[\![po\text{-}płaka\text{-}]\!] = \lambda e[\text{CRY}(e) \wedge m(e) = c_{\text{relatively.small}}]$

Given the meaning in (31), with P as a predicate over events and the measure function applied to an event, it is obvious that delimitative *po-* cannot combine with states like *stoit'* 'to cost' in (27), which only have the Kimian-state referential argument.

3.2 Combining the delimitative *po*-predicate and the iterative *-YVA*

According to the proposal in §2, diminutive-iteratives are derived from delimitative predicates in the way that the eventuality delimited/measured by *po-* is iterated by means of *-YVA*. In other words, delimitative *po-* is responsible for individuation and the iterative *-YVA* then for pluralization. Thus, given that for iteration and counting, discrete elements are necessary, the question arises as to which concept is responsible for the individuation here. Is quantization sufficient or is telicity necessary as well? According to Filip (2003: 91), delimitative *po-* ("attenuative" in her terms) makes predicates not only quantized, analogously to measure functions like *a (relatively) small quantity* and *a few*, but also telic given her definition of telicity based on atomicity (see Filip 2003: 60–61). However, if the standard adverbial test is a reliable diagnostic for (a)telicity, then delimitative *po*-verbs must be atelic because they are compatible with durative adverbials like 'an hour (long)' and incompatible with time-span adverbials like 'in an hour' in the relevant reading. In the light of grammatical diminutive-iteratives like *polëži-vat'* 'to lie from time to time' in (1), that in turn means that quantization brought about by an extensive measure function like the delimitative *po-* is sufficient for applying the iterative *-yva-* and that telicity is not a necessary condition in this case.

As mentioned in §1.1, the iterative marker brings about an unspecified number of instances of the particular eventuality and the cardinality of repetitions is contextually determined. Since it is difficult to determine the smallest number of repetitions here, I assume the weakest position and use the meaning of plurals as a base. This means that the cardinality of the iterated eventuality is greater than one, as shown in (33), where $|e|$ stands for the number of atomic events.

(33) $[\![\text{ITER}]\!] = \lambda P \lambda e[P(e) \wedge e = \sigma e'[P(e') \wedge e' \subset e \wedge |e| > 1] \wedge \forall e'[\text{ATOM}(e') \rightarrow \neg \exists e''[P(e'') \wedge e'' \subset e \wedge \text{ATOM}(e'') \wedge \tau(e') \supset\subset \tau(e'')]]]$

For deriving pluralities, usually, Link's (1983) σ-operator and the ∗-operator are used. Therefore, for the iterative *-YVA*, I use Kratzer's proposal (Kratzer 2008: 296, see also Boneh & Doron 2008: ex. (31) and Ferreira 2016: 358), according to

which the sum of all events e' that are proper parts of the event e and have the property P is identical to e. Since it brings about a weak notion of plurality, with singularities as special cases, I also use the conjunct $|e| > 1$ in (33), as discussed above.

The iterated events are not temporally adjacent, e.g. the meaning of the Polish *popłakiwać* 'to cry from time to time' in (2) is characterized by Żmigrodzki (2022) as *płakać z przerwami* 'to cry with pauses' (see also Katny 1994: 67). For this reason, I add to the meaning in (33) the restriction on temporal traces. Specifically, for every atomic proper part e' of the event e, it holds that there is no atomic subevent e'' with the property P that is a proper part of e and whose temporal trace abuts with the temporal trace of e'. Under the assumption that the abut relation (precluding any contact) is stronger than the overlap relation, it holds that if the temporal traces of e' and e'' do not abut, then they also do not overlap. The size of the time interval between e' and e'' is not defined here since it depends on the lexical meaning of the appropriate predicate and on the context. Moreover, intervals between the particular subevents can be of different sizes.

I do not use the classical non-overlap condition by Lasersohn (1995: 256) (see also e.g. Wood 2007: 126), with the function f standing for temporal, spatio-temporal or participant-based distributivity, since in the case of diminutive-itera-tive *po*-verbs, participants and spaces can overlap. As an illustration, consider example (34), in which the referent of the expression *Kasia* is identical for the crying subevents.[9] The crying subevents of example (34) also can (but do not have to) happen in an identical space.

(34) Kasia cały   dzień po-płak-iwa-ł-a.
      Kasia whole day   DEL-cry-ITER-PTCP.F
      'Kasia cried repeatedly the whole day.'   (Polish)

To receive the separated-in-time reading, Lasersohn (1995: 254) adds a between-ness condition to the non-overlap condition, which introduces a time that inter-venes between temporal traces of the singular events. I do not follow his proposal because I assume that the time variable is introduced by the aspectual head later in the derivation. Instead of the betweenness and non-overlap conditions, I use the abut condition, as stated in (33).

When the meaning of the iterative operator is applied to the meaning of the de-limitative predicate *popłaka(ć)* in (32), we obtain a diminutive-iterative predicate over events, as demonstrated in (35).

---

[9] With plural subjects, distributive readings are possible but not necessary.

(35) $[\![\textit{popłakiwa-}]\!]$
$= \lambda e [*\textsc{cry}(e) \wedge m(e) = c_{\text{relatively.small}} \wedge e = \sigma e'[*\textsc{cry}(e') \wedge m(e') = c_{\text{relatively.small}} \wedge e' \subset e \wedge |e| > 1] \wedge \forall e'[\textsc{atom}(e') \rightarrow \neg \exists e''[*\textsc{cry}(e'') \wedge m(e'') = c_{\text{relatively.small}} \wedge e'' \subset e \wedge \textsc{atom}(e'') \wedge \tau(e') \supset\subset \tau(e'')]]]$

Thus, having established the meaning of iterative -YVA, let us now look at the lower instantiation of the marker, progressive -YVA.

3.3 The progressive -YVA and imperfectivity

As already mentioned, when -YVA attaches to a delimitative predicate, the form cannot have the progressive meaning. It can only receive an iterative interpretation; consider e.g. the Russian *počityvat'* 'to read from time to time' in (10b) again. This results from the splitting of -YVA into two different syntactic positions and from the positioning of the delimitative *po-* between them, as discussed in §2. So, what is the meaning of the progressive -YVA?

Progressivity is often defined in terms of partitivity (Filip 1999: 171–175, 213 and references therein), as is (secondary) imperfectivity (e.g. Łazorczyk 2010: 134–139). Although progressivity and (Slavic) imperfectivity are close notions (Zucchi 1999: 200), they are not identical; see e.g. Comrie (1976: 33) and Dahl (1985: 92). Recall also from §3.2 that the aspectual projection, which is going to encode (im)perfectivity in the current proposal, occurs in a higher syntactic position, hence it is not identical to the progressive projection spelled out by the progressive *-yva-*. Thus, since the progressive brings about an internal part of the eventuality, I assume the (for simplicity extensional) meaning in (36) for the progressive -YVA.

(36) $[\![\textsc{prog}]\!] = \lambda P \lambda e \exists e'[P(e') \wedge e < e']$

It is based on Krifka (1992: 47) progressive operator but uses the proper-part-of relation instead of just part-of relation. The reason is that we need to exclude the possibility that the event culminates with the progressive -YVA.[10]

Given the proposal that there are two different syntactic positions (Prog and Iter; see (22) again) with distinct meanings that are spelled out as -YVA, the question arises whether the progressive -YVA can co-occur with the iterative -YVA. It is possible to test it with Czech since it allows combinations of more -YVA markers.

---

[10]Filip (2005) uses the part-of relation for the meaning of the imperfective morphological aspect since it can also derive the meaning of general-factuals, which can refer to a culminated eventuality. Also Łazorczyk (2010) uses the part-of relation in the meaning of her secondary imperfective operator.

Petr Biskup

Consider example (37), in which the imperfectivizing *-vá* is attached to the un-prefixed perfective stem *dá-* and then another *-(y)va* marker is adjoined, forming the imperfective *dávával*.[11]

(37) (*Včera) dá-vá-va-l peníze chudým.
     yesterday give-PROG/ITER-HAB-PTCP money poor.ADJ.DAT.PL
     (*Intended*:) '(Yesterday) he had the habit of giving money to the poor.'

Although the lower *-vá* can bring about the progressive or the iterative interpreta-tion, the higher *-va* can only bring about the habitual meaning in (37). Therefore, the sentence cannot be interpreted non-habitually. The ungrammatical status of the adverbial 'yesterday' indicates that an episodic reading is impossible. If an it-erative adverbial is used, the sentence must also receive a habitual interpretation, as shown by the translation in (38).

(38) Dá-vá-va-l peníze chudým dvakrát.
     give-PROG/ITER-HAB-PTCP money to.poor twice
     'He had the habit of giving money to the poor twice.'

For instance, a person could have the habit of giving money to the poor daily at 9 a.m. and 7 p.m. In contrast, it is not possible that there were only two occasions of giving money to the poor in total. Thus, the outer -YVA in cases like (38) always instantiates a habitual operator that is structurally higher than the iterative and progressive suffixes.[12]

The incompatibility of the iterative and progressive -YVA has a semantic rea-son. As discussed in §3.2, the iterative operator applies to quantized predicates. However, the progressive operator derives a homogeneous (i.e. cumulative and divisive) predicate, which cannot serve as an input for the iterative -YVA. Accord-ing to Łazorczyk (2010: 137–138), the secondary imperfective operator takes a telic predicate and returns a homogeneous subpart of it (see also Filip 1999: 167,

---

[11]The imperfective *dáva-(t)* can have either the progressive or the iterative interpretations, hence I gloss *-vá* with PROG/ITER.

[12]For more on the habitual -YVA, see Filip & Carlson (1997), Esvan (2007: 321–340), Berger (2009), Nádeníček (2011: 132–135), Nübler (2017), and also Biskup (2023), who shows that the habitual -YVA differs from the secondary imperfective -YVA – the progressive and iterative -YVA in the current approach – in phonological, aspectual and interpretational properties. Interpretation-ally, the habitual -YVA brings about a (generic) quantificational semantics in contrast to the pluractional iterative -YVA. Since the habitual -YVA is beyond the scope of this article, I will not discuss it any further here.

who argues that progressive sentences are cumulative).[13] Unsurprisingly, this behavior of the progressive operator corresponds with morphological aspect properties; we saw in §1.2 that the imperfectivizing -YVA selects a perfective stem and derives an imperfective predicate.[14]

Consequently, since the progressive -YVA selects a perfective, telic stem, the *po*-verbs *popisyvat'* 'to write from time to time', *popłakiwać* 'to cry from time to time' and *posedávat* 'to sit from time to time' cannot be derived in the way schematized in (39), i.e. as "delimited progressives". The reason for this is that in derivations like (39), in contrast to what was just said, the progressive -YVA would have to apply to stems that are imperfective, homogeneous and atelic: 'to write', 'to cry' and 'to sit'.

(39) a. * [po-[pis-yva]]-t'
           DEL-write-PROG-INF   (Russian)

      b. * [po-[płak-iwa]]-ć
           DEL-cry-PROG-INF   (Polish)

      c. * [po-[sed-á-va]]-t
           DEL-sit-TH-PROG-INF   (Czech)

Another alternative would be to assume that the progressive reading of *po*-verbs with -YVA is excluded because the delimitative *po-* applies first and makes the predicate quantized and individuated, with the consequence that proper parts of the denoted event are not accessible to the progressive -YVA (which applies as second). This proposal, however, would be very restricted because it could only work in the case of an incremental theme (see Filip 2005: 273). Let us test it with other prefixes.

---

[13]See also Tatevosov (2015: 489), who builds on Paslawska & von Stechow (2003: 346) and treats the imperfectivizing -YVA as an eventizer.

[14]Note that despite the fact that the iterative -YVA and the progressive -YVA cannot co-occur, sentences with a diminutive-iterative verb can receive a simultaneous interpretation when the reference time (expressed e.g. by the temporal clause) is included in the event time, as in (i), based on (34).

(i) Kiedy wszedł   do pokoju, Kasia po-płak-iwa-ł-a.
     when came.3SG.M into room   Kasia DEL-cry-ITER-PTCP-F
     'When he came into the room, Kasia was crying a little.'

Petr Biskup

(40) Pavel teď vy-pis-uj-e   / o-pis-uj-e   celou
     Pavel now out-write-PROG-3.SG   about-write-PROG-3.SG whole
     přednášku.
     talk
     'Pavel is excerpting/copying the whole talk right now.'   (Czech)

Although there is a secondary imperfective predicate with prefixes making the base verb quantized, combined with a quantified incremental theme in (40), the progressive reading is available. Thus, this alternative explanation of the impossibility of the progressive delimitative interpretation also does not work.

At this point, the question arises as to how the difference between progressivity and imperfectivity is modelled in the current approach. The meaning of the progressive operator in (36) shows that progressivity concerns the internal structure of eventualities. As to the morphological aspect, I make the standard assumption that it concerns temporal properties of eventualities, i.e. relates the event time to the reference time via the inclusion relation (see e.g. Paslawska & von Stechow 2003: 322). Concretely, in the case of the imperfective aspect, the reference time is included in the event time.

3.4 Morphosyntactic derivation

Building on Biskup (2019: 36–42, 2020), I assume that the value of the morphological aspect of diminutive-iterative *po*-verbs is determined in the aspectual head via the operation Agree. Combining the structure in (22) with the proposal that all aspectual markers are separated from the aspectual interpretation of the aspectual head (Biskup 2023), we receive (41), with parts relevant to our discussion.[15]

(41) [Asp$_{[Asp:?]}$ ... [Iter -YVA$_{[IPFV]}$ ... [Del *po*-$_{[PFV]}$ ... [Prog -YVA$_{[IPFV]}$ ...
     [√root]]]]]

It shows that aspectual markers like the iterative -YVA, the delimitative *po*- and the progressive -YVA bear an aspectual feature – either with the value [perfective] or [imperfective] – which can value the unvalued aspectual feature of the aspectual head.

---

[15]With the aspectual interpretation of the Asp head, the relation between the reference time and the event time is meant, as discussed in the last paragraph of §3.3. This interpretation is different from the semantics of particular aspectual markers, like prefixes, the semelfactive and habitual suffixes and the iterative and progressive -YVA, as discussed in §3.2 and §3.3 (see also Paslawska & von Stechow 2003, Grønn 2004, Tatevosov 2011, 2015).

Concerning the aspectual projection, it is standardly placed above the projection introducing the agent (see Babko-Malaya 2003, Błaszczak & Klimek-Jankowska 2012, Gribanova 2015). This placement is also supported by the fact that although Russian *-nie* nominals are aspectless (e.g. Schoorlemmer 1995, Gehrke 2008), they can have an agent argument, as demonstrated by the agent-oriented modifier in (42a) and the agentive *by*-phase in (42b).[16]

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

      b. soveršenie   prestuplenija licom…
         perpetration delict.GEN.SG person.INS.SG
         'a perpetration of the delict by a person'

                                              (Russian; Biskup 2023: ex. (69))

It has been argued that the secondary imperfective suffix merges inside the verbal domain below the agent (see e.g. Romanova 2004: 272 and Tatevosov 2015: 488 for Russian, Kwapiszewski 2022 for Polish and Biskup 2023 for Czech). However, given the splitting of -YVA into the iterative and progressive -YVA in the current proposal, we need to know more about the positioning of the agent. There are agent nominalizations ending in *-tel'* in Russian, *-ciel* in Polish and *-tel* in Czech that can contain the imperfectivizing -YVA. Semantically, the suffixes *-tel'*, *-ciel* and *-tel* (and others, like the agentive version of the Czech *-č* in (44)) relate to the projection containing the agent (external) argument since they introduce an entity – predominantly, a person – that carries out the action denoted by the predicate to which they are attached.[17] Thus, based on Baker & Vinokurova's (2009: 531) analysis of nominalizing affixes like the English *-er*, I consider the morphemes *-tel'*, *-ciel* and *-tel* to be nominal versions of the agentive Voice head. What is crucial for us is that the imperfectivizing -YVA is always closer to the root than the agentive nominalizing suffix, as demonstrated in (43a).

(43) a. do-pis-ova-tel
         to-write-ITER-NMLZ
         'correspondent'

---

[16]Regarding the aspectless status of Russian *-nie* nominals, a reviewer asks whether (42a) can have a durative modifier. Given that *for*-adverbials and *in*-adverbials diagnose (a)telicity rather than (im)perfectivity, this test is not effective.

[17]In the traditional terminology, nouns in *-tel* are called *nomina agentis*.

Petr Biskup

b. ob-jev-i-tel   Plut-a
   about-show-TH-NMLZ Pluto-GEN
   'the discoverer of Pluto'   (Czech)

Since (43a) refers to a person repeatedly performing the event of making a report, I take the *-ova* suffix to represent the iterative head. Note that the iterative meaning cannot be a property of *-tel* since this suffix also attaches e.g. to perfective predicates denoting a single event of discovering without changing the cardinality of the action, as shown in (43b). Thus, the order of the imperfectivizing -YVA and the nominalizing suffix can be taken to mean that the projection of Voice – hosting the agent – is structurally higher than the iterative projection spelled out by -YVA.[18]

I am not aware of diminutive-iterative *po*-predicates with *-tel'*, *-ciel* or *-tel* but there is at least the deverbal nominalization *pojídač* 'eater', in which the agentive *-č* occurs outside the diminutive-iterative predicate *pojída(t)* 'to eat from time to time', as shown in (44). The presence of the iterative meaning (in *-a*) is confirmed by the fact that the singular complement *krevety* is ungrammatical in contrast to the plural *krevet*.

(44)   po-jíd-a-č   krevet   / *krevety
      DEL-eat-ITER-NMLZ shrimp.PL   shrimp.SG
      'shrimp eater'   (Czech)

Building on these nominalization facts and the placement of the aspectual head above the projection introducing the agent, the relevant piece of morphosyntactic structure of diminutive-iterative *po*-verbs looks like (45), which is based on (41). Note that I added the standard verbalizing head *v* and that the structure does not contain the progressive projection now since the progressive reading of delimitative *po*-verbs with -YVA is excluded.

(45)   [Asp$_{[Asp:IPFV]}$ [Voice [Iter -YVA$_{[IPFV]}$ [Del *po*-$_{[PFV]}$ [*v* [√root]]]]]]

Given that the aspectual value is determined by the aspectual marker that is attached last (i.e. that is closest to Asp; see discussion in §2), when the aspectual head probes, then it finds the aspectual feature of the iterative -YVA first. Consequently, the Agree operation uses this [imperfective] feature and diminutive-iterative *po*-verbs always occur as imperfective, as shown in (45).

---

[18]The vocabulary item -YVA can be inserted into the head of the iterative and progressive projections since it is specified as [imperfective].

Structures of diminutive-iterative *po*-verbs are not difficult to linearize since all affixes are structurally higher than the root, as illustrated for the elements under discussion in (45). If the Affix-Specific Linearization by Harley (2013) is assumed – which encodes the prefixal versus affixal property directly in the specific marker –, then no head movement is necessary. Only argument phrases need to be evacuated from the extended verbal projection. Consider e.g. the relevant part of linearization of the Polish verb *popłakiwać* 'to cry from time to time' in figure 1.
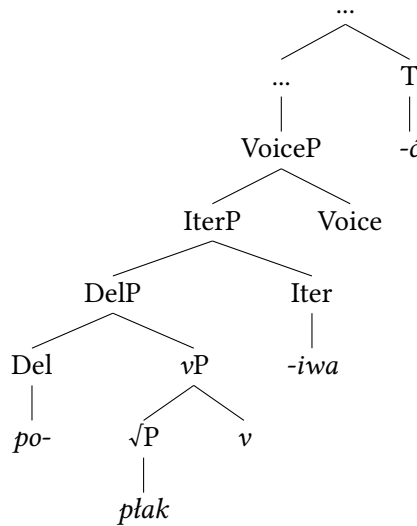
Figure 1: Linearization of Polish *popłakiwać* 'to cry from time to time'

The verbalizing head *v* is phonologically empty in this derivation but in other cases a theme vowel can be inserted. Since theme vowels are affixes, the *v* head is placed to the right also here. Delimitative *po-* and iterative *-iwa* are placed to the left and to the right, respectively, in accordance with their prefixal and suffixal status. I do not decompose imperfectivizing suffixes in this article but it is possible to split them, e.g. the Russian *-yva* into *-yv-a*, analogously the Polish *-ywa* into *-yw-a* and the Czech *-ova* into *-ov-a*. The second element (*-a*) could be analyzed as a theme vowel that, e.g. in the structure of *popłakiwać*, spells out the Voice head. Therefore, I put the head to the right in (1). Finally, *-ć* can be taken to represent the infinitival T head, linearized to the right in accordance with the suffixal status of the marker.

Petr Biskup

One could assume that (1) in fact represents the syntactic structure already before linearization (if one prefers a more powerful syntactic module). This would however go against the standard SVO analysis of Slavic languages, which supposes that heads take their complements to the right in Slavic (but see also Haider & Szucsich 2022: 34–35).

Delimitative *po-* is not the only prefix displaying the special, iterative type of behavior. Comitative *pod-*, e.g. in the Russian *podpevat'* and in the Polish *podśpiewywać*, both with the meaning 'to sing a little with sth.', seems to behave in the same way.

Flier (1985: 41) shows that perdurative ("delimitative" in his terms) *pro-* manifests very similar behavior; consider example (46).

(46) a. Často vesennie večera   ona pro-siž-iva-l-a   na vysokom
        often spring   evenings she through-sit-ITER-PST-F.SG on high
        kryl'ce.
        porch
        'She would often spend entire evenings in the spring sitting on the
        high porch.'

    b. * Ona dolgo prosiživala^IPFV   na vysokom kryl'ce.
        she long   through.sit.ITER.PST.F.SG on high   porch
        Intended: 'She would spend a long time sitting on the high porch.'

    c. * Ona sejčas prosiživaet^IPFV   na vysokom kryl'ce.
        she now   through.sit.ITER.PRS.SG on high   porch
        Intended: 'She is now spending time sitting on the high porch.'

                                                    (Russian; Flier 1985: 41)

As shown by (46a), the event of sitting is iterated. The incompatibility of the imperfective verb with the adverbial *dolgo* 'for a long time' in (46b) demonstrates that the durative meaning is excluded and the incompatibility of the verb in present with the adverbial *sejčas* 'now' in (46c) shows the impossibility of the progressive meaning. A larger context showing the impossibility of the progressive reading is provided in (47).

(47) * Ona pro-siž-iva-l-a   vesennie večera   na vysokom kryl'ce
      she through-sit-ITER-PST-F.SG spring   evenings on high   porch
      v tot moment, kogda bitva načalas'.
      in the moment when battle begin.PST
      Intended: 'She was spending entire evenings in the spring sitting on the
      high porch when the battle began.'

Thus, this perdurative *pro-* seems to be the next candidate for the positioning between the progressive and iterative -YVA.

The current proposal has consequences for the overall architecture of verbal predicates. Before the splitting of the secondary imperfective marker into the progressive and the iterative -YVA, there were two structural possibilities for verbal prefixes: below and above the secondary imperfective morpheme. Now, after the splitting, there are three options, as shown in (48).

(48) [Asp [Voice [SP$_{\text{high}}$ [Iter -YVA [SP$_{\text{high}}$ (Del *po-*) [Prog -YVA [SP$_{\text{low}}$ [v [√root [LP]]]]]]]]]]

Specifically, (i) below the progressive -YVA for (lexical, intermediate and lower superlexical) prefixes in predicates that can receive the progressive and iterative interpretations, like the Russian *pro-davat'*[IPFV] 'to sell'; (ii) between the progressive and the iterative -YVA for higher superlexical prefixes in predicates which cannot be progressivized but can have an iterative interpretation, like delimitative *po-* in diminutive-iteratives and perdurative *pro-* in (46); (iii) above iterative -YVA for higher superlexical prefixes in predicates which cannot be imperfectivized – like ingressive *roz-* in the Polish *rozboleć*[PFV] 'to start to ache' – and for higher superlexical prefixes which perfectivize a secondary imperfective predicate, like distributive *po-* in the Russian *po-vytalkivat'*[PFV] 'to push out one after another'.[19]

4 Conclusions

I have argued that diminutive-iterative *po-*verbs are derived from delimitative *po-*predicates. The secondary imperfective marker is split into two distinct elements, both syntactically and semantically: iterative -YVA and progressive -YVA. The iterative marker has a pluractional meaning and merges higher than delimitative *po-*, whereas progressive -YVA is a partitive operator that occurs below the prefix. In diminutive-iterative verbs, the progressive operator is not present and delimitative *po-*, with its measure function meaning, applies to the simplex predicate. Then, the event denoted by the quantized predicate is iterated by the pluractional

---

[19] Mehlig (2007, 2012) discusses examples of delimitative *po-*verbs in Russian like *pootkryvat' okno* 'to open a window for a while', which at first sight, suggest that the delimitative *po-* can also merge higher than the iterative *-yva*. Such predicates denote an attempt to attain the change of state through several different actions. These cases probably are not problematic since they could be analyzed in terms of the partitive progressive operator (which I argued to be lower than the delimitative *po-*), as in Tatevosov & Ivanov (2009).

Petr Biskup

imperfectivizing -ʏᴠᴀ. I have discussed certain restrictions on the formation of delimitative *po*-verbs and argued that only eventualities with a scalar structure and an event variable can be delimited. Static states, which contain the Kimian state referential argument, are not compatible with delimitative *po*-. As to morphosyntactic structure, I have argued that the iterative projection, spelled out by -ʏᴠᴀ, occurs inside the verbal domain below the Voice projection, which introduces the external argument. Morphological aspect properties of diminutive-iteratives are determined in the aspectual projection via Agree with the closest aspectual feature, i.e. the imperfective feature of the iterative head. We have also seen that the overall (aspectual) architecture of Slavic predicates is more fine-grained and that there are more prefixal positions than usually assumed.

Abbreviations

ᴅᴇʟ	delimitative
ꜰ	feminine
ɢᴇɴ	genitive
ʜᴀʙ	habitual
ɪɴꜰ	infinitive
ɪɴs	instrumental
ɪᴘꜰᴠ	imperfective
ɪᴛᴇʀ	iterative
ᴍ	masculine

ɴᴍʟᴢ	nominalizer
ᴘꜰᴠ	perfective
ᴘʟ	plural
ᴘʀᴏɢ	progressive
ᴘʀs	present
ᴘsᴛ	past
ᴘᴛᴄᴘ	participle
sɢ	singular
sɪ	secondary imperfective
ᴛʜ	theme (vowel)

Acknowledgments

References

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

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9(1). 5–16. DOI: 10.1007/BF00627432.

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 (Hg.) (eds.), *Von Zuständen, Dynamik und Veränderung bei Pygmäen und Giganten: Festschrift zum 60. Geburtstag von Walter Breu*, 25–43. Bochum: Universitätsverlag Dr. Norbert Brockmeyer.

Biskup, Petr. 2019. *Prepositions, case and verbal prefixes: The case of Slavic*. Amsterdam: John Benjamins.

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

Biskup, Petr. 2023. Aspect separated from aspectual markers in Russian and Czech. In Petr Biskup, Marcel Börner, Olav Mueller-Reichau & Iuliia Shcherbina (eds.), *Advances in formal Slavic linguistics 2021*, 61–98. Berlin: Language Science Press. DOI: 10.5281/zenodo.10123631.

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. https://lsa.umich.edu/content/dam/slavic-assets/slavic-documents/20FASL.pdf.

Boneh, Nora & Edit Doron. 2008. Habituality and the habitual aspect. In Susan Rothstein (ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*, 321–347. Amsterdam: John Benjamins. DOI: 10.1075/la.110.14bon.

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

Czochralski, Jan A. 1975. *Verbalaspekt und Tempussystem im Deutschen und Polnischen: Eine konfrontative Darstellung*. Warszawa: Państwowe Wydawnictwo Nauk.

Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.

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

Dostál, Antonín. 1954. *Studie o vidovém systému v staroslověnštině*. Praha: Státní pedagogické nakladatelství.

Dowty, David R. 1979. *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. Dordrecht, Boston & London: Reidel. DOI: 10.1007/978-94-009-9473-7.

Petr Biskup

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

Ferreira, Marcelo. 2016. The semantic ingredients of imperfectivity in progressives, habituals, and counterfactuals. *Natural Language Semantics* 24. 353–397. DOI: 10.1007/s11050-016-9127-2.

Filip, Hana. 1999. *Aspect, eventuality types and noun phrase semantics.* New York & London: Garland Publishing. DOI: 10.4324/9780203827413.

Filip, Hana. 2000. The quantization puzzle. In Carol Tenny & James Pustejovsky (eds.), *Events as grammatical objects: The converging perspectives of lexical semantics and syntax*, 39–95. Stanford, CA: CSLI Publications.

Filip, Hana. 2003. Prefixes and the delimitation of events. *Journal of Slavic Linguistics* 11(1). 55–101. https://www.jstor.org/stable/24599705.

Filip, Hana. 2005. Measures and indefinites. In Greg N. Carlson & Francis Jeffry Pelletier (eds.), *Reference and quantification: The Partee effect*, 229–288. CSLI Publications. DOI: 10.1007/1-4020-3232-3_7.

Filip, Hana & Gregory N. Carlson. 1997. Sui generis genericity. *University of Pennsylvania Working Papers* 4. 91–110. https://repository.upenn.edu/handle/20.500.14332/45451.

Flier, Michael S. 1985. The scope of prefixal delimitation in Russian. In Michael S. Flier & Alan Timberlake (eds.), *The scope of Slavic aspect*, 41–58. Columbus, OH: Slavica Publishers.

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

Grønn, Atle. 2004. *The semantics and pragmatics of the Russian factual imperfective.* Oslo: University of Oslo. (Doctoral dissertation).

Haider, Hubert & Luka Szucsich. 2022. Slavic languages – "SVO" languages without0 SVO qualities? *Theoretical Linguistics* 48(1–2). 1–39. DOI: 10.1515/tl-2022-2035. https://doi.org/10.1515/tl-2022-2035.

Isačenko, Aleksandr V. 1960. *Grammatičeskij stroj russkogo jazyka. Morfologija. Častj vtoraja [Grammatical structure of the Russian verb. Morphology. Part two].* Bratislava: Vydavatelstvo Slovenskej Akademie vied.

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

Jabłońska, Patrycja. 2004. When the prefixes meet the suffixes. *Nordlyd* 32(2). 363–401. DOI: 10.7557/12.73.

Kagan, Olga. 2015. *Scalarity in the verbal domain: The case of verbal prefixation in Russian.* Cambridge: Cambridge University Press. DOI: 10 . 1017 / CBO9781316136195.

Kagan, Olga. 2016. Measurement across domains: A unified account of the adjectival and the verbal attenuative *po-. Journal of Slavic Linguistics* 24(2). 301–336. DOI: 10.1353/jsl.2016.0013.

Karcevski, Serge. 1927. *Système du verbe russe: Essai de linguistique synchronique.* Praha: Plamja.

Karlík, Petr, Marek Nekuta & Zdenka Rusínová (eds.). 1995. *Příruční mluvnice češtiny.* Praha: Nakladatelství Lidové noviny.

Katny, Andrzej. 1994. *Zu ausgewählten Aktionsarten im Polnischen und deren Entsprechungen im Deutschen.* Rzesów: Wydawnictwo WSP.

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. https://ojs.ung.si/index.php/JSL/article/ view/99.

Klimek-Jankowska, Dorota & Joanna Błaszczak. 2024. The interaction of idioms and aspect in Polish. *Natural Language & Linguistic Theory* 42(1). 153–207. DOI: 10.1007/s11049-023-09577-2.

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

Kratzer, Angelika. 2008. On the plurality of verbs. In Johannes Dölling, Tatjana Heyde-Zybatow & Martin Schäfer (eds.), *Event structures in linguistic form and interpretation*, 269–300. Chicago, IL/Berlin: De Gruyter. DOI: 10 . 1515 / 9783110925449.269.

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: CSLI Publications.

Kwapiszewski, Arkadiusz. 2022. *Aspect and event structure: The morphosyntax of Polish verbs from a cross-linguistic perspective.* Oxford: University of Oxford. (Doctoral dissertation).

Lasersohn, Peter. 1995. *Plurality, conjunction and events.* Dordrecht: Kluwer. DOI: 10.1007/978-94-015-8581-1.

Łaziński, Marek. 2020. *Wykłady o aspekcie polskiego czasownika.* Warszawa: Wydawnictwo Uniwersytetu Warszawskiego. DOI: 10 . 31338 / uw . 9788323542346.

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).

Petr Biskup

Link, Godehard. 1983. The logical analysis of plural and mass terms: A lattice theoretical approach. In Rainer Bäuerle, Christoph Schwarze & Arnim von Stechow (eds.), *Meaning, use, and interpretation of language*, 302–323. Berlin: de Gruyter. DOI: 10.1515/9783110852820.302.

Maienborn, Claudia. 2003. *Die logische Form von Kopula-Sätzen*. Berlin: Akademie-Verlag. DOI: 10.1524/9783050082271.

Maienborn, Claudia. 2005. On the limits of the Davidsonian approach: The case of copula sentences. *Theoretical Linguistics* 31(3). 275–316. DOI: 10.1515/thli.2005.31.3.275. https://doi.org/10.1515/thli.2005.31.3.275.

Marušič, Franc. 2023. Circumfixation. In Peter Ackema, Sabrina Bendjaballah, Eulàlia Bonet & Antonio Fabregas (eds.), *The Wiley Blackwell companion to morphology*. Hoboken, NJ: John Wiley & Sons. DOI: 10.1002/9781119693604.morphcom013.

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, Leipzig 2007*, 397–410. Frankfurt am Main: Peter Lang.

Mehlig, Hans Robert. 2006. Glagol'nyj vid i vtoričnaja gomogenizacija oboznačaemoj situacii posredstvom kvantifikacii: K upotrebleniju delimitativnogo sposoba dejstvija v russkom jazyke. In Volkmar Lehmann (ed.), *Semantika i struktura slavjanskogo vida IV* (Slavolinguistica 7), 235–276. München: Sagner.

Mehlig, Hans Robert. 2007. Aspect and bounded quantity complements in Russian. In Susan Rothstein (ed.), *Theoretical and Crosslinguistic Appoaches to the Semantics of Aspect*, 257–290. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/la.110.11meh.

Mehlig, Hans Robert. 2012. Hybrid predications in Russian. *Journal of Slavic Linguistics* 20(2). 171–227. DOI: 10.1353/jsl.2012.0010.

Nádeníček, Petr. 2011. *Das tschechische Aspektsystem im Vergleich mit dem Verbalaspekt des Russischen und Polnischen*. Hamburg: Verlag Dr. Kovač.

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

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.

Petr, Jan. 1986. *Mluvnice češtiny 1: Fonetika – fonologie – morfonologie a morfemika – tvoření slov*. Praha: Academia.

Petruxina, Elena V. 2011. *Aspektual'nye kategorii glagola v russkom jazyke v so-postavlenii s češskim, slovackim, pol'skim i bolgarskim jazykami.* Second edition. Moskva: URSS.

Piñón, Christopher. 1994. Aspectual composition and the 'pofective' in Polish. In Sergey Avrutin, Steven Franks & Liljana Progovac (eds.), *Formal approaches to Slavic linguistics: The MIT meeting 1993 (FASL 2)*, 341–373. Ann Arbor, MI: Michigan Slavic Publications.

Ramchand, Gillian. 2004. Time and the event: The semantics of Russian prefixes. *Nordlyd* 32(2). 23–361. DOI: 10.7557/12.72.

Romanova, Eugenia. 2004. Superlexical versus lexical prefixes. *Nordlyd* 32(2). 255–278. DOI: 10.7557/12.69.

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

Schuyt, Roel. 1990. *The morphology of Slavic verbal aspect: A descriptive and his-torical study.* Amsterdam & Atlanta: Rodopi. DOI: 10.1163/9789004654044.

Smith, Carlota S. 1997. *The parameter of aspect.* 2nd edition. Dordrecht: Springer. DOI: 10.1007/978-94-011-5606-6.

Součková, Kateřina. 2004a. *Measure prefixes in Czech: cumulative* na- *and delimi-tative* po-. Tromsø: University of Tromsø. (MA thesis).

Součková, Kateřina. 2004b. There is only one *po-. Nordlyd* 32(2). 43–419. DOI: 10.7557/12.74.

Švedova, Natalija J. 1980. *Russkaja grammatika.* Moskva: Nauka.

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

Tatevosov, Sergei. 2009. Množestvennaja prefiksacija i anatomija russkogo gla-gola. In Ksenia Kisileva, Vladimir Plungjan, Ekaterina Raxilina & Sergej [=Sergei] Tatevosov (eds.), *Korpusnye issledovanija po russkoj grammatike*, 92–156. Moskva: Probel-2000.

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

Tatevosov, Sergei. 2015. Severing imperfectivity from the verb. In Gerhild Zy-batow, 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.

Tatevosov, Sergei & Mikhail Ivanov. 2009. Event structure of non-culminating accomplishments. In Lotte Hogeweg, Helen de Hoop & Andrej Malchukov (eds.), *Cross-linguistic semantics of Tense, Aspect, and Modality*, 83–130. Amsterdam/Philadelphia: John Benjamins. DOI: 10.1075/la.148.05tat.

Timberlake, Alan. 2004. *A reference grammar of Russian*. Cambridge: Cambridge University Press. DOI: 10.1007/s11185-006-0705-7.

Vinogradov, Viktor V. 1947. *Russkij jazyk: Grammatičeskoe učenie o slove*. Moskva & Leningrad: Izdat. "Vysšaja Škola".

Wood, Esther J. 2007. *The semantic typology of pluractionality*. Berkeley, CA: University of California. (Doctoral dissertation).

Zaliznjak, Anna A. 2017. Russian prefixed verbs of motion revisited. *Zeitschrift für Slawistik* 62(1). 1–28. DOI: 10.1515/slaw-2017-0001.

Zaliznjak, Anna A. & Aleksej D. Šmelev. 1997. *Lekcii po russkoj aspektologii*. München: Verlag Otto Sagner. DOI: 10.3726/b12573.

Zinova, Yulia. 2021. *Russian verbal prefixation: A frame semantic analysis*. Berlin: Language Science Press. DOI: 10.5281/zenodo.4446717.

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.

Żmigrodzki, Piotr. 2022. Wielki słownik języka polskiego PAN jako słownik tematyczny. *Studia Językoznawcze: Synchroniczne i diachroniczne aspekty badań polszczyzny* 21. 153–166. DOI: 10.18276/sj.2022.21-11. https://wsjp.pl/.

Zucchi, Sandro. 1999. Incomplete events, intensionality and imperfective aspect. *Natural Language Semantics* 7. 179–215. DOI: 10.1023/A:1008310800632. https://doi.org/10.1023/A:1008310800632.

Chapter 4

Bare nouns in Slavic and beyond

Olga Borik[a], Bert Le Bruyn[b], Jianan Liu[b] & Daria Seres[c]

[a]Universidad Nacional de Educacion a Distancia [b]Utrecht University [c]Univeristy of Graz

The article presents a study of the distribution of singular bare nouns in three Slavic languages, Russian, Polish and Macedonian, based on parallel translation corpora. The distribution of bare singulars in Russian and Polish shows that they freely appear in definite and indefinite contexts, which makes it possible to classify these languages as truly articleless. Macedonian bare singulars frequently appear in indefinite contexts, alongside with nouns accompanied by an indefinite marker *one*, whose status require further scrutiny. The data reported in this study call for a theoretical account of bare nouns which allows for fine-grained variation in their distribution across domains and languages, taking into consideration a broader/narrower use of competing expressions.

1 Introduction

Referring is one of the main functions of natural language, and speakers of different languages use a variety of linguistic means and mechanisms to express different types of reference. In the empirical study that we present here, we focus on the syntax-semantics interface of bare nouns (BNs) and examine their distributional properties in Russian, Polish and Macedonian, languages that belong to East Slavic, West Slavic and South Slavic subgroups respectively. In particular, we address the issue of a comparative distribution of bare singular nouns (BSs) in the definite and indefinite domain across the three languages.

In terms of definiteness/indefiniteness marking, Russian and Polish are typically classified as articleless languages (Dryer & Haspelmath 2013), that is, having no dedicated morphosyntactic marker to express definiteness or indefiniteness.

Borik, Le Bruyn, Liu, & Seres

We thus expect nominals to appear in their bare form in all argument positions in both languages. This straightforward expectation is in line not only with the traditional descriptive grammars, such as Švedova (1980), but also with some formal semantic literature, such as Chierchia (1998), Geist (2010), among others. Other formal approaches, most notably Dayal (2004, 2018), argue that number plays a crucial role in the distribution of BNs in articleless languages, making different predictions for bare plurals (henceforth BPs) and BSs. In particular, Dayal (2004) argues that BSs do not get an indefinite interpretation in languages without articles, while BPs can get narrow scope indefinite readings. Therefore, BSs are predicted to be largely restricted to definite contexts. Our focus on BSs allows us to check the predictions made by Dayal's theory as opposed to more traditional approaches.

Macedonian is usually described as a language with a definite article (Friedman 1993, Tomić 2006, among many others). The definite article in Macedonian is postpositive and morphologically bound. It is typically added to the first element of a nominal phrase[1] (e.g., *kuče-to* 'the dog', *ubavo-to kuče* 'the beautiful dog'), and is inflected for number and gender (e.g., *maž-ot* M.SG 'the man', *žena-ta* F.SG 'the woman', *dete-to* N.SG 'the child', *maži-te/ženi-te* M/F.PL 'the men/the women', *deca-ta* N.PL 'the children'). BSs are also admissible in argument positions in Macedonian, while there is no agreement on their interpretation in the literature (Weiss 2004, Topolinjska 2009, i.a.). The most widely accepted assumption is that BSs in Macedonian appear in indefinite contexts, although it has also been noticed that a determiner *eden* 'one' is often used to mark indefiniteness in this language (Tomić 2006).

Our study aims at answering the following research questions:

1. What is the distribution of BSs in Russian and Polish as languages without articles? Do they appear in both definite and indefinite domains or do we observe significant differences in the distribution of BSs across domains?

2. What is the status of BSs in Macedonian in indefinite contexts as compared to the one in Russian and Polish?

3. What is the status of BSs in Macedonian in definite contexts?

To address these questions, we run a parallel corpus study to analyse nominal phrases that appear in both definite and indefinite contexts in the three languages, with a critical look at the distribution of BSs in each of the domains.

---

[1]We use the term *nominal phrase* to abstract away from the DP/NP debate, prominent mostly in the syntactic literature on Slavic. See, for instance, Bošković (2008).

For the definite domain, the expectations are rather straightforward: both traditional descriptive and formal literature seems to converge on the idea that BSs freely appear in definite contexts in Russian and Polish, whereas in Macedonian we expect the definite article to dominate. However, the status of the definite marker as an article in Macedonian is not uncontroversial: Rudin (2021: 313), for instance, suggests that it might be a type of demonstrative rather than an article. Semantic literature repeatedly stresses similarities between demonstrative NPs and definite descriptions (e.g., Roberts 2002, Elbourne 2008), as well as the need to differentiate between the two (Lyons 1999). We include demonstrative nominals in our empirical study and look at the relative distribution of NPs specified by demonstratives vs. definite nominal phrases in Macedonian or BSs in Russian and Polish in the definite domain.

For the indefinite domain, existing analyses diverge when it comes to predictions. Traditional descriptions do not report any irregularities or asymmetries in the distribution of BSs across definite vs. indefinite domains, so they seem to predict BSs to freely appear in indefinite contexts. However, claims have been made that in Polish and Macedonian, the indefinite marker ONE[2] is acquiring (or has acquired) the status of an indefinite article (Hwaszcz & Kędzierska 2018, Molinari 2022 with reference to Polish; Tomić 2006 with reference to Macedonian). The prediction that these proposals make is that the marker ONE will frequently appear in the indefinite domain in these languages, competing with or prevailing over BSs. The same prediction is made by Dayal (2004, 2018), who takes Hindi as a representative example of an articleless language and argues that it typically resorts to a construction with ONE in those contexts where English uses the indefinite article. Applying Dayal's analysis to Russian and Polish,[3] we expect BSs in these languages to be severely restricted in the indefinite domain, as opposed to the construction with ONE, which should dominate. In other analyses, the marker ONE in Russian is assumed to mark specificity rather than function as an indefinite article (Ionin 2013), which predicts its appearance only in specific indefinite contexts, converging with the predictions of Geist (2010), who argues that BSs in Russian can only get a non-specific reading.[4]

To get a broader cross-linguistic perspective, we compare parallel-corpus data for Russian, Polish and Macedonian to corpus results for Mandarin and German,

---

[2]The English ONE is used as a cover term for language specific *odin* (Russian), *jeden* (Polish) and *eden* (Macedonian) and their respective forms.

[3]Dayal does discuss Russian, and we assume that the proposal extends to other languages without articles like Polish, as it is based on general, language-independent semantic principles and mechanisms.

[4]Geist's (2010) predictions should be relativized to the information structure since she argues that indefinite BSs cannot serve as aboutness topics.

two non-Slavic languages. Mandarin functions as a control language for Russian and Polish, as it is usually assumed to be an articleless language (Li 2021), whereas German functions as a control language for Macedonian, as both have a definite article.

In order to investigate the distribution of BSs in definite and indefinite contexts in the three Slavic languages we run a parallel-corpus study, described in detail in §2. We present the results of our study in §3, followed by a general discussion in §4. §5 concludes the paper.

2 Data and methodology

We use parallel corpora to study the distribution of grammatical items in different languages in parallel, an approach that has recently gained traction in the formal literature for the study of a variety of empirical domains, for example, tense and aspect (see – among others – Fuchs & González 2022; Gehrke 2022; Mo 2022; Mulder et al. 2022), negation (de Swart 2020) and reference (Bremmers et al. 2021). Parallel corpus research builds on the assumption that the meanings of the original and the translations are as closely related to each other as the grammars of the respective languages allow them to be. Another important assumption is that translations are representative of their target languages (*the target language representativeness hypothesis*). For a more detailed discussion of the methodology and its caveats, see Le Bruyn et al. (2022), Le Bruyn & de Swart (2022).

This study uses a translation corpus built on the first chapter of J. K. Rowling's *Harry Potter and the Philosopher's Stone*, a novel written in English and translated into many typologically diverse languages. English grammatically marks the distinction between definiteness and indefiniteness, which allows us to easily detect all definite and indefinite referential expressions in the source text. We selected all (in)definite referential expressions (*a N, the N, N-s, the N-s*) with their aligned translations in Russian, Polish, Macedonian, Mandarin and German (n=284) and manually annotated the corresponding NP forms in all the target languages.[5] At this point, it is important to highlight that our methodology involves the annotation of forms (but not meanings) in the same contexts across the languages under study.

---

[5]Because some referential expressions are not translated and because of issues of automatic alignment, some data are literally lost in translation. Our dataset for this study includes referential expressions that have translations in all five languages under scrutiny. These numbers are expected not to be identical to the ones in Liu et al. (2023), a study that we conducted for a wider set of languages using the same methodology.

Since this paper focuses on the singular domain, we limit our quantitative analysis to the singular paradigm only.[6] Apart from theoretical reasons discussed in §1, plurals were excluded due to the relatively low frequency in our dataset and the interaction of plural definites with proper names (e.g., *The Potters, The Dursleys*). Thus, our final dataset includes the translations of *a N* (n=82) and *the N*$_{sing}$ (n=124) constructions into Russian, Polish, Macedonian, as well as Mandarin and German, which are used as control languages in this study.

An example of an English source *the N* expression (1a) and its translations from the parallel corpus is shown below.

(1)   a. Mr Dursley might have been drifting into an uneasy sleep, but *the cat* on the wall outside was showing no sign of sleepiness.

b. Dolgoždannyj i nespokojnyj son uže prinjal v svoi ob"jatija mistera Darsli, a sidevšaja na ego zabore *koška* spat' soveršenno ne sobiralas'.

Russian [N]

c. Pan Dursley zapadł w niezbyt zresztą spokojny sen, ale *kot* na murku nie okazywał najmniejszych oznak senności.   Polish [N]

d. Gospodinot Darsli možebi potona vo nemiren son, no *mačkata* na dzidot nadvor ne pokažuvaše ni troška sonlivost.

Macedonian [N+the]

e. Mr Dursley mochte in einen unruhigen Schlaf hinübergeglitten sein, doch *die Katze* draußen auf der Mauer zeigte keine Spur von Müdigkeit.   German [the N]

f. Désīlǐ xiānshēng mímíhúhú, běnlái kěnéng húluàn shuì-shàng yí jiào, kě huāyuán qiángtóu shàng *nà zhī māo* què méiyǒu sīháo shuìyì.

Mandarin [demonstrative+classifier+N]

In the definite domain, we examine the forms that Russian, Polish and Macedonian use for the translation of the English *the N*. In particular, we check whether and to what extent BSs that we expect to find in Russian and Polish, and singular definites that we expect to find in Macedonian, interact with demonstratives in the definite domain. We then contrast the results obtained for the three Slavic languages with Mandarin (as a control for Russian and Polish) and German (as a control for Macedonian).

In the indefinite domain, we look to determine which forms are used for the translations of the English *a N* in all three languages. We evaluate to which extent BSs are used in singular indefinite contexts in Russian vs. Polish vs. Macedonian and check for the interactions with the forms using the marker ᴏɴᴇ. Once

---

[6]Although there will be a short discussion of plurals in §4.1.2.

again, we compare the results obtained for Slavic languages with the results for Mandarin in the indefinite domain.

3 Results

3.1 Singular definite contexts

As far as definite contexts are concerned, there are no major surprises found in our data. The overall results are presented in Figure 1, which reflects absolute frequencies and includes all translations in the target languages. The category *Rest* contains all those translations that do not present any immediate interest for us (e.g., pronouns, possessives, etc.).

As we can observe, BSs are, indeed, the default option for rendering English *the N* both in Russian and in Polish, as shown in Figure 1. The differences in the occurrence of bare nominals in definite contexts are not significant for these two languages (p = 0.37, Fisher's Exact Test (FET)). Regarding Macedonian, the most prominent form in singular definite contexts is the one with a definite article, a result which is also fully in accordance with our initial expectations. In all three languages, there are practically no demonstratives used in singular definite contexts.

Comparing the results of Russian and Polish with their control language, Mandarin, we see that BNs in Mandarin are the most frequent form in the definite context as well. However, we also observe an important difference in the relative distribution of BNs vs. NPs specified by demonstratives in the definite domain: in Mandarin, the tendency to resort to demonstratives is higher. The differences are significant for the comparison of Mandarin and Polish (p < 0.001, FET), and for Mandarin and Russian (p = 0.016, FET).

As for Macedonian and its control language German, the two languages are quite uniform in the distribution of nominal forms in the definite domain. BSs and demonstrative NPs are either absent or clearly outnumbered in singular definite contexts in both Macedonian and German.[7] In §4.1.1, we come back to the issue of definiteness marking in Macedonian and discuss some of the examples with BSs.

---

[7]More specifically, there is only one BS found in German and three in Macedonian.
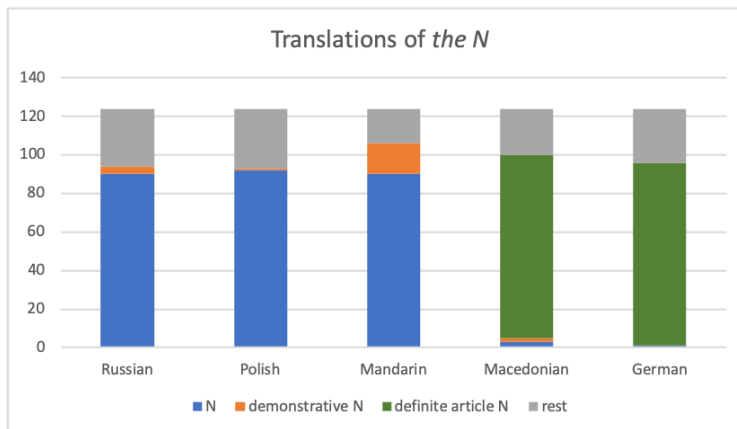
Figure 1: Russian, Polish and Mandarin BNs vs. demonstrative-N; Macedonian and German the-N vs. demonstrative-N

3.2 Singular indefinite contexts

The parallel-corpus data showed that a bare noun is the default option for rendering singular indefinite nominals in both Russian and Polish (see Figure 2). These two languages do not use the ONE+N construction in indefinite contexts in a statistically relevant way. The differences in distributions of bare nominals and nominals preceded by ONE are not significant for Russian and Polish ($p = 0.5$ FET).

In Macedonian, however, while a BS is still the most frequent form in the indefinite domain, the English *a N* construction is more often translated with the numeral ONE than in Russian or Polish. The differences are significant for the comparison of both Macedonian and Russian ($p < 0.001$, FET), and Macedonian and Polish ($p < 0.001$, FET).

As for the control language, Mandarin, where the numeral ONE precedes the nominal in a large number of cases in indefinite contexts, it shows a sharp contrast with Russian and Polish, which hardly ever use this structure. Moreover, Mandarin also shows contrast with Macedonian, where the use of ONE is not as frequent. The differences are significant for Mandarin and Russian ($p < 0.001$, FET), and Mandarin and Polish ($p < 0.001$, FET), as well as for Macedonian and Mandarin ($p < 0.001$, FET).
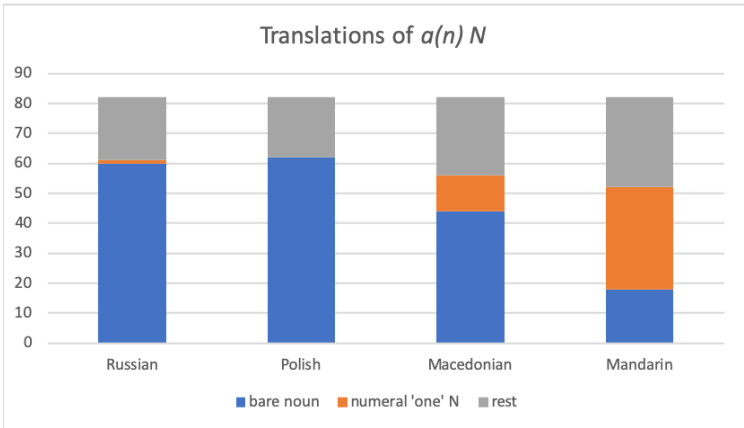
Borik, Le Bruyn, Liu, & Seres

Figure 2: Russian, Polish, Macedonian and Mandarin bare nominals vs. ONE+N

3.3 Recap

Summing up the results of our parallel corpus study, it can be said that Russian and Polish freely use bare nouns in both singular indefinite and singular definite contexts, in accordance with the Slavic descriptive literature. They are, however, in sharp contrast with Mandarin, where the numeral ONE seems to be the default option in the indefinite domain and the demonstrative is competing with bare NPs in the definite domain.

As for Macedonian, in the indefinite domain it seems to occupy an intermediate position between Russian and Polish, on the one hand, and Mandarin, on the other: ONE+N construction appears in the translations of *a N* quite frequently, but not as often as in Mandarin. In the definite domain, Macedonian uses NPs with a definite marker in *the N_{sing}* contexts as often as German.

4 Discussion

We structure our main discussion points in the same way we presented the results of the study, that is, according to the distribution of various forms in a specific context. We begin by evaluating the parallel corpus results obtained for the definite domain. In the discussion of the indefinite domain, we reflect not only on the distribution of BSs, but also on the role of ONE+N construction in the grammar of all three target languages.

4.1 Definite contexts in Russian, Polish and Macedonian

Both traditional descriptions and formal semantic analyses seem to be in full agreement on attributing a possible definite reading to BSs in languages without articles. Our data cast no doubt on this claim for Russian or Polish: BSs prevail in definite contexts in both languages and the distributional behaviour of BSs is therefore in full accordance with their standard semantic descriptions and/or analyses. In Macedonian, the prevailing form is the definite singular, i.e., our data also confirm the status of Macedonian as a language with a definite article. Even though the interpretation of our main results seems to be rather straightforward, there are two points of interest that we would like to discuss.

The first observation concerns the distribution of BSs vs. NPs specified by demonstratives in the definite domain. In the previous section, we pointed out that demonstratives do not seem to occupy a prominent place in either of the three Slavic languages in *the N_{sing}* contexts. Russian and Polish as languages without articles can be contrasted to Mandarin in this respect, one of the two control languages used in this study, where the higher rate of demonstratives in the definite domain suggests that the demonstrative plays a much more significant role in definiteness marking in Mandarin. In fact, Liu et al. (2023) hypothesise that Mandarin is developing a definite article (and an indefinite one), but in this paper, we limit ourselves to empirical statements with respect to Mandarin. In Macedonian, there were only three contexts where a demonstrative was used, all corresponding to anaphoric uses of *the N* in the source text.

The second point that we would like to discuss is the status of the definite article in Macedonian, which will be examined in the next subsections.

4.1.1 The Macedonian definite article in the singular domain

The status of the Macedonian definite article has been subject to some debate in the semantic literature, as pointed out in §1. No consensus emerges from the literature concerning the semantic contribution of this marker. One of the features that it exhibits (and that distinguishes it from a typical definite article) is that it morphologically marks a proximal-neutral-distal distinction, just like demonstratives in many languages do (Lyons 1999). It should be noted that the question about the status of a definite marker in any language is essentially semantic and cannot be definitively resolved without looking into the meaning of this expression, but the distribution of any definite marker/article also plays a significant role in a potential analysis and this is what our study can inform about.

We looked into the properties of the definite article in Macedonian by running a comparative study of the distribution of the definite article in Macedonian and

Borik, Le Bruyn, Liu, & Seres

German. In particular, we measured their co-occurrence in the same contexts by calculating Normalized Pointwise Mutual Information (NMPI, Bouma 2009), which is a bidirectional measure for parallel data (Le Bruyn et al. 2022). The result shows that the NPMI of the two articles reaches 0.48 (with a maximum of 1). That means that the likelihood of the articles in the two languages occurring in the same contexts is higher than chance but not at ceiling. In other words, the bi-directional mapping pattern suggests that the distribution of definite articles in German and Macedonian across the definite contexts is not completely identical.

Table 1: Bi-directional mapping patterns between the German and Macedonian (singular) definite article

		Macedonian		
		definite article N	rest	
German	definite article N	88	24	112
	rest	20	74	94
		108	98	206

Looking into the contexts where Macedonian and German did not coincide in the use of the definite article, we find some interesting examples of BSs. For instance, in (2), Macedonian uses a BS, while German opts for a definite article, just like the English source:

(2) a. At half past eight, Mr Dursley picked up his briefcase, pecked Mrs Dursley on *the cheek* and tried to kiss Dudley goodbye but missed...
English (source): [the N]

b. Vo osum i pol gospodinot Darsli ja zede svojata aktenčanta, ja kolvna gospoǵáta Darsli vo *obraz* i se obide da go bakne Dadli za razdelba, no ne uspea... Macedonian: [N]

c. Um halb neun griff Mr Dursley nach der Aktentasche, gab seiner Frau einen Schmatz auf *die Wange* und versuchte es auch bei Dudley mit einem Abschiedskuss... German: [the N]

Another example of the same type of article mismatch is presented in (3):

(3) a. A man appeared on the corner the cat had been watching, appeared so suddenly and silently you'd have thought he'd just popped out of *the ground.* English: [the N]

b. Na agolot što go nabljuduvaše mačkata se pojavi čovek, tolku nenadejno i tivko, kako da izniknal od *zemja.* Macedonian [N]

c. An der Ecke, die sie beobachtet hatte, erschien ein Mann, so jäh und
lautlos, als wäre er geradewegs aus *dem Boden* gewachsen.

German: [the N]

Although we cannot reach any firm conclusions on the basis of only few exam-
ples, we can hypothesise that they both present cases of weak definites (posses-
sive weak definites, Barker 2005, in the case of (2)), so that the contexts where
the uniqueness of a definite description is questioned are potentially very good
candidates for the absence of the definite article in Macedonian. Needless to say,
additional empirical investigation is needed to check this hypothesis.

4.1.2 Some remarks on the plural definite domain

Even though we did not run statistical analyses for the plural domain and there
were not too many data points in our dataset, we would like to draw the attention
of the reader to some observations concerning the use of the definite article with
plurals in Macedonian that appear important. For instance, Macedonian seems
to use definite articles in plural generic contexts, while English resorts to bare
plurals and German presents variation.

(4) a. *Cats* couldn't read maps or signs. English (source): [Ns]

b. *Mačkite* ne možat da čitaat ni mapi ni oznaki. Macedonian: [Ns+the]

c. *Katzen* konnten weder Karten noch Schilder lesen. German: [Ns]

The use of the definite article with generic plurals as illustrated in (4) may suggest
that Macedonian – at least in some aspects – is rather comparable to Romance
languages in its use of definite plurals than to Germanic languages.

Existential contexts in the plural definite domain require further scrutiny. We
detected several examples in our dataset where both Macedonian and German
use a definite article whereas English uses a bare plural in the same context. One
of those examples is (5).

(5) a. And finally, *bird-watchers* everywhere have reported that the nation's
owls have been behaving very unusually today. English:[Ns]

b. I konečno *nabljuduvačite* na ptici od site strani javija deka buvo - vite
vo našata zemja deneska se ondesuvale mnogu neobično.

Macedonian: [Ns + the]

c. Und hier noch eine Meldung. Wie *die Vogelkundler* im ganzen Land
berichten, haben sich unsere Eulen heute sehr ungewöhnlich
verhalten. German: [the Ns]

Borik, Le Bruyn, Liu, & Seres

In this particular case, the presence of the article in Macedonian could be due to a specific syntactic construction used in the example (prepositional phrase *watchers of birds* instead of the nominal compound *bird-watchers* in the source text). This, however, would not explain the presence of the definite article in German. It might also be the case of a so-called FUNCTIONAL reading of BPs in English, discussed at length for English by Condoravdi (1994). The availability of this reading for BPs is language-specific, so we conclude that our data demonstrate some cross-linguistic variation worth a more systematic investigation. It is not surprising to see this variation in the distributional patterns, as cross-linguistic differences in the use of the definite article are very well documented and widely discussed in the literature. The corpus data of the current study is not sufficient to arrive at any firm conclusions, but it may be reasonably suggested that German and Macedonian, just like many other languages with grammatical marking of definiteness, do not fully coincide in definiteness marking patterns: the overlap in the use of the definite article is only partial, not absolute.

4.2 Indefinite contexts in Russian, Polish and Macedonian

Indefinite contexts constitute the most interesting case in our study, as they convincingly illustrate several theoretically relevant points. First, there is variation both within and outside the Slavic family in the distribution of BSs in the indefinite domain, which has direct repercussions for existing theoretical analyses of BSs. Second, intricate interactions of ONE+N with BSs in the indefinite domain can elucidate the grammatical status of ONE in a given language. Third, our data pose some very specific constraints and requirements for an accurate and empirically adequate theoretical analysis of BSs in languages without articles. We discuss each of these points in the three subsections that follow.

4.2.1 BSs in Russian, Polish and Macedonian

One of the main results of our study concerns the distributional pattern of BSs in Russian and Polish. In particular, the data from the parallel corpus show that in Russian and Polish BSs freely appear in indefinite singular contexts as counterparts of *a N* in the source text. One rather typical example of an indefinite in an existential context is given below:

(6) a. English (source): [a N]
         There was *a tabby cat* standing on the corner of Privet Drive, but
         there wasn't a map in sight.

   b. Russian: [N]
     Na uglu Praivet Draiv dejstvitel'no sidela *polosataja koška*, no nikakoj karty vidno ne bylo.

   c. Polish: [N]
     Na rogu Privet Drive rzeczywiście stał *bury kot*, ale nie studiował żadnej mapy.

Our Russian and Polish data directly support traditional descriptive approaches to BSs in Slavic languages without articles and those formal approaches which do not rule out an indefinite interpretation for BSs, e.g., Chierchia (1998), Krifka (2003). The results of our study are also compatible with the proposal that bare NPs in Russian are essentially indefinite and a definite reading is achieved through pragmatic strengthening (Seres & Borik 2021).

    On the other hand, our empirical findings are in conflict with Dayal's (2004) proposal, whose prediction – as we mentioned in §1 – is that BSs should never give rise to indefinite readings in regular argument position in languages without articles. Dayal examines the behavior of BSs in Hindi, Russian and Mandarin, and argues that an overt indefiniteness marker has to appear in those contexts where an indefinite reading has to be expressed. This prediction holds for Hindi, where ONE functions as such a marker,[8] but it is very clear that Russian and Polish behave differently. In fact, ONE has only been used twice in Russian in the indefinite domain, whereas Polish data do not contain a single occurrence of this item. Thus, our data allow us to conclude that both Russian and Polish are truly articleless languages where BSs dominate in both definite and indefinite contexts. No competing forms emerge in our study in either of the two contexts in either of the two languages.

    In contrast to Russian and Polish, Macedonian uses both BSs and ONE+N constructions. Our data show that Macedonian differs from truly articleless languages, and the construction ONE+N competes with BSs in the indefinite domain in Macedonian. This difference can be illustrated with the translation of example (6a) from above into Macedonian: where Russian and Polish use a BS, Macedonian uses ONE+N.

(7)   Na agolot na Šimširovata uliča stoeše *edna neobična šarena mačka*, no nikade nemaše mapa.

If we look outside the Slavic family, our control language, Mandarin, shows a strong tendency for the ONE+N construction to appear in singular indefinite

---

   [8]This result has been confirmed by a parallel corpus study reported in Liu et al. (2023).

contexts (see Figure 2). Macedonian clearly occupies an intermediate position between Mandarin (relatively low percentage of BSs) and Russian/Polish (predominantly BSs) with respect to the use of BSs in the singular indefinite domain.

Note that this kind of variation in the use of BSs comes out unexpected on most analyses. In general, articleless languages are perceived as a homogeneous group that either do or do not use BSs in a certain domain, but the kind of variation that we see in our data is rather challenging for theoretical approaches. We will come back to this point at the end of this section, but first we get a better look at the closest competitor of a BS in the indefinite domain, the indefinite marker ONE.

4.2.2 The status of ONE in the indefinite domain

It is well known that the numeral ONE is a predecessor of the indefinite article in many languages (Heine 1997, van Gelderen 2011, among many others). Looking once again at the distribution of nominal forms in the indefinite domain in Figure 2, we observe a clear interaction between BSs and ONE+N constructions: the frequency of ONE+N in our data goes from being at floor in Russian and Polish to a significant percentage in Macedonian and to predominance in Mandarin. This raises a question about the grammatical status of the marker ONE in different languages.

The differences in the use of the ONE+N construction across languages may be accounted for by different stages of its grammaticalisation as an article. Typically, the stages of grammaticalisation of the indefinite article are defined in the following order: 1. the numeral, 2. the presentative marker, 3. the specificity marker, 4. the non-specific marker, 5. the generalised article (Givón 1981, Heine 1997, i.a.).[9] Even though defining the exact stage of grammaticalisation of ONE in the languages under study is out of the scope of this paper, our data offer several discussion points relevant for the issue.

Our empirical findings for Russian and Polish, where BSs overwhelmingly dominate in the indefinite domain, seem to be in conflict with the proposal of Hwaszcz & Kędzierska (2018), who claim that in Russian ONE is grammaticalised as a presentative marker, that is, it marks a newly introduced referent, which is intended to be used in the subsequent discourse and is usually specific and topical. The authors also claim that in Polish ONE is grammaticalised even further, being used as a specific and sometimes as a non-specific marker. Neither of the two claims is confirmed by our data, as some representative examples can illustrate:

---

[9]These stages are coarsely defined and may have substages.

(8) a. English (source): [a N]
The Dursleys had *a small son* called Dudley and in their opinion there was no finer boy anywhere.

b. Russian: [N]
U mistera i missis Darsli byl *malen'kij syn* po imeni Dadli, i, po ix mneniju, ėto byl samyj čudesnyj rebenok na svete.

c. Polish: [N]
Syn Dursleyów miał na imię Dudley, a rodzice uważali go za najwspanialszego chłopca na świecie.

(9) a. English (source): [a N]
He was sure there were lots of people called Potter who had *a son* called Harry.

b. Russian: [N]
Mister Darsli legko ubedil sebja v tom, čto v Anglii živet množestvo semej, nosjaščix familiju Potter i imejuščix *syna* po imeni Garri.

c. Polish: [N]
Mnóstwo ludzi może się nazywać Potter i mieć *syna* Harry'ego.

Example (8) is a typical context where a new specific referent is introduced by a modified indefinite in the source text, which is then rendered by a BS both in Russian and in Polish, just as a non-specific indefinite *a son* in (9). At least in Russian, ONE+N cannot be used instead of N in (8) and (9), unless ONE is interpreted as a numeral.[10] Our data show no sign of any significant difference between Russian and Polish with respect to the grammatical status of ONE: this marker does not show up regularly or systematically in either presentative or specific or any other type of context.

Macedonian ONE, on the other hand, is more frequent. We have not conducted any specific study of the contexts where ONE appears in Macedonian, as our dataset is too small to yield sensible results, but we can provide some indicative examples here that can help us map out a path for future research. For instance, Macedonian uses a BS in translations of both example (8) and (9) above, but there are other specific and non-specific contexts where ONE+N construction appears:

(10) a. English (source): [a N]   (non-specific)
My dear Professor, surely *a sensible person* like yourself can call him by his name?

---

[10]We thank an anonymous reviewer for stressing this point.

b. Macedonian [one N]
Draga moja profesorke, ne misliš li deka *edna tolku razumna ličnost'*
kako što si ti slobodno može da go narekuva po ime?

(11) a. English (source): [a N]   (specific)
Professor McGonagall pulled out *a lace handkerchief* and dabbed at
her eyes beneath her spectacles.

b. Macedonian [one N]
Profesorkata Mekgonagl izvadi *edno tanteleno maramče* i gi protri
očite pod očilata.

The mixed data across specific and non-specific contexts indicate that the ONE+N
construction is not really established in these types of contexts. The data ob-
tained in our study are, in principle, in line with Hwaszcz & Kędzierska (2018),
who claim that ONE in Macedonian is used with both specific and non-specific
indefinite NPs. Our Macedonian data show that both specific and non-specific in-
definite NPs may also appear as bare, as illustrated in the above examples, which
may indicate a certain degree of optionality in the use of ONE for marking specific
and non-specific nominals.[11] This flexibility (possibly translated as optionality)
provides a contrast with English and German, languages where an indefinite ar-
ticle is obligatorily used in all the examples discussed in this subsection. Thus,
Macedonian does differ from languages with established indefinite articles, and
we therefore conclude this discussion by saying that the status of ONE cannot be
unequivocally defined as an indefinite article in Macedonian, contra, e.g., Tomić
(2006).[12] Rather, ONE is an indefinite marker that might evolve into an article,
but further research is needed to substantiate this claim.

4.2.3 Theoretical implications

As the discussion in the previous sections indicates, the main challenge that our
data pose for theoretical approaches striving for empirical adequacy is the prob-
lem of language variation. The variation in the definite domain, especially in the
distribution of the definite article across languages, is relatively well known and
discussed in the semantic literature (e.g., Dryer 2005). Our analysis of the definite
article in Macedonian vs. German adds one more study case to this discussion.

---

[11]One of the limitations of corpus studies is that it is impossible to determine the optionality of
an element. In order to research the (non-)obligatoriness of ONE in certain linguistic environ-
ments, linguistic experiments with native speakers need to be carried out.

[12]In this respect, the Macedonian data resemble the situation in Bulgarian, as reported in Geist
(2013).

In the indefinite domain, however, variation in the distribution of BSs in languages without articles (or without an indefinite article) is less expected. For instance, the approach to BNs in general and BSs in particular developed in Dayal (2004), Dayal (2018), and Dayal & Sağ (2020) is based on the claim that BSs do not allow for indefinite readings in articleless languages. The formal machinery of this approach does not leave much room for variation: the denotation of a noun in regular argument positions is derived by type-shifting operators and, crucially, Dayal's analysis cuts off the possibility of an existential type-shift for BSs. The logic behind this move, we believe, applies universally. Our data for Russian and Polish, though, strongly suggest that there should be an easy way to allow for a BS to appear in the singular indefinite domain, which may be achieved via standard type-shifting operations, like an existential type shift. However, allowing for this type shift to be subject to parametric variation will considerably weaken Dayal's formal theory, at least in the absence of any independent principle underlying such variation.

The Macedonian data, where we see a competition between BSs and the ONE+N construction, suggest that there should be a way to allow for BSs in those contexts where the other construction does not appear on a regular basis. In other words, there should be an account of an interaction between nominal forms that coexist in the indefinite domain. Dayal's approach cannot easily accommodate such interaction either, because ONE+N is predicted to be the only option in the indefinite domain in the absence of an indefinite article. Thus, we conclude that the semantic theory of bare nominals advocated in Dayal (2004), Dayal (2018), and Dayal & Sağ (2020) has considerable difficulties accounting for an overall empirical picture that emerges from our data.[13]

Mandarin, our control language, clearly prefers the ONE+N construction to BNs in the indefinite domain. As Liu et al. (2023) argue, this fact does not really follow from Dayal's analysis either, since in Mandarin, which lacks grammatical number, BNs are expected to easily get an indefinite reading, just like BPs in other languages do. If Mandarin BNs behave like BPs rather than BSs, they are predicted to get a narrow scope indefinite reading and hence, they should be visibly prominent in indefinite (singular and plural) contexts. In our data, however, the ONE+N construction wins over BNs in the singular indefinite domain. In fact, it looks like what Dayal (2004, 2018) predict for Mandarin occurs in Russian and Polish, with a proviso for number marking, and what her analysis predicts for Russian and Polish seems to hold for Mandarin.

An analysis that our data calls for should allow for a formal way to derive an

---

[13]See also Liu et al. (2023).

existential interpretation of a BS via type-shifting, but only if there is no competing form with an overt marker that would block this shift. Chierchia's (1998) or Krifka's (2003) classical analyses, for instance, state that while in some languages type shifts are indicated by overt determiners, in languages that lack them, type shifts apply covertly whenever the linguistic context requires it. Covert type-shifting is restricted by the Blocking Principle, which roughly states that if a language has an overt means to express a type shift, then it must be used. This analysis seems to be much better equipped to handle our data. For instance, we have seen no evidence that ONE+N in Russian (*odin N*) and Polish (*jeden N*) functions as an article-like expression. Thus, the covert application of the existential type-shift is not blocked, which allows for BSs to be freely used in indefinite contexts. For Macedonian, a language with an emerging indefinite marker ONE, the existential type shift would be blocked for a BN only in those contexts where *eden* appears. Our cross-linguistic data provide a serious argument in favor of a classical blocking semantic analysis of bare nominals, in which fine-grained variation in the distribution of bare nominals follows from the broader/narrower use of article-like expressions.

5 Conclusions

In this paper, we have reported the results of a parallel translation corpus study on the distribution of BSs in three Slavic languages, Russian, Polish and Macedonian. We built our corpus on the text of the first chapter of *Harry Potter and the Philosopher's Stone* and complemented the results obtained for Slavic languages with the results for Mandarin as a control language for Russian and Polish, and German as a control language for Macedonian.

In the view of the empirical data presented here, it can be concluded that Russian and Polish are truly articleless languages and freely allow their BSs to take on definite and indefinite readings across domains. In Macedonian, BSs are restricted to the indefinite domain where they compete with the indefinite marker ONE, whereas in the definite domain, Macedonian uses the definite article, just as expected. Therefore, we conclude that Macedonian is a language with a definite article and with an emerging indefinite marker whose exact grammatical status requires further empirical investigation.

Slavic languages present challenging theoretically relevant contrasts with their control languages. In case of Macedonian, we have stressed the need to further scrutinize the conditions and the contexts where the definite article is used because we have shown the overlap between the definite articles in Macedonian

and German is partial. We also see the need to extend the investigation to the plural domain to get a full picture of the distribution of the definite article in Macedonian. As for Russian and Polish, they present a striking contrast with Mandarin in the indefinite singular domain, where the two Slavic languages show a clear preference for BSs and Mandarin opts for ONE+N construction as a counterpart of the English *a N*. Macedonian occupies an intermediate position: ONE+N is used rather frequently in Macedonian, but not as often as in Mandarin singular indefinite contexts.

We have argued that these contrasts call for a theoretical approach where the observed variation in the distribution of BSs and competing forms can be naturally accounted for. We suggest that the Blocking Principle as formulated in Chierchia (1998) can serve as a foundation for such an approach.

Abbreviations

F	feminine	PL	plural
M	masculine	SG	singular
N	neuter		

Acknowledgments

References

Barker, Christian. 2005. Possessive weak definites. English (US). In Kim, Ji-yung, Lander, Yury, Partee & H. Barbara (eds.), *Possessives and Beyond: Semantics and Syntax.* 89–113. GLSA Publications.

Bošković, Zeljko. 2008. What will you have, DP or NP? In Emily Elfner & Martin Walkow (eds.), *Proceedings of NELS 37*, 101–114. University of Illinois Urbana-Champaign.

Borik, Le Bruyn, Liu, & Seres

Bouma, Gerlof J. 2009. Normalized (pointwise) mutual information in collocation extraction. In Christian Chiarcos, Richard Eckart de Castilho & Manfred Stede (eds.), *Proceedings of the Biennial International Conference of the German Society for CLLT*, 31–40. Tübingen: Gunter Narr Verlag.

Bremmers, David, Jianan Liu, Martijn van der Klis & Bert Le Bruyn. 2021. Translation mining: Definiteness across languages (A reply to Jenks 2018). *Linguistic Inquiry* 53(4). 735–752. DOI: 10.1162/ling_a_00423.

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

Condoravdi, Cleo. 1994. *Descriptions in context*. New Haven, CT: Yale University. (Doctoral dissertation).

Dayal, Veneeta. 2004. Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* 27(4). 393–450. DOI: 10.1023/b:ling.0000024420.80324.67.

Dayal, Veneeta. 2018. (In)definiteness without articles: Diagnosis, analysis, implications. In Ghanshyam Sharma & Rajesh Bhatt (eds.), *Trends in Hindi linguistics*, 1–26. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110610796-001.

Dayal, Veneeta & Yağmur Sağ. 2020. Determiners and bare nouns. *Annual Review of Linguistics* 6(1). 173–194. DOI: 10.1146/annurev-linguistics-011718-011958.

de Swart, Henriëtte. 2020. Double negation readings. In Viviane Déprez & M. Teresa Espinal (eds.), *The Oxford Handbook of Negation*, 479–496. Oxford: Oxford University Press. DOI: https://doi.org/10.1093/oxfordhb/9780198830528.013.26.

Dryer, Matthew. 2005. Definite articles. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 154–157. Oxford: Oxford University Press.

Dryer, Matthew S. & Martin Haspelmath (eds.). 2013. *WALS online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. https://wals.info/.

Elbourne, Paul D. 2008. Demonstratives as individual concepts. *Linguistics and Philosophy* 31(4). 409–466. DOI: 10.1007/s10988-008-9043-0.

Friedman, Victor A. 1993. Macedonian. In Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic Languages*, 249–305. London / New York: Routledge.

Fuchs, Martín & Paz González. 2022. Perfect-perfective variation across Spanish dialects: A parallel-corpus study. *Languages* 7(3). 166. DOI: 10.3390/languages7030166.

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.

Geist, Ljudmila. 2010. Bare singular NPs in argument positions: Restrictions on indefiniteness. *International Review of Pragmatics* 2(2). 191–227. DOI: 10.1163/187731010x528340.

Geist, Ljudmila. 2013. Bulgarian *edin*: The rise of an indefinite article. In Uwe Junghanns, Dorothee Fehrmann, Denisa Lenertová & Hagen Pitsch (eds.), *Formal Description of Slavic Languages: The Ninth Conference. Proceedings of FDSL 9, Göttingen 2011*, 125–148. Frankfurt am Main: Peter Lang.

Givón, Talmy. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* 15(2-1). 35–54. DOI: doi:10.1515/flih.1981.2.1.35.

Heine, Bernd. 1997. *Cognitive Foundations of Grammar*. New York/Oxford: Oxford University Press. DOI: 10.1093/oso/9780195102512.001.0001.

Hwaszcz, Krzysztof & Hanna Kędzierska. 2018. The rise of an indefinite article in Polish: An appraisal of its grammaticalisation stage (Part 1). *Studies in Polish Linguistics* 13. 93–121. DOI: 10.4467/23005920SPL.18.005.8744.

Ionin, Tania. 2013. Pragmatic variation among specificity markers. In Cornelia Ebert & Stefan Hinterwimmer (eds.), *Different kinds of specificity across languages*, 75–103. Dordrecht: Springer. DOI: 10.1007/978-94-007-5310-5_4.

Krifka, Manfred. 2003. Bare NPs: Kind-referring, indefinites, both, or neither? In Robert B. Young & Yuping Zhou (eds.), *Proceedings of SALT 13*, 180–203. DOI: https://doi.org/10.3765/salt.v13i0.2880.

Le Bruyn, Bert & Henriëtte de Swart. 2022. *Cross-linguistic semantics: Methodological advances*. Course materials, ESSLLI. shorturl.at/doqV6.

Le Bruyn, Bert, Martín Fuchs, Martijn van der Klis, Jianan Liu, Chou Mo, Jos Tellings & Henriëtte de Swart. 2022. Parallel corpus research and target language representativeness: The contrastive, typological, and translation mining traditions. *Languages* (3). 176. DOI: 10.3390/languages7030176.

Li, Xuping. 2021. The Semantics of Chinese Noun Phrases. In *Oxford Research Encyclopedias: Linguistics*. Oxford: Oxford University Press. https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-886.

Liu, Jianan, Shravani Patil, Daria Seres, Olga Borik & Bert Le Bruyn. 2023. 'Articleless' languages are not created equal. In Maria Onoeva, Anna Staňková & Radek Šimík (eds.), *Proceedings of Sinn und Bedeutung 27*, 381–398. DOI: 10.18148/sub/2023.v27.1076.

Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511605789.

Mo, Chou. 2022. *The compositionality of Mandarin Aspect: a parallel corpus study*. Utrecht: University Utrecht. (Doctoral dissertation). https://www.lotpublications.nl/Documents/628_fulltext.pdf.

Molinari, Luca. 2022. The syntax of Polish *jeden* 'one' as an indefinite determiner. *Annali di Ca' Foscari. Serie occidentale* 56. 63–84. DOI: http://doi.org/10.30687/AnnOc/2499-1562/2022/10/004.

Mulder, Gijs, Gert-Jan Schoenmakers, Olaf Hoenselaar & Helen de Hoop. 2022. Tense and aspect in a Spanish literary work and its translations. *Languages* 7(3). 217. DOI: https://doi.org/10.3390/languages7030217.

Roberts, Craige. 2002. Demonstratives as definites. In Kees van Deemter & Roger Kibble (eds.), *Information sharing: Reference and presupposition in language generation and interpretation*, 89–136. Standford: CSLI Press.

Rudin, Catherine. 2021. Demonstratives and definiteness: Multiple determination in Balkan Slavic. In Andreas Blümel, Jovana Gajić, Ljudmila Geist, Uwe Junghanns & Hagen Pitsch (eds.), *Advances in formal Slavic linguistics 2018* (Open Slavic Linguistics 4), 305–338. Berlin: Language Science Press. DOI: 10.5281/zenodo.5483114.

Seres, Daria & Olga Borik. 2021. Definiteness in the absence of uniqueness: the case of Russian. In Andreas Blümel, Jovana Gajić, Ljudmila Geist, Uwe Junghanns & Hagen Pitsch (eds.), *Advances in formal Slavic linguistics 2018* (Open Slavic Linguistics 4), 339–363. Berlin: Language Science Press. DOI: 10.5281/zenodo.5155544.

Švedova, Natalija J. 1980. *Russkaja grammatika*. Moskva: Nauka.

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

Topolinjska, Zuzanna. 2009. 15. Definiteness (synchrony). In Sebastian Kempgen, Peter Kosta, Tilman Berger & Karl Gutschmidt (eds.), *Die slavischen Sprachen / The Slavic Languages*, 176–188. Berlin, New York: De Gruyter Mouton. DOI: doi:10.1515/9783110214475.1.3.176.

van Gelderen, Elly. 2011. *The linguistic cycle: Language change and the language faculty*. New York: Oxford University Press. DOI: 10.1093/acprof:oso/9780199756056.001.0001.

Weiss, Daniel. 2004. The rise of an indefinite article: The case of Macedonian *eden*. In Walter Bisang, Nikolaus P. Himmelmann & Björn Wiemer (eds.), *What makes grammaticalization? A look from its fringes and its components*, 139–168. Berlin, New York: De Gruyter Mouton. DOI: doi:10.1515/9783110197440.2.139.

Chapter 5

Multiple wh-fronting in a typological setting: What is behind multiple wh-fronting?

Željko Bošković

University of Connecticut

The paper establishes broad typological correlations between multiple wh-fronting (MWF) and other phenomena in an attempt to understand what is behind MWF. In particular, the paper establishes a correlation between MWF and the morphological shape of wh-words, which is argued to be responsible for MWF. MWF languages are also shown to be characterized by a particular status regarding articles: they either lack articles or have affixal articles (the difference is shown to matter for superiority effects). Certain cases of non-wh indefinite interpretations of wh-phrases and the exceptional behavior of D-linked wh-phrases regarding MWF – they are not subject to it – are also discussed and captured (including Hungarian, where D-linked wh-phrases are not exceptional in this respect, which is tied to another exceptional property of Hungarian).

1 What is special about multiple wh-fronting?

The goal of this paper is to shed light on what is behind one particular language type regarding multiple questions. Most languages front one question word/wh-phrase or leave them all in situ in multiple questions. The former type is illustrated by English (1) and the latter by Chinese (2).

(1)　What did John give to who?

(2)　John gei-le　　shei shenme?　　　　　　　　　　　　　(Chinese)
　　　John give-PFV who what
　　　'What did John give to who?'

Željko Bošković

There is another pattern, which is not frequent crosslinguistically: the so-called multiple wh-fronting languages (MWF), which front all wh-phrases in questions. The pattern is illustrated by Serbo-Croatian (SC) examples in (3) (note that SC is an SVO language).[1]

(3) a. Ko   šta   kupuje?   (SC)
         who what buys
         'Who is buying what?'

     b. *Ko   kupuje šta?
         who buys   what

There have been quite a few works on MWF in the generative tradition since the seminal paper by Rudin (1988) (MWF has been discussed less outside of that tradition, but see e.g. Mycock 2007). These works generally focus on examining the structure and the derivation of MWF constructions. However, they do not attempt to understand what is really behind MWF, why some languages employ this strategy.

This paper aims to address that question, but from a broad typological perspective, in particular, by establishing correlations between MWF and other phenomena. Its scope will be limited – I will not go into the derivation and the structure of

---

[1]There are some highly specific contexts where MWF languages need not front wh-phrases (just like there are contexts where English can employ wh-in-situ). I will generally not be concerned with those exceptional contexts here (apart from D-linking), just with the broad, main pattern. I merely note that, as discussed in Bošković (2002), several of those exceptional contexts involve PF issues, e.g. the case where the fronted wh-phrases would yield a sequence of homophonous elements, like Romanian (i). Bošković (2002) shows that such cases are exceptional only superficially – they still involve MWF in the syntax, with pronunciation of a lower copy of a moved wh-phrase (second *ce* in (i.a), which is motivated by PF considerations. Thus, the second wh-phrase in (i.a) licenses parasitic gaps (see (ii), which is a test for movement in overt syntax (compare (iii.a) and (iii.b).

(i)   a. Ce   precede ce?   (Romanian)
          what precedes what

      b. *Ce   ce   precede?
          what what precedes

(ii) Ce   precede ce   fără   să   influenţeze?   (Romanian)
      what precedes what without SUBJ.PART influences
      'What precedes what without influencing?'

(iii) a. *What precedes what without influencing?
      b. What did Mary promote without influencing?

MWF constructions; the goal of the paper is simply to establish, and understand, prerequisites for the MWF pattern, in an effort to understand what is behind this strategy of forming multiple questions. The discussion will be based on the following 18 (typologically diverse) MWF languages: SC, Romanian, Polish, Russian, Bulgarian, Macedonian, Czech, Slovenian, Ukrainian, Yiddish, Lithuanian, Hungarian, Basque, Mohawk, Georgian, Ossetic, Svan, and Latin.[2] Latin will turn out to be particularly useful, since it can be compared with modern Romance languages.

What will be important for our purposes is the notion of indeterminates (the term goes back to Kuroda 1965, who actually took it from traditional Japanese grammars, which use the term "indeterminate words"). In many languages, the same forms that are used for wh-words have a variety of usages, like existentials, universal quantifiers, negative concord/negative polarity items, free choice, depending on the context where they occur (for much relevant discussion, see Haspelmath 1997). They are referred to as indeterminates since their exact quantificational force is not inherently determined – it is determined by the licensing context in which they are found.

Cheng (1991), a predecessor of this work, observes that Bulgarian, Polish, and Hungarian have indeterminate systems. It turns out that all MWF languages from above have a productive indeterminate system, which suggests that the indeterminate system is a prerequisite for MWF. But there is more to it. There are different types of indeterminate systems. I define here a particular type, which I will refer to as the sub-wh system. It is a fully productive system where addition of an inseparable affix to a wh-phrase results in a series of meanings shown in SC (4).

(4) a. ko 'who' (SC)
  b. *i*ko 'anyone'
  c. *n*iko 'no one'
  d. *ne*ko 'someone'
  e. *sva*ko 'everyone'
  f. *bilo* ko 'whoever'

---

[2]The list includes languages I was able to identify as having MWF (and determine for them the additional information that is needed in the discussion below) based on literature surveys (most of them are well-known as MWF languages; for some less-known cases, see Ledgeway 2012 for Latin, Baker 1996 for Mohawk, Gillon & Armoskaite 2015 for Lithuanian, Erschler 2012 for Ossetic, Erschler 2015 for Georgian and Svan).

Željko Bošković

There is a morphological subset-superset relationship between the wh/question usage and other usages, as stated in (5) regarding 'who'.

(5) sub-wh system: *who*+X for other pronouns (inseparable, fully productive, order doesn't matter)

What is not a sub-wh system is the situation found in Chinese, where the same form can have different functions, as illustrated by (6), or Japanese, where a particle occurs on each function – in some cases inseparable (namely, existential), in some cases separable – note that *-ka*, which is always separated on the wh-usage in Standard Japanese, need not be separated in Okinawan, as illustrated by (7).

(6) a. Ni   xiang mai *shenme* (ne)?   (Chinese)
         you want buy what   Q
         'What do you want to buy?'

     b. Wo bu   xiang mai *shenme.*
         I   not want buy anything
         'I don't want to buy anything.'

     c. Wo xiang mai yi-dian *shenme.*
         I   want buy one-CL something
         'I want to buy something.'

(7) a. Taruu-ja *nuu* koota-*ga.*   (Okinawan, Kinjo & Oseki 2016)
         Taro-TOP what bought-Q
         'What did Taro buy?'

     b. Taruu-ja *nuu-ga* koota-ra.
         Taro-TOP what-Q bought-RA
         'What did Taro buy?'

It should be noted that it has been argued that the Q marker starts with the wh-phrase even in Standard Japanese (just as in Okinawan), see e.g. Hagstrom (1998). This is then a rather different system from SC, where the wh-form is a subset of everything.[3]

   English also does not have a sub-wh system since the relevant system is not fully productive in English (compare *some*where, *every*where, *no*where, *any*where with **some*who/*every*who/*no*who, **no*what/*no*when/*no*how), i.e., it is lexicalized

---

[3]Japanese is, however, more similar to SC in the relevant respect than Chinese is, which may not be surprising in light of the discussion below given that Japanese in fact used to be a MWF language (i.e. Old Japanese appears to have been a MWF language; see Aldridge 2009, Dadan 2019).

(Cheng 1991 suggests that the good cases are lexically incorporated forms, essentially compounds).

Returning to MWF languages, it turns out that all MWF languages have exactly the sub-wh type of indeterminates, which leads me to posit (8) (note that this is a one-way correlation).

(8) If a language has multiple wh-fronting, it has a sub-wh indeterminate system.

This was illustrated above with SC in (4). Additional confirmations of (8) are provided by the MWF languages in Tables 1–4 (the data in Tables 1–7 are from, or based on, Haspelmath 1997; only partial paradigms are given below, and not all series are illustrated – all these languages have additional series; for more complete paradigms, see Haspelmath 1997).[4]

Table 1: Russian indeterminate series

	interrogative	existential	neg-concord	free choice
person	kto	kto-to	ni-kto	kto ugodno
thing	čto	čto-to	ni-čto	čto ugodno
place	gde	gde-to	ni-gde	gde ugodno
time	kogda	kogda-to	ni-kogda	kogda ugodno
manner	kak	kak-to	ni-kak	kak ugodno

Particularly interesting for our purposes is Romance. Latin was clearly a MWF language (see Ledgeway 2012 and Dadan 2019 for extensive discussion) and had a fully productive sub-wh system. The fully productive sub-wh system got lost in

---

[4]I do not consider German as having a productive sub-wh system since in German only one series, the *irgend*-series (but not the *etwas*- or *n*-series, which are the respective second and fourth examples in (i)), is related to wh-words, as shown by (i) (data from Haspelmath 1997; note, however, that (8) is a *one-way* correlation).

(i) a. *person*: wer, jemand, irgend-wer / irgend-jemand, niemand
     b. *thing*: was, etwas, irgend-was / irgend-etwas, nichts
     c. *place*: wo, –, irgend-wo, nirgends
     d. *time*: wann, –, irgend-wann, nie
     e. *manner*: wie, –, irgend-wie, (auf keine Weise)
     f. *determiner*: welche, (ein), irgend-ein / irgend-welche, kein

Željko Bošković

Table 2: Bulgarian indeterminate series

	interrogative	existential	neg-concord	free choice
person	koj	nja-koj	ni-koj	koj to i da e
thing	što	ne-što	ni-što	što to i da e
place	kâde	nja-kâde	ni-kâde	kâde to i da e
time	koga	nja-koga	ni-koga	koga to i da e
manner	kak	nja-kak	ni-kak	kak to i da e

Table 3: Hungarian indeterminate series

	interrogative	existential	neg-concord	free choice
person	ki	vala-ki	sen-ki	akár-ki
thing	mi	vala-mi	sem-mi	akár-mi
place	hol	vala-hol	se-hol	akár-hol
time	mikor	vala-mikor	sem-mikor	akár-mikor
manner	hogy(an)	vala-hogy(an)	se-hogy(an)	akár-hogy(an)

Table 4: Basque indeterminate series

	interro-gative	*bait*-series (non-emphatic)	*i*-series (NPI)	*edo*-series (free choice)	*nahi*-series (free choice)
person	nor	nor-bait	i-nor	edo-nor	nor-nahi
thing	zer	zer-bait	e-zer	edo-zer	zer-nahi
place	non	non-bait	i-non	edo-non	non-nahi
time	noiz	noiz-bait	i-noiz	edo-noiz	noiz-nahi
manner	nola	nola-bait	i-nola	edo-nola	nola-nahi
determiner	zein	–	–	edo-zein	zein-nahi

all modern Romance languages except one: Romanian, which is the only modern Romance language that still has MWF, a strong confirmation of (8). A partial illustration of the Romance situation is given in Tables 5–7.

Table 5: Latin indeterminate series

	interrogative	existential	polarity	free choice
person	quis	ali-quis	quis-quam	qui-vis
thing	quid	ali-quid	quid-quam	quid-vis
place	ubi	ali-cubi	usquam	ubi-vis
time	quando	ali-quando	umquam	–

Table 6: Italian

	interrogative	existential	neg-concord
person	chi	qualcuno	nessuno
thing	che	qualche cosa, qualcosa	niente, nulla
place	dove	in qualche luogo	in nessun luogo
time	quando	qualque volta	(mai)

Table 7: Romanian indeterminate series

	interrogative	existential	free choice	*oare*-series
person	cine	cine-va	ori-cine	oare-cine
thing	ce	ce-va	ori-ce	oare-ce
place	unde	unde-va	ori-unde	oare-unde
time	cînd	cînd-va	ori-cînd	oare-cînd

I conclude therefore that a sub-wh system is a prerequisite for MWF. I will now briefly discuss why that is the case.

The crucial point is that *ko* in (4b) is actually not 'who', i.e. it does not correspond to English *who*. The form is a true indeterminate, which means that it does not have an inherent quantificational force (see below for evidence to this effect). It requires licensing, which also determines its quantificational force (i.e. its exact meaning).

Željko Bošković

The particles that indeterminates merge with normally do that – they determine the exact quantificational force, and the meaning of the indeterminate in cases like those given in SC (9) as a partial illustration of the relevant SC paradigm.[5]

(9) a. i+*ko* 'anyone'   (SC)
     b. n+i+*ko* 'no one'

Importantly, in a sub-wh system, the *only* usage on which the indeterminate is not merged with a particle is the wh-usage, which means that we are dealing here with an unlicensed indeterminate. I suggest that this is what requires fronting. The indeterminate is licensed as a wh-phrase by moving to an interrogative projection (which determines its meaning). The movement thus does not occur because of a property of the interrogative head (which is the case in English, where only one wh-phrase fronts because of that), but because of indeterminate licensing – this is why they *all* need to undergo fronting, resulting in MWF.

In short, in the sub-wh system, affixes merged with an indeterminate determine its quantificational force and license the indeterminate. When there is no such affix, the indeterminate is licensed as a wh-phrase by movement to an interrogative projection.

MWF languages do however have certain cases where the wh-phrase itself (so the form that is used in wh-questions) receives a different, non-wh interpretation, like the wh-existential in (10) (see e.g. Izvorski 1998, Bošković 2002, Šimík 2011).

(10) a. Ima ko šta da ti proda.   (SC)
         has who what that you sells
         'There is someone who can sell you something.'
      b. *Ima ko da ti proda šta.
         has who that you sells what

Importantly, the relevant elements must front here. The fronting does not occur to the interrogative projection, since the relevant clause is simply not interrogative. I suggest that since *ko* and *šta* are not merged with an indefinite particle in these cases, they are licensed as indefinites by moving to a special indefinite licensing position. What is relevant here is languages like Kaqchikel, where the

---

[5]In these particular cases, the morphology is rather transparent. *I-* also means 'even', on the connection between 'even' and NPIs see e.g. Rooth (1985), Haspelmath (1997), Giannakidou (2007), Crnič (2011); *n-* may indicate a connection with negation. At any rate, these details are not important for our purposes.

exact same form functions as interrogative or indefinite, and must be fronted on both functions, with the landing site of the interrogative being higher than the indefinite licensing projection, as discussed in detail in Erlewine (2016). What Kaqchikel shows is that there is a pattern where the indefinite meaning of an indeterminate is licensed by movement to a special projection that licenses this meaning (see Erlewine 2016). The suggestion is that this is precisely what happens in (10) (the movement is not to the interrogative CP projection since the relevant clauses are clearly not interrogative; note that this (i.e. (10)) can also be taken to confirm that the relevant elements are not inherently wh-phrases but bare indeterminates).[6]

It is worth noting that a number of Australian languages have the same form for wh-phrases and indefinites but while the morphology is the same the syntax is not: as wh-phrases they must front, as indefinites they stay in situ (these languages cannot be checked for MWF since they do not allow multiple questions in the first place, see Cheng 1991 for relevant discussion of these languages).

(11) a. ngana nganhu wartirra nyina-nguru karra-ngka
who that.NOM woman sit-PRS scrub-LOC
muyinu-npi-rra?
hidden-INCH-CTEMP
'Who is that woman hiding in the scrub?'

b. ngayu nyina-lha martama-l.yarra palykura-la
1SG.NOM sit-PRS press.on-CTEMP groundsheet-LOC
nganangu-la.
someone.GEN-LOC
'I sat down on someone's groundsheet, holding it down.'

(Martuthunira; Dench 1987)

(12) a. ngatha ngananhalu nhantha-nnguli-nha.
1SG.NOM something.INS bit-PASS-PST
'I was bitten by something.'

---

[6]The movement strategy just discussed and the affixation strategy for licensing indefinites can be combined, though this option is slightly disfavored, possibly due to a parallelism for indefinite licensing being favored.

(i) ? Ima ko da ti proda nešto.   (SC)
has who that you sells something
'There is someone who can sell you something.'

Željko Bošković

    b. ngananha-ma-rna nyinta   ngunhalku?
       what-CAUS-PST   2SG.NOM that.ACC
       'What have you done to him?'   (Panyjima; Dench 1981)

There is a parallel situation with MWF languages. In particular, there are similar wh-indefinites in Slavic MWF languages, as illustrated by Russian (13) (see e.g. Zanon 2022, Hengeveld et al. 2023).

(13)   Možet, kto prixodil.
      maybe who came
      'Maybe someone came.'   (Russian; Hengeveld et al. 2023)

This usage is very restricted in Slavic; in SC even more so than in Russian – (13) is in fact unacceptable in SC; regarding Russian, see especially Zanon (2022), who argues that the relevant elements are licensed by a semantically motivated and constrained null operator, which essentially plays the role of the licensing affixes discussed above hence this kind of analysis of the usage in question can be adjusted to the system developed here. Alternatively, it is possible that an indeterminate that does not have a licensing particle attached and does not move to an indeterminate-licensing projection or has a linking index (see the discussion right below) is interpreted by a default rule for unlicensed indeterminates, which would apply in the relevant contexts in the languages that allow this usage (they also differ regarding such contexts), as a simple indefinite. In this respect, it is worth noting that such indefinites cannot occur in wh-questions (see, e.g., Zanon 2022 and Hengeveld et al. 2023), which can be taken to confirm the default nature of the licensing in question – it is available only if another way is not available.[7]

    Interestingly, Zanon (2022) and Hengeveld et al. (2023) observe that these wh-indefinites cannot be focused. What is important here is that real MWF/wh-fronting in Slavic has been analyzed as focus-movement (e.g. Bošković 2002, see also Stepanov 1998 for Russian as well as the discussion below), i.e. it is essentially focusing. It then makes sense that if the relevant element is focused it would be interpreted as a wh, not a non-wh (i.e. indefinite), hence the non-wh-indefinite usage does not allow focalization. Hengeveld et al. (2023) actually observe that the non-focusing requirement is not general – it does not hold in Dutch. Given the current discussion, Dutch-like exceptions should not be possible in MWF languages.

---

[7]Note that these indefinites are different from those in wh-existentials like (10) – e.g. Ksenia Zanon (p.c.) notes that the former cannot be coordinated, see Zanon (2022), while the latter can be.

A different (and independent) exception to MWF concerns D-linked wh-phrases, which need not undergo fronting, as illustrated below by SC (14). (Note, however, that this is not the case in all MWF languages; they must front in Hungarian, which is discussed in §2.)[8]

(14) Ko   kupuje koju   knjigu?   (SC)
      who buys   which book
      'Who is buying which book?'

Two issues are relevant here. First, *koju* is not an indeterminate but a wh-specific form (this may not be a general situation though). Second, as briefly noted above, Bošković (2002) argues that MWF is actually movement to a focus projection, this means that the relevant licensing takes place in the Spec of a focus-licensing head; this by itself is not surprising – focus/interrogativity connection has often been noted.[9] Furthermore, Bošković (2002) observes that D-linking is very different from focus. With D-linked wh-phrases the range of felicitous answers is restricted by a set of objects that is familiar to the speaker and the hearer as a result of it being referred to/salient in the context. In other words, the range

---

[8]It may be worth noting here that D-linked wh-phrases more generally can be special, and subject to ill-understood language variation. Thus, there are languages that disallow multiple questions, e.g. Hong Kong Sign Language (HKSL), Italian, and Mandinka. ((i)–(iv) are taken from Gan 2022).

  (i)   * WHO BUY WHAT?   (HKSL)

  (ii) * Chi ha scritto che   cosa?   (Italian)
          who has written what

Gan (2022) shows that D-linking improves multiple questions in HKSL and Mandinka, but not in Italian (it is not out of question that there is some connection here with the SC vs. Hungarian difference regarding D-linked MWF questions).

  (iii) STUDENT WHO BUY COMPUTER BUY-WHICH?   (HKSL)
         'Which student bought which book?'

  (iv) * Quale studente comprerà quale libro?   (Italian)
          which student   will-buy which book

[9]I assume that as a result of this connection, indeterminates can still be licensed as interrogative in such a projection. (Possibly, being in such a projection would enable them to undergo unselective binding with interrogative C in spite of the issue noted in §2.1 (i.e. without a null operator, the intuition being that it is not needed in this case since the relevant element is located in an operator, in fact the right operator, position—SpecFocP), which would license their interrogative interpretation.)

of reference of D-linked wh-phrases is discourse-given. Due to their discourse givenness, such wh-phrases are not focused, hence they are not subject to focus movement. (One wh-phrase always must front for clausal typing as discussed in Cheng 1991 so when only a D-linked wh-phrase is present it fronts but Bošković 2002 shows that the landing site is different; for special behavior of D-linked wh-phrases regarding MWF see also Diesing 2003 on Yiddish, which disallows MWF with D-linked wh-phrases.)

Regarding the interpretation of D-linked wh-phrases, Enç (2003) proposes that specific arguments have a linking index ℓ which identifies the set of individuals of which the argument is a member (i.e. it gives the set which that argument must belong to). Non-specific arguments have no such index. Shields (2008) extends this to wh-phrases: D-linked wh-phrases are specific and therefore have a set-denoting (linking) index, which non-D-linking wh-phrases do not have. The linking index points to the set of entities in the discourse that a specific expression is required to be a member of.

Indeterminate pronouns are normally non-specific, D-linked ones (i.e. D-linked wh-phrases) are not. The interpretation of the latter is essentially determined by their semantics, no further licensing is needed (essentially, an indeterminate with a linking index is interpreted as D-linked – the linking index points to the set of entities in the discourse that the relevant element is required to be a member of). It is also possible that the linking index allows D-linked wh-phrases to undergo unselective binding by interrogative C and that they are licensed in that way (see Pesetsky 1987 on unselective binding of D-linked wh-phrases; see also §2).[10]

In conclusion, this section has established a correlation between MWF and another phenomenon. In particular, MWF languages have been shown to have

---

[10]There is an alternative account. A number of authors (e.g. Belletti 2004, Lacerda 2020) have argued for several languages that they have a low topic projection. It is possible that D-linked wh-phrases are licensed in a low topic-like projection (see Grohmann 2006 for D-linking as topichood). On this analysis, the D-linked wh-phrase in (14) would not actually be in situ (SC and Hungarian could then differ here regarding topic movement; see, however, below). It is worth noting here that (i) is also acceptable. Bošković (2002), however, shows that the D-linked wh-phrase in such cases is lower than the second wh-phrase in examples like (3a), i.e. it is not the case that the D-linked wh-phrase simply optionally undergoes movement that the second wh-phrase must undergo in (3) (examples like (i), i.e. optional fronting, is actually not allowed in all MWF languages, see Bošković 2002, Pesetsky 1987, Wachowicz 1974).

(i) Ko koju knjigu kupuje? (SC)
who which book buys
'Who is buying which book?'

a sub-wh indeterminate systems, which forces MWF (except with D-linked wh-phrases).

2 Multiple wh-fronting and articles

2.1 Another generalization

I will now show that there is another property that MWF languages have in common, which is in principle independent of the one presented in §1 (in the sense that if one of the generalizations in questions turns out not to be correct the other one would not necessarily be affected).[11] In particular, they all either lack definite articles or have affixal definite articles (15). The relevant language cut is given in (16).[12]

---

[11]But see the generalization regarding indeterminates themselves in Oda (2022) that would actually relate (8) and (15). Oda also provides an alternative deduction of (8) based on my earlier version of this generalization given in Bošković (2020) where the prerequisite for MWF was a broader indeterminate system than the sub wh-system.

[12]For most of the languages listed in (16b), their affixal status is well-known. For arguments that Hungarian definite article is affixal (more precisely, a prefix), see MacWhinney (1976), Oda (2022), and Lewis (2024). MacWhinney observes that it undergoes a morphophonemic alternation that is typical of affixes, while Oda and Lewis observe typological generalizations where Hungarian patterns with languages with affixal articles (languages with affixal articles actually pattern with languages without articles regarding those generalizations). Regarding the affixal status of the definite article in Yiddish, which might be the least discussed case here, see Oda (2022). To mention some relevant arguments, Talić (2017) and Oda (2022) observe that languages with affixal definite articles allow article omission in contexts where such omission is not possible in free-standing article languages like English. Oda notes that this is especially the case in PPs, where due to article omission a bare noun can even receive a definite interpretation in (some) affixal article languages, which is never possible in languages with non-affixal definite articles, where a definite article is required for definite interpretation (see Bošković 2016; Oda argues that in the relevant cases the preposition essentially functions as the definite article). Thus, Zwicky (1984: 119) observes regarding (i.b): "The phrase *in gloz* in 'in the glass' is a typical example. The noun *gloz* in this expression is understood definitely, and can even be anaphoric."

(i) a. lebn tir
     near door
     'near the door'

    b. in gloz = in the glass   (Yiddish; Zwicky 1984)

Bošković (2016) also notes that, for the purposes of Bošković's NP/DP generalizations (see below for some relevant discussion), definite articles have a form distinct from demonstratives. Definite articles in Yiddish have the same form as demonstratives, with stress distinguishing them. Margolis (2011: 122) in fact states that: "this/these" is identical to the definite article with

Željko Bošković

(15)   MWF languages either lack articles or have affixal definite articles.

(16)   a.   No articles: SC, Polish, Russian, Czech, Slovenian, Ukrainian, Mohawk, Latin, Georgian, Lithuanian, Ossetic, Swan
        b.   Affixal articles: Romanian, Bulgarian, Macedonian, Basque, Hungarian, Yiddish

Turning to the deduction of (15), in a series of works (e.g. Bošković 2012), based on a number of syntactic and semantic typological generalizations, where languages with and without definite articles consistently differ regarding a number of syntactic and semantic phenomena, I argued that languages without definite articles do not project DP (i.e., there are no null definite articles in such languages).

Talić (2017) argues for a refinement of the NP/DP language distinction; she shows that in many respects languages with affixal definite articles behave like a separate type (see also Oda 2022, Lewis 2024), in that they sometimes behave like languages with articles and sometimes like those without articles.[13]

In Bošković (2020) I suggested an implementation of this observation for the affixal article languages that have MWF: there is D in such languages, but there is no DP. The affixal article is base-generated adjoined to N (more precisely, its host). It should be noted that there is nothing strange about this theoretically: Adjunction through movement can involve either phrasal or head adjunction, the same should hold for adjunction through base-generation (for much relevant discussion regarding definite articles, see also Oda 2022; regarding indefinite articles, see Wang 2019).
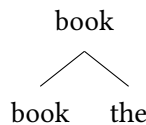
book

book   the

Figure 1: Noun-article base-generation

Recall now that in a sub wh-system, only on the wh-usage the indeterminate does not occur with a licensing particle. I suggest then that, in principle, such

---

added stress. Essentially following Oda (2022), I thus consider Yiddish to be an affixal article language, the definite article being an affixal, hence unstressed, version of the demonstrative (there may be a change under way regarding the status of the relevant element where dialectal differences may also be relevant; not all dialects of Yiddish in fact have MWF, see Diesing 2003).

[13]Below, for ease of exposition I will simply use the term (affixal) article, though what matters here (and what matters for Bošković's NP/DP generalizations) is definite articles only.

indeterminates can still be licensed at a distance *in situ*, with a null operator in SpecDP that is unselectively bound by interrogative C. This is not possible in MWF languages due to the lack of a DP projection that would be capable of such licensing. The only way to license the indeterminate on the wh-usage is then to front it to an interrogative position.[14]

A confluence of independent factors, namely the sub wh-system and a particular status regarding articles, is what is behind MWF: MWF languages have a sub-wh indeterminate system, and either lack articles or have affixal articles, which are the typological findings of this paper.

Regarding the relevance of the latter property, in languages without articles and languages with affixal articles the possibility of wh-licensing *in situ* by interrogative C through unselective binding is blocked because such licensing is done through a null operator in SpecDP (except with D-linked wh-phrases), which is absent in languages without articles and languages with affixal articles (in the former, because DP itself is lacking, and in the latter because the affixal article is base-generated adjoined to N, which means that in such languages there is D, but there is still no DP, hence no null operator in SpecDP).

2.2 Superiority variation regarding basic Superiority effects

I turn now to a case of variation within MWF languages which will also shed light on the exceptional status of Hungarian regarding D-linked wh-phrases, noted in §1. Already Rudin (1988) observed that MWF languages differ regarding whether they show ordering, i.e. Superiority, effects with MWF. Regarding basic cases like those shown in (17)–(18), SC does not show them, while Bulgarian does show them.

(17)   a. Koj kakvo e kupil?   (Bulgarian)
           who what   is bought

       b. * Kakvo koj   e kupil?
            what   who is bought
      (*Intended*:) 'Who bought what?'

(18)   a. Ko   šta   kupuje?   (SC)
           who what buys

       b. Šta   ko   kupuje?
           what who buys

---

[14]As suggested above, D-linked wh-phrases may be able to undergo unselective binding even in the absence of DP for independent reasons, namely, due to the presence of the linking index.

Željko Bošković

'Who is buying what?'

A survey of the literature shows the following language cut regarding Superiority effects in basic cases of this sort.

(19) a. No Superiority effects: SC, Polish, Czech, Russian, Slovenian, Ukrainian, Mohawk, Lithuanian, Georgian, Ossetic, Svan, Hungarian

      b. Superiority effects: Romanian, Bulgarian, Macedonian, Basque, Yiddish

It turns out that the cut is not arbitrary – there is a correlation with (the type of) articles. Putting Hungarian aside (taking Hungarian into consideration we would have a one-way correlation in (21), which was actually noted in Bošković 2008), we have (20).

(20) MWF languages without articles do not show basic Superiority effects, those with affixal articles do.

(21) MWF languages without articles do not show basic Superiority effects.

Below, I will briefly outline a deduction of (20) that will also accommodate the Hungarian exception (given the affixal status of the Hungarian definite article, see fn. 12), tying it to another Hungarian exception, namely the exceptional behavior of Hungarian regarding D-linking.

Bošković (2002) argues that Superiority effects arise with MWF to SpecCP (English-style wh-movement), not with MWF to a lower position, which means that SC MWF targets a lower position than Bulgarian MWF (see Bošković 2002 for evidence to this effect). Now, if Superiority is taken to be a sign of true, English-style wh-movement, this can be generalized in such a way that languages with articles (non-affixal or affixal) must have true English-style wh-movement to SpecCP when fronting wh-phrases. Bošković (2008) in fact suggests that the D-feature is crucially involved in movement to SpecCP. Affixal article languages still have the D-feature, which means that they have wh-movement to SpecCP, which is Superiority inducing. This then captures (20). But what about Hungarian?

Superiority as a test for wh-movement is confirmed by single-pair (SP)/pair-list (PL) answers. Bošković (2001, 2002) shows that overt wh-movement languages require a PL answer for examples like (22). (22) cannot be felicitously asked in the following situation: John is in a store and sees somebody buying an article of clothing, but does not see who it is and does not see exactly what the person is buying. He goes to the sales clerk and asks (22).

148

(22) Who bought what?

Whereas German patterns with English, wh-in-situ languages Japanese, Hindi, and Chinese allow SP answers in such questions (see Bošković 2001). Importantly, French allows SP answers, but only with in-situ questions like (23a), not (23b).

(23) a. Il a   donné quoi à qui?   (French)
          he has given what to who

      b. Qu' a-t-il   donné à qui?
          what has-he given to who
      'What did he give to who?'

Based on this, Bošković (2001, 2002) argues that the availability of SP answers depends on the possibility of not moving any wh-phrase to SpecCP overtly (see Bošković 2001 for an account of this generalization).

Turning to MWF languages, SC allows SP answers, while Bulgarian does not, which confirms that SC MWF lands in a lower position than Bulgarian MWF (see Bošković 2007 and references therein for additional languages confirming this).

As noted above, Bošković (2002) argues that MWF involves focus. Now, Bošković (1999) argues that movement-attracting heads can differ regarding the specification of the movement-attracting feature. They can be specified to attract one element with the relevant feature, call it F, or all elements with the F feature. English interrogative C is an attract 1-F head – it attracts one (in particular, the highest) element with the wh-feature. In SC, wh-phrases undergo focus movement; the relevant head has the specification Attract All-focus. Bulgarian is a combination of English and SC: It has single-fronting wh-movement as in English (Attract 1-wh) and MWF for focus (Attract All-focus, see Bošković 1999 and fn. 16). Importantly, from this perspective, Superiority is not a diagnostic of wh-movement, but single fronting.[15] In this respect, Bošković (2002) shows that there are selec-

---

[15]Given the economy-of-derivation condition that every requirement be satisfied through the shortest movement possible, Attract 1-F heads will always attract the highest phrase with the relevant feature: thus, in (i), the relevant formal inadequacy of the interrogative C is checked through a shorter movement in (i.a) than in (i.b) (cf. the pre-wh-movement structure in (i.c)).

(i) a. Who$_i$ did Mary tell t$_i$ to buy the book?

     b. * What$_i$ did Mary tell who to buy t$_i$?

     c. Mary tell who to buy what

With Attract All-F heads, like the SC focus-licensing head, all relevant elements must move: Regardless of the order of movement, the same number of nodes are crossed with such movement, hence the order of movement of wh-phrases is free (see Bošković 1999 for a more detailed discussion).

Željko Bošković

tive Superiority effects in Bulgarian. Only the first wh-phrase, which is the only wh-phrase that undergoes wh-movement, is subject to Superiority effects, other wh-phrases are not. Thus, the indirect object wh-phrase must precede the direct object wh-phrase in (24) (because it is higher than the object wh-phrase before wh-fronting) but not in (25), where a subject wh-phrase, which is higher than both indirect and direct object wh-phrase before wh-fronting, is present.[16]

(24) a. Kogo kakvo e pital Ivan?   (Bulgarian)
          whom what is asked Ivan

      b. ?* Kakvo kogo e pital Ivan?
             what   whom is asked Ivan
         'Who did Ivan ask what?'

(25) a. Koj kogo kakvo e pital?   (Bulgarian)
          who whom what   is asked

      b. Koj kakvo kogo e pital?
          who what   whom is asked
         'Who asked who what?'

All this raises a question: Is there a MWF language where D-linked wh-phrases also must front? That would be a true MWF counterpart of English, with an Attract All-wh specification (note that Attract All-wh affects D-linked wh-phrases, in contrast to Attract All-focus). As noted above, and as discussed in Bošković (2007) and É. Kiss (2002), both D-linked and non-D-linked wh-phrases must move in Hungarian. This is illustrated by (26)–(27).

(26) a. * Ki   irt   mit?
             who wrote what

      b. Ki   mit   irt?
          who what wrote

---

[16]Note that, as discussed in Bošković (1999), it is the same head, interrogative C, that has the relevant properties (Attract 1-wh, Attract All-focus) in Bulgarian. Given that the first wh-phrase that moves to SpecCP automatically satisfies the Attract 1-wh requirement (see Bošković 1999), the highest wh-phrase must move first, then the order of movement does not matter, since Attract All-focus does not care about the order of movement, as noted in fn. 15. (Note that, as standardly assumed, the order of fronted wh-phrases reflects the order of their movement, see Rudin (1988), Richards (2001) for different implementations of this, i.e. the wh-phrase that is first in the linear order is the one that moves first, hence the highest wh-phrase must move first when Superiority is in effect.)

   c. Mit ki   irt?
      what who wrote
      (*Intended*:) 'Who wrote what?'   (Hungarian; Bošković 2007)

(27) a. * Ki   irta   melyik levelet?
          who wrote which letter

     b. Ki   melyik levelet irta?
        who which letter wrote

     c. Melyik levelet ki   irta?
       which letter who wrote
      (*Intended*:) 'Who wrote which letter?'   (Hungarian; Bošković 2007)

Importantly, Hungarian MWF questions also disallow SP answers (see, e.g., Surányi 2005) and do not show Superiority effects (see (26)), which is exactly the behavior expected of a true MWF counterpart of English (there are no Superiority effects since we are dealing only with an Attract-All fronting and SP answers are disallowed because the fronting is to SpecCP).[17] What appeared to be an exceptional behavior of Hungarian regarding Superiority and D-linking is thus explained, in fact in a uniform manner.

    At any rate, the discussion from §2 is summarized below in table form (where the left column gives the relevant language types – there are two types for affixal article languages, depending on whether D-linked wh-phrases are also subject to MWF).

Table 8: Summary

	MWF	Superiority with MWF	SP with wh-fronting
Free-standing article	*	N/A	*
Affixal article		Yes	*
Affixal article + D-linking MWF		No	*
No article		No	

---

[17]Horváth (1998), Puskás (2000), Lipták (2001), and É. Kiss (2002) suggest that the wh-phrase that is closest to the verb in Hungarian MWF questions undergoes focus-movement, other wh-phrases undergo movement that non-wh-quantifiers undergo, but see Surányi (2005) for arguments against this position.

Željko Bošković

3 Conclusion

The paper has established correlations between MWF and other phenomena, in an attempt to understand what is behind MWF. In particular, MWF languages have been shown to have a sub-wh indeterminate system, which was suggested to force MWF. In such a system, an inseparable affix is attached to the indeterminate, with the exact quantificational force of the indeterminate determined by the affix that merges with it. What is traditionally considered to be wh-phrases in the sub-wh indeterminate system are not really wh-phrases but bare indeterminates; they are not licensed *in situ* because they are bare – no licensing affix is attached to them – hence they must front to a position in the left periphery to get licensed, which in turn determines their interpretation. This yields MWF.

MWF languages are also characterized by a particular status regarding articles – they either lack articles or have affixal articles. It was argued that in these language types the possibility of wh-licensing *in situ* by interrogative C through unselective binding is blocked because such licensing is done through a null operator in SpecDP, which is absent in languages without articles and languages with affixal articles for a principled reason. The distinction between the lack of articles and affixal articles in MWF languages was, however, shown to have an effect on the presence/absence of Superiority effects. The exceptional behavior of D-linked wh-phrases regarding MWF (they don't need to undergo it) was also captured (including the Hungarian pattern, where D-linked wh-phrases are not exceptional in this respect – they are subject to MWF). Certain cases of non-wh indefinite interpretations of wh-phrases were also discussed.

All in all, the paper has established the following generalizations regarding MWF, where Hungarian was shown to be exceptional regarding (28c) but for a principled reason, which was tied to its exceptional behavior regarding D-linking; the reader should thus bear in mind that the way (28c) is deduced in the paper does leave room for principled exceptions.

(28) a. If a language has multiple wh-fronting, it has a sub-wh indeterminate system.
      b. MWF languages either lack articles or have affixal definite articles.
      c. MWF languages without articles do not show basic Superiority effects, those with affixal articles do.

At any rate, the main typological finding of this paper is that a confluence of independent factors, namely the sub-wh indeterminate system and a particular status regarding articles, is what is behind MWF.

Abbreviations

1	first person	PART	particle
2	second person	PASS	passive
ACC	accusative	PFV	perfective
CAUS	causative	PL	pair-list
CL	classifier	PRS	present
CTEMP	contemporaneous relative	PST	past
GEN	genitive	Q	question particle
HKSL	Hong Kong Sign Language	SC	Serbo-Croatian
INCH	inchoative	SG	singular
INS	instrumental	SP	single-pair
LOC	locative	SUBJ	subjunctive
NOM	nominative	TOP	topic
MWF	multiple wh-fronting		

Acknowledgments

References

Aldridge, Edith. 2009. Short wh-movement in Old Japanese. In Shoishi Iwasaki, Hajime Hoji, Patricia M. Clancy & Sung-Ock Sohn (eds.), *Japanese/Korean linguistics, Vol. 17*, 549–563. Stanford, CA: CSLI Publications.

Baker, Mark. 1996. *The polysynthesis parameter*. Oxford: Oxford University Press.

Belletti, Adriana. 2004. Aspects of the low IP area. In Luigi Rizzi (ed.), *The structure of CP and IP: The cartography of syntactic structures, Vol. 2*, 16–51. Oxford: Oxford University Press.

Bošković, Željko. 1999. On multiple feature-checking: Multiple wh-fronting and multiple head-movement. In Samuel Epstein & Norbert Hornstein (eds.), *Working minimalism*, 159–187. Cambridge, MA: MIT Press.

Bošković, Željko. 2001. On the interpretation of multiple questions. *Linguistic Variation Yearbook* 1(1). 1–15. DOI: 10.1075/livy.1.03bos.

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

Željko Bošković

Bošković, Željko. 2007. A note on wh-typology. In Peter Kosta & Lilia Schürcks (eds.), *Linguistic investigations into formal description of Slavic languages: Contributions of the 6th European conference held at Potsdam University*, 159–170. Frankfurt/M.: Peter Lang.

Bošković, Željko. 2008. What will you have, DP or NP? In Emily Elfner & Martin Walkow (eds.), *NELS 37: Proceedings of the 37th annual meeting of the North East Linguistic Society*, 101–114. Amherst, MA: GLSA Publications.

Bošković, Željko. 2012. On NPs and clauses. In Günther Grewendorf & Thomas Ede Zimmermann (eds.), *Discourse and grammar: From sentence types to lexical categories*, 179–242. Berlin: Mouton de Gruyter. DOI: 10.1515/9781614511601.179.

Bošković, Željko. 2016. On second position clitics crosslinguistically. In Franc Lanko Marušič & Rok Žaucer (eds.), *Formal studies in Slovenian syntax: In honor of Janez Orešnik*, 23–44. Amsterdam: John Benjamins. DOI: 10.1075/la.236.02bos.

Bošković, Željko. 2020. *Comparative syntax lectures, fall 2020*. University of Connecticut.

Cheng, Lisa Lai Shen. 1991. *On the typology of wh-questions*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Crnič, Luka. 2011. *Getting even*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Dadan, Marcin. 2019. *Head labeling preference and language change*. Storrs, CT: University of Connecticut. (Doctoral dissertation).

Dench, Alan. 1981. *Panyjima phonology and morphology*. Canberra: The Australian National University. (MA thesis).

Dench, Alan. 1987. *Martuthunira: A language of the Pilbara region of Western Australia*. Canberra: The Australian National University. (Doctoral dissertation).

Diesing, Molly. 2003. On the nature of multiple fronting in Yiddish. In Cedric Boeckx & Kleanthes Grohmann (eds.), *Multiple wh-fronting*, 51–76. Amsterdam: John Benjamins. DOI: 10.1075/la.64.06die.

É. Kiss, Katalin. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.

Enç, Mürvet. 2003. *The syntax of specificity*. Ms. University of Wisconsin-Madison.

Erlewine, Michael Yoshitaka. 2016. Anti-locality and optimality in Kaqchikel agent focus. *Natural Language & Linguistic Theory* 34. 429–479. DOI: 10.1007/s11049-015-9310-z.

Erschler, David. 2012. From preverbal focus to preverbal ''left periphery'': The ossetic clause architecture in areal and diachronic perspective. *Lingua* 122(6). 673–699. DOI: 10.1016/j.lingua.2012.01.009.

Erschler, David. 2015. Embedded questions and sluicing in Georgian and Svan. *Languages of the Caucasus* 1(1). 38–74. DOI: 10.5070/L911029246.

Gan, Linghui. 2022. *Syntactic structure of argument wh-questions in Hong Kong Sign Language (HKSL)*. Presentation at the 28th Annual Conference of The International Association of Chinese Linguistics (IACL-28), Hong Kong. https://drive.google.com/file/d/1DFwkndEPovMonqPa0ibrhg_pw-CAF7HQ/view.

Giannakidou, Anastasia. 2007. The landscape of EVEN. *Natural Language & Linguistic Theory* 25. 39–81. DOI: 10.1007/s11049-006-9006-5.

Gillon, Carrie & Solveiga Armoskaite. 2015. The illusion of the NP/DP divide. *Linguistic Variation* 15(1). 69–115. DOI: 10.1075/lv.15.1.03gil.

Grohmann, Kleanthes. 2006. Top issues in questions: Topics – topicalization – topicalizability. In Lisa Cheng & Norbert Corver (eds.), *Wh-movement moving on*, 249–288. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/7197.003.0017.

Hagstrom, Paul. 1998. *Decomposing questions*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Hengeveld, Kees, Sabine Iatridou & Floris Roelofsen. 2023. Quexistentials and focus. *Linguistic Inquiry* 54(3). 571–624. DOI: 10.1162/ling_a_00441.

Horváth, Julia. 1998. Multiple wh-phrases and the wh-scope-marker strategy in Hungarian interrogatives. *Acta Linguistica Hungarica* 45. 31–60. DOI: 10.1023/A:1009656907847.

Izvorski, Roumyana. 1998. Non-indicative wh-complements of possessive and existential predicates. In Pius N. Tamanji & Kiyomi Kusumoto (eds.), *Proceedings of NELS 28*, 159–173. Amherst, MA: GLSA Publications.

Kinjo, Kunio & Yohei Oseki. 2016. Wh-concord in Okinawan = syntactic movement + morphological merger. *University of Pennsylvania Working Papers in Linguistics* 22(1). 177–186. https://repository.upenn.edu/handle/20.500.14332/45087.

Kuroda, Sige-Yuki. 1965. *Generative grammatical studies in the Japanese language*. Cambridge, MA: Massachusetts Institute of Technology. (Doctoral dissertation).

Lacerda, Renato. 2020. *Middle-field syntax and information structure in Brazilian Portuguese*. Storrs, CT: University of Connecticut. (Doctoral dissertation).

Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford: Oxford University Press.

Lewis, Beccy. 2024. *Cross-linguistic variation in the expression of heterogeneous plural marking and its theoretical consequences*. Storrs, CT: University of Connecticut. (Doctoral dissertation).

Željko Bošković

Lipták, Anikó. 2001. *On the syntax of wh-items in Hungarian* (LOT Dissertation Series 45). Utrecht: LOT.

MacWhinney, Brian. 1976. Hungarian research on the acquisition of morphology and syntax. *Journal of Child Language* 3(3). 397–410. DOI: 10.1017/S0305000900007261.

Margolis, Rebecca. 2011. *Basic Yiddish: A grammar and workbook*. London: Routledge.

Mycock, Louise. 2007. Constituent question formation and focus: A new typological perspective. *Transactions of the Philological Society* 105. 192–251. DOI: 10.1111/j.1467-968X.2007.00188.x.

Oda, Hiromune. 2022. *The NP/DP-language distinction as a scale and parameters in minimalism*. Storrs, CT: University of Connecticut. (Doctoral dissertation).

Pesetsky, David. 1987. Wh-in situ: Movement and unselective binding. In Eric J. Reuland & Alice G. B. ter Meulen (eds.), *The representation of (in)definiteness*, 98–129. Cambridge, MA: MIT Press.

Puskás, Genoveva. 2000. *Word order in Hungarian: The syntax of A-bar positions*. Amsterdam/Philadelphia: John Benjamins.

Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.

Rooth, Mats. 1985. *Association with focus*. Amherst, MA: University of Massachusetts. (Doctoral dissertation).

Rudin, Catherine. 1988. On multiple questions and multiple wh fronting. *Natural Language & Linguistic Theory* 6. 445–501. DOI: 10.1007/BF00134489.

Shields, Rebecca. 2008. *What's so special about d-linking*. Paper presented at NELS 39, Cornell University. https://www.researchgate.net/publication/252605974_What's_so_Special_about_D-Linking.

Šimík, Radek. 2011. *Modal existential wh-constructions* (LOT Dissertation Series 269). Utrecht: LOT.

Stepanov, Arthur. 1998. *On wh-fronting in Russian*. Pius N. Tamanji & Kiyomi Kusumoto (eds.). Amherst, MA: GLSA Publications. 453–467.

Surányi, Balázs. 2005. Triggering wh-fronting. In Christopher Piñón & Siptár Péter (eds.), *Approaches to Hungarian Vol. 9: Papers from the Düsseldorf conference*, 231–259. Budapest: Akadémiai Kiadó.

Talić, Aida. 2017. *Functional and bare projections in the domain of N and A*. Storrs, CT: University of Connecticut. (Doctoral dissertation).

Wachowicz, Krystyna A. 1974. Against the universality of a single wh-question movement. *Foundations of Language* 11. 155–166. https://www.jstor.org/stable/25000765.

Wang, Shuyan. 2019. An intermediate stage of spec-to-head reanalysis: Evidence from Mandarin and Cantonese. In Sae-Youn Cho (ed.), *Proceedings of the 12th Generative Linguistics in the Old World in Asia & the 21st Seoul international conference on generative grammar*, 594–603. Seoul: Hankook Munhwasa.

Zanon, Ksenia. 2022. *Wh-indefinites in Russian.* Paper presented at FDSL 15, HU Berlin.

Zwicky, Arnold M. 1984. "Reduced words" in highly modular theories: Yiddish anarthrous locatives reexamined. *Ohio State University Working Papers in Linguistics* 29. 117–126.

Chapter 6

A quantification-based approach to plural pronoun comitatives

Madeleine Butschety

University of Nova Gorica

Plural pronoun constructions (PPCs) consist of a plural pronoun and a comitative (i.e. *with*-) phrase. In sentences such as *My s Petej pojdëm domoj* (lit. 'We with Petja will-go home') from Russian, PPCs are ambiguous between a default interpretation according to which a plural referent *we* will go home with Petja (=ePPC), and an unexpected interpretation according to which Petja and the speaker will go home (=iPPC). I show that this ambiguity can be derived under the assumption that plural pronouns and (universal) quantifiers have some striking properties in common. In particular, I argue that the unexpected iPPC reading arises if the comitative phrase occurs inside the restrictor of the plural pronoun (which is similar to a quantifier's restrictor), and an ePPC reading arises if it occurs elsewhere in the structure. My account further offers an explanation regarding the availability and distribution of iPPC interpretations within and across Slavic languages.

1 Introduction

Plural pronoun comitatives (often also dubbed "Plural pronoun constructions" in the literature; henceforth PPCs) are complex expressions that consist of a plural pronoun and a comitative (i.e. *with*-) phrase. In many Slavic languages, PPCs can give rise to two different interpretations, paraphrased as (a) and (b) in the examples below.

(1) My s   Petej   pojdëm   domoj.
    we with Petja.INST go.FUT.1PL home

<div align="right">(Russian; Vassilieva & Larson 2005: 101)</div>

Madeleine Butschety

    a. 'We will go home with Petja.'

    b. 'I and Petja will go home.'

(2)   *pro* S   Mariju   smo   otišli u muzej.   (Torlakian BCMS)
        with Maria.INST AUX.1PL went in museum

    a. 'We went to the museum with Maria.'

    b. 'Maria and I went to the museum.'

The availability of an apparent singular ('I') interpretation of the (dropped) pronoun under reading (b) is unexpected given that its surface form is plural.[1] This reading (henceforth: iPPC; following Feldman's 2003 terminological distinction) contrasts with the default 'we' interpretation (henceforth: ePPC) of the plural pronoun under the reading paraphrased in (a).[2] In the light of the ambiguity between (a) and (b), it may seem tempting to assume that plural pronouns are ambiguous between a singular and a plural interpretation in general. As Vassilieva & Larson (2001) pointed out already, however, this cannot be the case. Constructions like (3) not involving a comitative phrase do not give rise to the ambiguity observed for the otherwise parallel example (1).[3]

(3)   My pojdëm   domoj.   (Russian; Vassilieva & Larson 2001: 449)
        we go.FUT.1PL home

    a. 'We will go home.'

    b. *'I will go home.'

Common analyses have usually taken one of two explanations: either that the reference of a plural pronoun in a PPC is composed of its singular counterpart

---

[1]To talk about the pronoun's "surface form" is actually a bit misleading here. Example (2) from Torlakian BCMS involves *pro*-drop. So, strictly speaking, there is no such thing as a surface form of the plural pronoun in this sentence. Although this example is also felicitous in its version with an overt plural pronoun *mi* 'we', I omit it here and in the succeeding examples for reasons that will be discussed in §4. Furthermore, we can infer the "underlying" form, or rather the features of *pro* from the features on the verb – which is 1st person plural in both readings of (2).

[2]The distinction between "iPPC" and "ePPC" as made in Feldman (2003) refers to an "inclusive" and an "exclusive" interpretation of the plural pronoun, respectively. That is, under an "inclusive" interpretation of a PPC, the referent from the comitative phrase (e.g. *Petja* in (1)) is apparently included in the overall reference of the plural pronoun; whereas under an "exclusive" interpretation, the reference of the plural pronoun does not include the referent from the comitative phrase – i.e., under the ePPC reading of (1), the 1st person plural pronoun refers to the speaker and someone else (but Petja).

[3]The original transliteration of example (3) was altered to be in line with the scientific transliteration of Cyrillic.

and the referent of the comitative (cf. Vassilieva & Larson 2001, Vassilieva & Larson 2005), or that PPCs involve asymmetric coordination from a syntactic perspective, but have the same (or similar) semantics as symmetric coordination (cf. Dyła 1988, McNally 1993, Feldman & Dyła 2008). I will spell out the core assumptions of these and other approaches in §2.

A real challenge for pretty much any theory of PPCs (apart from apposition-based ones such as Cable 2017, perhaps) is that the plural pronoun and the comitative phrase can occur as a discontinuous constituent. And yet, they give rise to iPPC interpretations (at least in most Slavic languages). Consider the contrast between (4) and (5) from Torlakian, which is mainly syntactic but not semantic.

(4) *pro* S   Mariju   smo   otišli u muzej.   (Torlakian BCMS)
        with Maria.INST AUX.1PL went in museum

    a. 'We went to the museum with Maria.'   ePPC

    b. 'Maria and I went to the museum.'   iPPC

(5) *pro* Otišli smo   u muzej   s   Mariju.   (Torlakian BCMS)
        went AUX.1PL in museum with Maria.INST

    a. 'We went to the museum with Maria.'   ePPC

    b. 'Maria and I went to the museum.'   iPPC

To the best of my knowledge, no analysis has yet come up with a neat answer to the question of how an iPPC reading can arise for such "split" PPCs like (5). My attempt to explain this fact will be outlined in §3.2.1. In particular, I propose a novel analysis of PPCs in Slavic languages that is based on the assumption that plural pronouns and (universal) quantifiers behave alike in various respects. In particular, I argue that the difference between an iPPC and an ePPC interpretation can be derived in terms of whether the comitative phrase resides inside the restrictor of the plural pronoun (which is similar to a quantifier's restrictor) or not. My analysis is essentially based on data from Torlakian BCMS and from Bulgarian. These two languages are very suitable as a starting point, as they have different properties regarding available readings for split PPCs – and from these differences, I derive one of the core assumptions of my analysis.[4] Moreover, to the best of my knowledge, nothing has been said in the literature yet about the behavior of PPCs in Torlakian BCMS and Bulgarian.

This article proceeds as follows. In §2 I briefly summarize the main findings and claims of previous approaches to PPCs, and point out a few problems that

---

[4]However, I also occasionally show how (and to what extent) the generalizations derived from those distinct features are applicable to other Slavic languages such as Russian or Polish.

arise from each. I present my proposal in §3. First, in §3.1, I present my analysis of the internal structure of plural pronouns in general. I claim that plural pronouns, just as quantifiers, select a restrictor argument and that it is the precise internal structure of this restrictor which determines the overall referential properties of the respective plural pronoun. In §3.2 I then show the implications of those assumptions for an analysis of PPCs. Specifically, I argue that the comitative phrase occurs inside of the plural pronoun's restrictor in iPPCs, but outside of it (that is, elsewhere in the syntactic structure) in ePPCs. I outline my explanation of why split PPCs can have iPPC interpretations in some Slavic languages in §3.2.1. Furthermore, I also offer an explanation as to why we find iPPC readings for split PPCs in precisely those Slavic languages in which we find them (such as Torlakian BCMS), but not in others (such as Bulgarian). In §3.2.2 and §3.2.3 I present data in favour of my analysis related to Subject Control constructions and binding. However, I want to point out in advance that the data from those two sections could probably also be correctly derived under other approaches to PPCs. Nevertheless, it is my intention to discuss them rather than to leave them as implicit evidence. So in the end, the virtue of my analysis is that it can properly predict when, how, and why iPPC readings for split PPCs arise – and thus, that it fills an explanatory gap that exists among previous analyses of PPCs.

2 Previous analyses of PPCs in Slavic

Existing approaches to PPCs can basically be divided into three categories: those assuming that PPCs have an underlying coordinative structure; those which treat PPCs as a kind of appositive construal involving ellipsis; and finally, those which are based on the idea that the comitative phrase acts as a complement of the plural pronoun in such constructions. In this section, I briefly illustrate the main claims of previous analyses as well as some issues that remain open in the light of the respective theories. My own analysis draws on the assumptions of analyses of the third category, i.e. on approaches which treat the comitative phrase as a complement of the plural pronoun. However, my proposal takes one step further in observing and implementing some parallels between plural pronouns and quantifiers.

2.1 Based on coordination

As has been observed in the literature on (primarily) Russian and Polish, "ordinary" or "regular" comitative constructions (i.e. those which are not headed by

a plural pronoun) behave differently with regard to whether they trigger singular or plural agreement on the verb; see example (6). Traditional analyses (cf. Dyła 1988, McNally 1993, Feldman & Dyła 2008) anchor the alternation of verbal number agreement in diverging underlying syntactic structures. Specifically, those comitatives that trigger plural agreement on the verb are considered to be coordinative (that is, conjunctive) construals, whereas comitatives that trigger singular agreement on the verb are commonly treated as adjuncts to VP.[5]

(6) a. Maša   s   Dašej   xodjat v školu.
        Maša.NOM with Daša.INST go.1PL to school
        'Maša and Daša go to school.'   *s*-CONJUNCTION

     b. Anja   s   Vanej   pošla   v biblioteku.
        Anja.NOM with Vanja.INST went.SG.F to library
        'Anja went to the library with Vanja.'   *s*-ADJUNCTION
                                                 (Russian; Feldman 2003)

I do not want to enter this debate to any extent and I will have little to say about such "regular" comitatives in this article. My account is based on the specific properties of plural pronouns and does not intend to make generalizations to any other kinds of DPs. Nonetheless, both structural options from (6) have been considered in the light of PPCs – although coordination-based accounts (such as the influential paper by McNally 1993) are more numerous than analyses that assume that PPCs have an adjunction structure (see Ladusaw 1989).

A recent example of a coordinative treatment of PPCs is Sokolová's (2019) analysis. It is claimed there that Slovak PPCs classify as "coordinate comitatives", i.e. the plural pronoun DP and the comitative phrase are assumed to have the same structural rank. Under Sokolová's (2019) approach, the denotation of a plural pronoun such as *my* 'we' in (7) consists of two referents (or "participants", as dubbed in the original article) – namely the speaker (= first participant), and a second participant that gets lexically specified by the referent of the comitative.

(7) My s   Evou   chodíme do rovnakej školy.
     we with Eva.INST go.1PL to same   school
     'Eva and I go to the same school.'   (Slovak; Sokolová 2019: 101)

The iPPC interpretation of a sentence like (7) is assumed to arise on the basis of an "absorption" mechanism. The apparent double occurrence of the second participant conflates into one via the overall reference of the plural pronoun. Sokolová

---

[5]Transliterations were adjusted here.

Madeleine Butschety

(2019) is not explicit about how this "absorption of a referent" (a notion based on Daniel 2000) is supposed to work out on a derivational level, however. Likewise is Daniel's (2000) account not concerned with a precise formal underpinning. The original idea, though, is that plural pronouns are only explicit about the first referent in general (such as "speaker" or "addressee" in the case of 'we' and 'you$_{PL}$', respectively). Under some discourse conditions, an explication of the second participant, which otherwise remains implicit, is required. This explication can be resolved via different means in (morpho-)syntax; and so in the end, PPCs are just one (morpho-)syntactic conventionalization to feed these pragmatic needs for specification of the other referent of the plural pronoun (Mikhail Daniel, p.c.).[6]

Analyses such as McNally's (1993), on the other hand, assume an asymmetric kind of coordination. McNally (1993) does not deal with PPCs in particular, but suggests that PPCs such as (8a) and "regular" comitative constructions like (8b) share the same underlying structure illustrated in Figure 1.

(8) a. Oni   s   Petej   pridut.
         they.NOM with Petja.INST come.3PL
         'He and Petja are coming.   (Russian; McNally 1993: 359)

    b. Anna   s   Petej   napisali pis'mo.
         Anna.NOM with Petja.INST wrote.3PL letter
         'Anna and Petja wrote a letter.'   (Russian; McNally 1993: 347)
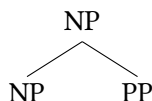
Figure 1: Asymmetric coordination structure (McNally 1993: 359)

McNally (1993) claims that from a semantic point of view, the structure in Figure 1 has the same interpretation as a structure involving symmetric coordination.

---

[6]Note that an explanation along these lines, once incorporated into a formal discourse framework, would probably suffice the purpose of predicting when 2p-interpretations arise – namely if and only if the identity of the second participant otherwise remains unclear. Evidence in favor of this view comes from infelicitous sequences such as (i).

(i) Nie otidohme v muzeja. #Večerta nie s   Peter gledahme   star Disni film.
     we went.1PL in museum evening we with Peter watched.1PL old Disney movie
     'We$_i$ went to the museum. In the evening, #[I and Peter]$_i$ watched an old disney movie.'
                                                                        (Bulgarian)

Several problems arise from a coordination-based approach to PPCs. A salient challenge has to do with the fact that the plural pronoun can be dropped in a PPC (in Slavic languages that permit *pro*-drop, such as Torlakian BCMS, Bulgarian, Polish, or Slovak), but not in other (regular) forms of conjunction. But if we assume that a PPC is (structurally speaking) a conjunction of the plural pronoun and the comitative phrase, then we would expect that PPCs and regular coordination show the same syntactic pattern – contrary to the actual facts. Dropping the first conjunct (even if pronominal) in regular coordination is impossible, compare (9a) and (9b) from Polish.

(9)   a.   On/*pro* z   bratem poszli do kina.
         he   with brother went   to   cinema
         'He and his brother went to the cinema.' (Polish; Trawinski 2005: 385)

     b.   On/*pro i   Maria poszli do kina.
         he   and Maria went   to   cinema
         (*Intended*): 'He and Maria went to the cinema.'

                                        (Polish; Trawinski 2005: 384)

Moreover, it remains puzzling why the order of plural pronoun and the other DP involved is fixed in PPCs; i.e. why a construction like (10a) is ungrammatical. Because, as can be seen from (10b), such a reversed order is (although slightly deviant) sometimes possible in coordinated structures.

(10)   a.   * Petja   s   nami   tancevali.
         Petja.NOM with us.INST danced.PL
         Intended: 'Petja and us danced.'

     b.   ? Petja   i   ty   tancevali.
         Petja.NOM and you.NOM danced.PL
         'Petja and you danced.'   (Russian; Vassilieva & Larson 2005: 114)

Under the assumption that PPCs have an underlying structure like the one in Figure 1, we could potentially explain these differences to regular conjunction structures in terms of asymmetric coordination. But to the best of my knowledge, no such account has been presented yet, and thus, the exact reasons for those differences remain to be spelled out in detail.

2.2 Based on apposition

This line of approach originates from Cable (2017). The basic claim of Cable's (2017) analysis is that PPCs are appositive constructions involving ellipsis. In

particular, he assumes that PPCs in their iPPC versions contain an elided instance of the respective plural pronoun's singular counterpart; such that a sentence like (11a) has the underlying structure in (11b).

(11) a. My   s   Petej   pojdëm   domoj.
         we.NOM with Petja.INST go.FUT.1PL home
         'Petja and I will go home.'   (Russian; Vassilieva & Larson 2005: 101)

     b. My   [⟨ja⟩ s   Petej]   pojdëm   domoj.
         we.NOM I.NOM with Petja.INST go.FUT.1PL home
         Approx.: 'We, I and Petja, will go home.'   (Russian; Cable 2017: 8)

A plural pronoun in a PPC is taken to denote the sum of the two referential expressions occurring in the appositive, i.e. the sum of *ja* ('I') and *Petja* in (11b).

While such an account seems appealing, it is not without complications. First of all, it raises the question of what licenses ellipsis in the apposition. It cannot be deletion under identity in a strict sense (see Lipták 2015 for an overview, and the references therein), because according to Cable's (2017) proposal, the plural pronoun denotes the sum of the elided element and the comitative referent. Hence, if anything, deletion should target both *ja* and *Petej* in (11b); and then, there would be no overt material left in the appositive apart from the comitative element *s* itself. At best, we could assume that only *ja* gets elided because ellipsis would have to target a discontinuous constituent otherwise, or because it bears the same grammatical case as the plural pronoun (whereas *Petej* bears instrumental case, assigned by the comitative element). It is hard to figure out which deletion mechanisms are supposed to be at stake in a structure like (11b) – especially as long as the assumptions concerning the appositive structure are not embedded in any tradition of analyzing appositions. However, Cable (2017) seems to be aware of this problem and mentions that it could also be the case that the comitative element itself has a meaning akin to appositive structure. How this meaning would need to be defined in particular is left as an open issue.

Furthermore, it seems that phonological deletion poses a problem to any apposition-based approach to (Slavic, at least) PPCs in general. Because as varied as theories of appositives may be, there is a syntactic property which appositions are uncontroversially assumed to have. Namely, that there is an anchor expression to which the appositive attaches. The anchor expression of the apposition ⟨*ja*⟩ *s Petej* from (11b) would (also quite undisputedly) have to be *my*, i.e. the plural pronoun. While this is unproblematic for the Russian case at hand, we run into complications as soon as we want to apply this analysis to Slavic languages that allow *pro*-drop structures. Take example (12) from Torlakian BCMS, for instance.

The version in (12a) with the overt plural pronoun could be treated along the lines of (11b), as shown in (12b). But for its *pro*-dropped version in (13a), this does not work out. Because we would have to assume that appositions can attach to silent anchor expressions, as pictured in (13b).[7]

(12) a. Juče   smo   mi   s   Mariju   otišli u bioskop.
         yesterday AUX.1PL we.NOM with Maria.INST went in cinema

      b. Juče   smo   mi   [⟨ja⟩   s   Mariju]   otišli u bioskop.
         yesterday AUX.1PL we.NOM I.NOM with Maria.INST went in cinema
         Approx.: 'Yesterday we, I and Maria, went to the cinema.'

                                                                    (Torlakian BCMS)

(13) a. Juče   smo   *pro* s   Mariju   otišli u bioskop.
         yesterday AUX.1PL   with Maria.INST went in cinema

      b. # Juče   smo   *pro* [⟨ja⟩   s   Mariju]   otišli u bioskop.
           yesterday AUX.1PL   I.NOM with Maria.INST went in cinema
         Intended: 'Yesterday we, I and Maria, went to the cinema.'

                                                                    (Torlakian BCMS)

If it were indeed possible to adjoin an apposition or any kind of additional syntactic material (such as a PP, for instance) to an attachment site that is not overtly present in the syntactic structure we would expect to find constructions, such as (14b) or (14d), regularly. But as we can see, this expectation is not met.

(14) a. Mi,   (naime) ja   s   Petra,   idemo sutra   u bioskop.
         we.NOM namely I.NOM with Peter.INST go.1PL tomorrow in cinema
         'We, namely I and Peter, will go to the cinema tomorrow.'

      b. # *pro* Naime ja   s   Petra,   idemo sutra   u bioskop.
              namely I.NOM with Peter.INST go.1PL tomorrow in cinema
         Intended: 'We, namely I and Peter, will go to the cinema tomorrow.'

      c. Ona s   plavom kosom ide   sutra   u bioskop.
         she with blond hair go.3SG tomorrow in cinema
         'She with the blond hair will go to the cinema tomorrow.'

      d. # *pro* S   plavom kosom ide   sutra   u bioskop.
              with blond hair go.3SG tomorrow in cinema
         'She will go to the cinema tomorrow with the blond hair.'
         Intended: 'She with the blond hair will go to the cinema tomorrow.'

                                                                    (Torlakian BCMS)

---

[7]The hashtag # in (13b) is intended to indicate that it should be regarded with suspicion whether such a syntactic configuration is even possible.

The sentence in (14b) is ungrammatical in an out-of-the-blue context as well as in any syntactic context that does not contain a salient anchor or anaphoric expression. More precisely, the only syntactic environment which could save (14b) is a preceding sentence like *We/The best of friends will go to the cinema – naime ja s Petra....* Example (14d) on the other hand is simply infelicitous. It cannot mean what it is supposed to, i.e. that he/she and the blond-haired he/she will go to the cinema tomorrow. Instead, the sentence only has a rather awkward interpretation according to which she (an individual salient from the discourse) will go to the cinema tomorrow taking the blond hair (literally) with her.

Abstracting from (im)possible anchor expressions, an apposition-based analysis makes unwelcome predictions regarding the spell-out of an iPPC. Appositives are usually articulated with an intonational/phonological break or boundary. Thus, if for instance (11a) has the underlying structure in (11b), such a break or boundary should reflect in the pronunciation of such a sentence; see (15), where "(–)" signals the points where the breaks should occur.

(15)   My (–) s Petej (–) pojdem domoj.

This prediction is not borne out, since a PPC (no matter which reading is intended) does not surface any kind of apposition-typical intonational breaks in an unmarked context. It rather has a regular pronunciation.

2.3 Comitatives as pronoun complements

The analyses in Vassilieva & Larson (2001) and Vassilieva & Larson (2005) treat (Russian) plural pronouns as incomplete expressions, comprising a singular nucleus and an unsaturated element $\triangle$ in their meaning. These two components are elements of an ordered pair $\langle X, Y \rangle$. X is obligatorily taken by the plural pronoun's singular counterpart (in accordance with its person feature), and Y gets (per default) saturated by some σ from the context. The resulting (distributive) semantics of Russian *my* 'we', *vy* 'you.PL' and *oni* 'they' can be represented as in (16) below.[8]

---

[8]The formalism that Vassilieva & Larson (2001) and Vassilieva & Larson (2005) make use of was adopted from Larson & Segal (1995), where it is assumed that the truth value assigned to a sentence is dependent on context sequences σ – such that $\sigma(a), \sigma(b), \sigma(c)$ and $\sigma(d)$ are associated with speaker, addressee, speaker time and speaker location, respectively, from the context. Indexical pronouns are claimed to get their values via these means. Other instances of such a sequence, dubbed $\sigma(n)$ determine the reference of non-indexical pronouns on the other hand. Since my analysis does not make use of this particular formal framework, I will not go into this topic any further here.

(16)   a.   VAL($\langle X, Y\rangle$, [$_\mathrm{D}$ *my*], σ) iff $|(\{\sigma(a)\} \cup Y) - X| = 0$
        '(all of) speaker + others Y'

     b.   VAL($\langle X, Y\rangle$, [$_\mathrm{D}$ *vy*], σ) iff $|(\{\sigma(b)\} \cup Y) - X| = 0$
        '(all of) addressee + others Y'

     c.   VAL($\langle X, Y\rangle$, [$_\mathrm{D}$ *oni*], σ) iff $|(\{\sigma(i)\} \cup Y) - X| = 0$
        '(all of) he/she/it + others Y'

<div align="right">(Vassilieva & Larson 2005: 119)</div>

With regard to PPCs and their iPPC interpretations, the idea outlined in Vassilieva & Larson (2005) is that the comitative phrase occupies a complement position to the plural pronoun. The Y-slot in the meaning of the plural pronoun then gets filled by the referent of the comitative, i.e. as a matter of syntactic means. The underlying structure of the iPPC reading is given in Figure 2, and its semantics in (17).
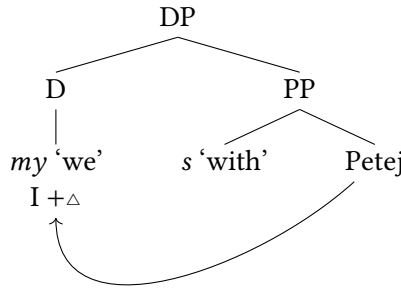
Figure 2: iPPC structure (Vassilieva & Larson 2005: 120)

(17)   VAL($\langle X$, *my s Petej*, $\sigma\rangle$ iff $|(\{\sigma(a)\} \cup \{\text{Petja}\}) - X| = 0$; i.e.
      VAL($\langle X$, *my s Petej*, $\sigma\rangle$ iff $|(\{\sigma(a), \text{Petja}\}) - X| = 0$
      '(all of) speaker + Petja'   (Vassilieva & Larson 2005: 120)

Hence, Vassilieva & Larson tie the availability of an iPPC interpretation to the specific syntactic configuration depicted in Figure 2. However, there are two pieces of data that they explicitly leave unaccounted for. Firstly, it remains unclear under this analysis why PPCs (in their iPPC meaning) can occur split, i.e. why the plural pronoun and the comitative phrase can be discontinuous and still, an iPPC interpretation is available; see (18a). Secondly, it is puzzling why *wh*-questions such as (18b) do actually not give rise to an iPPC meaning.

(18) a. My   pojdëm   zavtra   s   Ivanom v magazin i   vsë
             we.NOM go.FUT.1PL tomorrow with Ivan.INST to store   and all
             kupim.
             buy.FUT
             'Tomorrow, we will go to the store with Ivan and buy all we need.'
                                                         ePPC
             'Tomorrow, Ivan and I will go to the store and buy all we need.' iPPC

      b.   S   kem   my xodili v magazin?
            with whom we went to store
            'With whom did we go to the store?'   ePPC
            Unavailable iPPC reading: 'I and who went to the store?'

                                    (Russian; Vassilieva & Larson 2005: 122)

If the comitative phrase acts as a complement of the plural pronoun, we do not expect (18a) to give rise to an iPPC reading, because the plural pronoun and the comitative phrase do not form a constituent in this sentence – yet, an iPPC interpretation is available. This poses a serious challenge to Vassilieva & Larson's (2005) analysis since the availability of an iPPC reading for split PPCs is widespread across Slavic languages (with some exceptions). My analysis, to which we turn next, offers an explanation of these facts as well as a novel perspective on PPCs in general.

3 Proposal

In this section, I present my approach to PPCs in Slavic. My analysis is based on observations regarding similarities in the semantic and syntactic behaviour of plural pronouns and quantifiers; or rather, of plural pronominal DPs and quantificational (noun) phrases (henceforth: QPs).[9] Therefore I first outline my claims regarding the internal structure of plural pronouns in §3.1 and show how it is similar to the internal structure of QPs. On this basis, I derive the two readings of PPCs (i.e. iPPC and ePPC) in §3.2 and then illustrate in §3.2.1 why and how split PPCs give rise to iPPC interpretations in some Slavic languages, but not in

---

[9]It is important to note here that I do not claim that plural pronouns are quantificational elements (in a broad sense) by definition. While these two types of expressions exhibit some striking parallels, my analysis is not intended to treat them as completely analogous. That is, just because a (universal) quantifier has such and such properties or shows such and such behaviour (of a syntactic or semantic nature), this does not automatically also have to apply to plural pronouns. For the time being, I restrict the similarities between plural pronouns and quantifiers to what is explicitly diagnosed in the course of my analysis outlined in this article.

others. Data related to Subject Control structures and binding are presented in §3.2.2 and §3.2.3, respectively.

3.1 Restrictor sets for plural pronouns

Plural pronouns have often been treated analogously to definite descriptions in that both denote pluralities of individuals (cf. Link 1983, Nunberg 1993, Elbourne 2008, Büring 2011, among many others). If we leave figurative uses aside, then this plurality obligatorily includes the speaker (1st person plural pronoun) or addressee (2nd person plural pronoun) of the utterance. In my analysis, I follow Link (1983) and others in assuming that plural pronouns denote pluralities of individuals. But the suggested way in which this plurality gets composed differs from previous approaches. I propose a slightly modified variant of the "plural pronouns as definite descriptions" view. In particular, I claim that plural pronouns are more similar to quantifiers after all. What I am arguing for in this and subsequent sections is that plural pronouns and (universal) quantifiers have some striking properties in common – syntactic, semantic, and pragmatic ones.

Let us start with a very basic yet deep parallel regarding the syntactic arguments these expressions take. Quantifiers are uncontroversially assumed to take two arguments: restrictor and scope. The former typically corresponds to the NP the quantifier forms a QP with. The latter typically corresponds to the predicate (i.e. the VP, roughly speaking). While it is probably also quite uncontroversial that a plural pronoun (in subject position) combines with a predicate (a scope argument so to speak), I claim here that the internal syntactic structure and semantic composition of such an expression also involves a restrictor argument. More specifically, the restrictor of a plural pronoun is argued to be the decisive factor in determining the plural pronoun's overall reference. So, what must the restrictor of a plural pronoun look like by analogy with the restrictor of a quantificational expression? And what semantic or pragmatic properties can we find?

There has been a vast debate in the literature on quantifiers concerning the question whether quantificational determiners presuppose that their restrictor sets must not be empty (see Heim & Kratzer 1998, and Szabolcsi 2010 for an overview). By now, it is more or less commonly agreed that at least strong quantifiers such as *all, most*, or *each* presuppose non-emptiness of their restrictor's denotation. So if, as claimed here, requiring a restrictor argument is one of the properties that plural pronouns share with quantifiers, we would probably expect plural pronouns to trigger an equivalent presupposition regarding the denotation of their restrictor sets. While there seems to be nothing wrong in principle at first glance with the assumption that a plural pronoun like *we* presupposes that what

Madeleine Butschety

it ranges over must not be the empty set, such a bare existential presupposition does not suffice under closer inspection. The reason is that non-emptiness alone does not account for referential properties related to a plural pronoun's person feature. Specifically, a felicitous use of *we* not only requires that the pronoun ranges over a plurality made up of whatever individuals, but rather over a plurality which obligatorily contains the speaker of the utterance. I will first and foremost tie these requirements directly to the lexico-semantic properties of plural pronouns.[10]

The lexical entries of *we* and *you*$_\text{PL}$ in (19) and those expressions' underlying structures in Figures 3–5 further illustrate the assumed parallel to a quantifier's restrictor.[11]

(19)   a. $[\![we]\!], [\![you_\text{PL}]\!] = \lambda P.\lambda Q.P \subseteq Q$

   b. $[\![\text{OP}_\text{U}]\!] = \lambda x.\lambda y.\lambda z.z \le x \vee z = y$ such that $x \ne y$
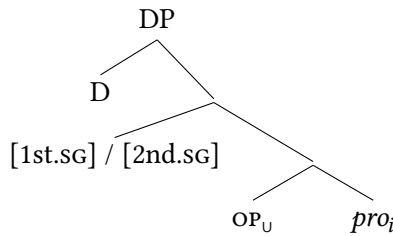
Figure 3: Underlying syntactic structure of 1st/2nd person plural pronoun

---

[10]So far, nothing hinges on this decision. I will not discuss any data that is concerned with presupposition projection or presupposition failure. Note, though, that presuppositional content of *we* that matches the conditions stated above has occasionally been suggested in the literature (see Stokke 2022 and the references therein).

[11]I concentrate on 1st and 2nd person plural pronouns only throughout this article. The reason for this restriction is twofold. On the one hand, 3rd person plural pronouns do not have a fixed first referent like 1st and 2nd person plural pronouns do (i.e. speaker or addressee, respectively). On the other hand, iPPC interpretations are often harder to obtain for PPCs in the Slavic languages I investigate, if they are available at all. And this, in turn, might be related to the lack of a fixed first referent. In particular, I found that iPPC readings are often judged infelicitous in out-of-the-blue contexts – and only in suitable contexts or in follow-up sentences to an explicit QUD highlighting which individual is intended to be the first referent, iPPC interpretations seemed more readily available.

$$\lambda Q.(\lambda z.z \leq g(i) \vee \text{SPKR}(z) \wedge \text{ATOMIC}(z)) \subseteq Q$$

we

$\langle I \rangle$

SPKR

OP_U $g(i)$
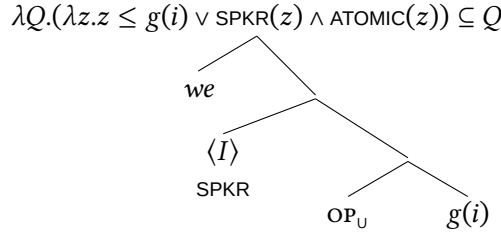
Figure 4: Semantic interpretation 1st person plural pronoun

$$\lambda Q.(\lambda z.z \leq g(i) \vee \text{ADDR}(z) \wedge \text{ATOMIC}(z)) \subseteq Q$$

you$_\text{PL}$

$\langle you_\text{SG} \rangle$

ADDR

OP_U $g(i)$
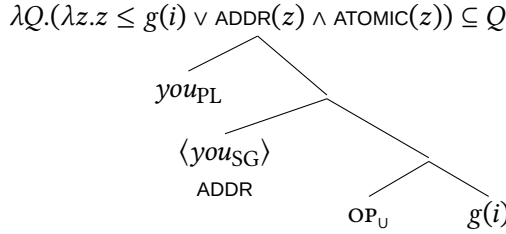
Figure 5: Semantic interpretation of 2nd person plural pronoun

The restrictor argument of a plural pronoun thus consists of three components: A silent instance of SPKR (i.e. reference to the speaker of the utterance) or ADDR (reference to the addressee of the utterance), a silent operator OP_U, and a contextual assignment function $g(i)$ which surfaces as a silent pronominal form *pro*$_i$ (interpreted as $[\![pro_i]\!]^g$, i.e. $g(i)$ again) in the structure. The central idea is that the two referential instances within the restrictor are conflated by OP_U. Roughly speaking, the mechanism of OP_U can be seen as set union. OP_U takes two arguments x, y to form a z such that z is identical to y (i.e. the speaker SPKR or addressee ADDR) or z is less or equal to x (i.e. $g(i)$). Note that being less or equal to $g(i)$ will come out as plain $g(i)$ in any case. Actually, OP_U's λz-part is just a formal workaround. Given how this function is defined it acts more or less like a type-shifter in that it takes two type e expressions as input and forming a set containing precisely and only those individuals. So what the 1st and 2nd plural pronouns refer to according to (19) can be paraphrased as 'the speaker and other(s)' and 'the addressee and other(s)', respectively.

3.2 Restrictors in PPCs

In the previous section, I suggested that plural pronouns have a (syntactically speaking) restrictor argument (and semantically speaking, introduce a restrictor set), and that this restrictor (or restrictor set) determines the plural pronoun's

Madeleine Butschety

overall reference. The main hypothesis of my proposal regarding PPCs is that the difference between an iPPC and an ePPC interpretation boils down to whether the comitative phrase resides inside the restrictor of the plural pronoun or not – in the former case, an iPPC reading arises, whereas in the latter case, only an ePPC reading should be available. In particular, the structure of a PPC such as Russian (20) under its iPPC meaning is as pictured in Figure 6.

The crucial aspect in Figure 6 is that the comitative element *s* 'with' is assumed to have the very same semantics as OP_U and that the referent of the comitative phrase occupies the position of the pronominal element interpreted as $g(i)$ in the default plural pronominal structure. That is to say that the comitative phrase *s Petej* 'with Petja.INST' as a whole acts as a spell-out of the more general OP_U plus $g(i)$-part from Figure 3 – and consequently, *my* in the structural configuration (6) refers to just the speaker and Petja. Put differently, the comitative element serves the same purpose as OP_U here, i.e. it forms a set from its two arguments, namely, the speaker and Petja.

(20) My   s   Petej   …
      we.NOM with Petja.INST
      Intended (iPPC): 'I and Petja …'

where $[\![s]\!] = [\![\text{OP}_\text{U}]\!]$

$\lambda Q.(\lambda z.z \leq \text{Petej} \vee \text{SPKR}(z) \wedge \text{ATOMIC}(z)) \subseteq Q$ …

my

⟨I⟩
SPKR

s   *Petej*
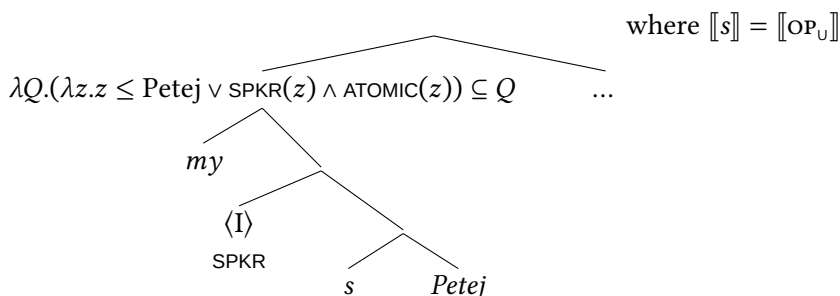
Figure 6: Semantic interpretation of iPPC structure

However, how to derive the ePPC interpretation of (20) is not entirely straightforward. Note that we cannot simply adjoin the comitative phrase to the (fully determined) plural pronominal DP – because on the one hand, it would not be straightforward how to integrate the comitative phrase into the scope argument of the plural pronoun. On the other hand, we might run into incorrect predictions after all. The data that will be presented in the remainder of this section strongly suggests that the comitative phrase is actually rather an adjunct to VP in the ePPC cases. I thus claim at this point that we are dealing with adjunction

to VP, and refer to the forthcoming sections for further evidence in favour of this view.

The predictions of this theory should now be straightforward: if and only if the comitative phrase occurs inside the restrictor of a plural pronoun, an iPPC interpretation arises. On the other hand, if the comitative phrase occupies a position adjacent to the VP, only an ePPC reading is predicted to be available. We turn to data from Torlakian BCMS and Bulgarian which support the suggested "quantificational" treatment of plural pronouns next. This choice regarding investigated languages is not due to arbitrary reasons. As will soon become clear, these two languages have very illustrative distinct properties when it comes to PPCs. Moreover, nothing has been said in the literature yet about PPCs in Bulgarian or Torlakian BCMS to the best of my knowledge – and hence, they are definitely worth a closer inspection.

To start with, we find a direct syntactic-semantic reflection of the predictions just stated in Bulgarian. In particular, a structure like (21a) in which the plural pronoun and the comitative phrase are "tied together" has an iPPC reading only. In contrast, a split PPC like (21b) exclusively gives rise to an ePPC interpretation.

(21) a. Nie/*pro* s   Peter otidohme v muzeja.   (Bulgarian)
         we.NOM with Peter went.1PL   in museum.DEF
         Unavailable ePPC reading: 'We went to the museum with Peter.'
         'Peter and I went to the museum.'   iPPC

      b. Nie/*pro* otidohme v muzeja s   Peter.
         we.NOM went.1PL in museum with Peter
         'We went to the museum with Peter.'   ePPC
         Unavailable iPPC reading: 'Peter and I went to the museum.'

In accordance with the comitative phrase's structural proximity to the plural pronoun, we could argue that *s Peter* occurs within the plural pronoun's restrictor in (21a), but not in (21b). But while Bulgarian exhibits a structure-meaning correspondence that perfectly matches the predictions of my theory, it happens to be the case that not all Slavic languages (and we might in fact be talking about a minority here) behave like Bulgarian in this respect. In Torlakian BCMS, for instance, both sentences from (22) are ambiguous between an iPPC and an ePPC interpretation. That is, no matter whether the plural pronoun and the comitative phrase appear structurally very close to one another or as separate constituents, both readings are available.

(22) a. *pro* s   Mariju   smo   otišli u muzej.   (Torlakian BCMS)
           with Maria.INST AUX.1PL went in museum
           'We went to the museum with Maria.'   ePPC
           'Maria and I went to the museum.'   iPPC

      b. *pro* otišli smo   u muzej   s   Mariju.
           went AUX.1PL in museum with Maria.INST
           'We went to the museum with Maria.'   ePPC
           'Maria and I went to the museum.'   iPPC

The crucial question is how to account for the iPPC interpretation of (22b) under the analysis proposed in this article. Because note that (22a)'s ePPC reading can be easily explained via the assumption that *s Mariju* is an adjunct to any suitable XP from the sentence except inside the plural pronominal DP, or directly attached to it.

3.2.1 Together vs. apart: Split PPCs as floated constructions

If a plural pronominal DP is indeed similar to a QP, then we would expect that these two phrasal expressions do not only share striking semantic properties, but also syntactic ones – such as the possibility to detach restrictor and head: Quantifier-floating (henceforth: Q-floating; for a general discussion see Sportiche 1988, Shlonsky 1991, Merchant 1996, Bošković 2004, Fitzpatrick 2006, among many others).[12] I argue here that a similar movement operation is the source of (22b)'s iPPC interpretation, i.e. that the plural pronoun and the comitative phrase were base-generated within the same DP, but got detached via syntactic movement. But note that I remain intentionally vague in saying only that it is a "similar" movement operation – because strictly speaking, the floated element would be

---

[12]At this point, one could object that this is not an accurate description of Q-floating. In some approaches, for example, it is assumed that a quantifier is not a syntactic head in its own right, but merely occupies the specifier position of the restrictor NP/DP. Yet other accounts suggest that quantifiers might be simply adjoined to their restrictor arguments. To enter such a debate is beyond the scope of this article and must be postponed to another occasion. What I assume, though, is a kind of generalized structure for which the crucial point is this: there is a syntactic element (i.e. the quantifier or the plural pronoun) which selects, or at least combines with a restrictor argument. These two expressions can be separated or dislocated from each other via a syntactic movement operation. Note that this is also to say that I assume with Sportiche (1988), Shlonsky (1991), Merchant (1996), or Bošković (2004) Q-floating to involve syntactic movement. There, too, is an ongoing debate in the literature concerning the question whether this is indeed the case, or whether the quantifier acts more like an adverbial in such constructions (see for instance Bobaljik 1998), or whether both are lively options (see especially Fitzpatrick 2006 for a detailed cross-linguistic investigation).

the restrictor in the case of split PPCs, not the head (as in Q-floating structures). I will set this issue aside for the time being and briefly return to it in §4.

However, notably, an explanation along the lines of an analogy to Q-floating cannot be applied to Bulgarian split PPCs such as (21b) from above since Bulgarian does not have "true" Q-floating structures.[13] This is shown by the ungrammatical transformation of (23a) into Q-floated (23b). What might look like a Q-floating structure in (23c) at first glance should rather be considered as topicalization since the clitic would not occur in the non-topicalized structure.

(23)   a.   Vsički bademi sa na masata.
            all   almonds are on table.DEF
            'All almonds are on the table.'

      b.   * Bademite   sa vsički na masata.
             almonds.DEF are all   on table.DEF
             Intended: 'The almonds are all on the table.'

      c.   Knigite   gi   pročetoh vsičkite.
             books.DEF them.CL read.1SG all.DEF
             'As for the books, I read them all.' (*Intended*: 'The books I read all.')

                                    (Bulgarian; Vulchanova & Giusti 1995: 55)

In contrast, Q-floating structures exist in Torlakian BCMS, as illustrated in (24a-24b) – and moreover, there seems to be a remarkable correlation in general between the availability of iPPC readings for split PPCs and whether the respective language also permits Q-floating. Take Russian as a further example, where we find the very same pattern as in Torlakian BCMS: split PPCs can give rise to iPPC interpretations and Q-floating structures are possible; see (25a) and (25b)–(25c), respectively.[14]

(24)   a.   Svi bademi su   na sto.   (Torlakian BCMS)
            all almonds AUX.3PL on table
            'All almonds are on the table.'

---

[13]By "true", I mean Q-floating structures that do not make use of a resumptive pronoun or any other resumptive linguistic device.

[14]One could also mention Polish in this context, which seems to pattern more or less with Bulgarian. The picture is, however, not entirely clear. While it has been stated in Feldman & Dyła (2008) that the comitative phrase has to occur adjacent to the plural pronoun in order to obtain iPPC readings, my informants found iPPC interpretations somewhat available for split PPCs. Likewise, structures with floated *wszystkie* 'all' were occasionally (and given certain intonational circumstances) found acceptable. After all, it thus seems that Polish occupies an intermediate position.

Madeleine Butschety

    b. Bademi su   svi na sto.
       almonds AUX.3PL all on table
       'The almonds are all on the table.'

(25) a. My pojdëm   zavtra   s   Ivanom   v magazin.
       we go.FUT.1PL tomorrow with Ivan.INST to store
       'Tomorrow, we will go to the store with Ivan.'   ePPC
       'Tomorrow, Ivan and I will go to the store.'   iPPC

    b. Prišli   vse deti.
       came.PL all children
       'All the children came.'

    c. Deti   prišli   vse.
       children came.PL all
       'The children all came.'   (Russian; Fitzpatrick 2006: 144)

The parallel between split PPCs and Q-floating does not only account for the availability of iPPC readings for split PPCs, but also for the other piece of data that remained puzzling under Vassilieva & Larson's (2005) approach – namely the lack of an iPPC interpretation of *wh*-questions like (18b), repeated in (26) for convenience.

(26) S   kem   my xodili v magazin?
     with whom we went to store
     Unavailable iPPC interpretation: 'With whom did I go to the store?'
                         (Russian; Vassilieva & Larson 2005: 122)

In the terms of my account, a structure like (26) would have to involve the reverse kind of Q-floating movement operation; and carrying out such a movement does simply not yield grammatical results. We would thus expect it to be unavailable for QPs and plural pronominal DPs alike. Examples analogous to the Russian (26) do also not give rise to iPPC readings in Torlakian BCMS:

(27) S   koj ste   prekjuče   išli u prodavnicu?
     with who AUX.2PL the.day.before.yesterday gone in shop
     'With whom did you_PL go to the store the day before yesterday?'
     Unavailable: 'With whom did you_SG go to the store the day before
     yesterday?'   (Torlakian BCMS)

The same holds, unsurprisingly, of Bulgarian (28a). The only *wh*-question structure for which iPPC interpretations arise are echo questions like (28b), i.e. syntactic constructions in which the *wh*-element remains *in situ*.

(28) a. S   koj vie   ste   bili v muzeja?   (Bulgarian)
          with who you.PL are.2PL been in museum.DEF
          'With whom have you_PL been to the museum?'

  b. Vie   s   KOJ otidokhte v muzeja?
      you.PL with who went.2PL in museum.DEF
      'You_SG went to the museum with WHOM?'

Such cases are covered by the theory presented in this article, since there is
nothing in the structure of (28b) that would suggest that the comitative phrase
does not occur inside of the plural pronoun's restrictor – rather, the comitative-
internal DP just got replaced by a respective wh-phrase.

3.2.2 Subject Control

Further support for the view put forth here comes from Subject Control (hence-
forth: SC) constructions: Consider the contrast between (29a) and (29b) from
Torlakian BCMS with regard to (im)possible interpretations to start with.

(29) a. *pro* pokušali smo   juče   s   Mariju   da   PRO
                tried   AUX.1PL yesterday with Maria.INST COMP

      popravljamo krov.
      repair   roof
      'Yesterday, we tried to repair the roof with Maria.'   ePPC
      'Yesterday, Maria and I tried to repair the roof.'   iPPC

  b. *pro* pokušali smo   juče   da   PRO popravljamo krov s
            tried   AUX.1PL yesterday COMP   repair   roof with

      Mariju.
      Maria.INST
      'Yesterday, we tried to repair the roof with Maria.'   ePPC
      Unavailable iPPC interpretation: 'Yesterday, Maria and I tried to repair
      the roof.'

                                                        (Torlakian BCMS)

If, as in (29a), the comitative phrase occurs in the matrix, we get the standard
ambiguity between an iPPC and an ePPC reading. But if, on the other hand, the
comitative phrase occupies a syntactic position inside of the embedded complex
popravljamo krov 'repair the roof', as in (29b), an iPPC interpretation is not avail-
able. Given the observations from the previous section, i.e. that split PPCs usually
also give rise to iPPC interpretations in Torlakian BCMS, the question arises why
this reading is unavailable for (29b).

It is commonly assumed that PRO is a null pronoun that lacks any phonological content. Thus, PRO is treated as an empty category (see Chomsky 1981 for the origins of this notion). Other instances falling into this class are traces of moved phrases, including wh-movement traces, and (dropped) *pro*. But in contrast to *pro*, for instance, there is no kind of overt NP or DP that corresponds to PRO.[15] PRO is simply a silent anaphoric element that gets bound by an antecedent expression. To the best of my knowledge, no specific assumptions have been made in the literature regarding any sort of internal structure of PRO. Hence we should not go all wrong if we presume that there is none, i.e. if we take PRO just to be a syntactic subject placeholder element that has to occupy SpecTP in SC constructions due to the EPP (Extended Projection Principle).

So since PRO has no internal structure, the comitative phrase in (29b) cannot occur inside of it – because PRO has no such thing as what I dubbed a restrictor here. Consequently, *s Mariju* can only be part of the plural pronoun's scope argument in (29b). Given that we are dealing with an SC construction, no movement of the plural pronoun beyond *da* can be argued to be involved. Or, put differently, there is just no instance whatsoever (neither overt nor covert) of the plural pronoun inside of (29b)'s PRO-part such that we could assume the comitative phrase to be a leftover of.[16] On the one hand, we cannot claim that the comitative phrase was base-generated within PRO (for the reasons just mentioned) and that the two got separated via a syntactic movement operation akin to Q-floating. On the other hand, we also cannot claim that the (dropped) plural pronoun does, or at any point of the derivation did, appear inside of the *da*-complement of the matrix. So in sum, my analysis correctly rules out iPPC interpretations for sentences like (29b), but correctly predicts the availability of an iPPC reading for sentences such as (29a), where the PPC also occurs split, but the (dropped) plural pronoun and the comitative phrase are inside the same clause – and that is all the more important: Because if the movement operation behind split PPCs indeed resembles Q-floating, this is another reason why (29b) has no iPPC interpretation; even under an account to SC such as Hornstein's (1999). Q-floating is clause-bound (cf.

---

[15]Whereas for *pro*, the corresponding overt expression would be the pronoun with the respective number and person features, of course.

[16]It has been suggested in Hornstein (1999), however, that SC does indeed involve movement of the subject from the infinitival clause to a higher syntactic position in the matrix. That is, PRO, in Hornstein's (1999) sense, would rather be a trace than a silent anaphoric pronominal. Under such an analysis, the argument concerning (29b) just made above might seem a bit misguided. However, even if the movement-based approach to SC constructions was the more suitable solution (for a discussion, see Landau 2003), there is another issue standing in the way of *s Mariju* originating in the restrictor of the plural pronoun in (29b) – namely clause-boundedness of floated quantifiers. I return to this issue in more detail below.

Kayne 1981), i.e. the moved constituent cannot cross certain syntactic boundaries such as (full) clauses. In this sense, the complementizer *da*, or as we could also say, PRO constitutes such a crucial border: a PPC that appears (linearly speaking) behind it cannot, and cannot have been part of the plural pronoun's restrictor. Thus, only ePPC readings are available for structures such as (29b). But a PPC that occurs in the matrix, may it be split or not (as long as both parts reside inside the matrix) can give rise to iPPC interpretations.

The significance of these SC examples can easily be overlooked, and what they indicate can just as easily be misunderstood. So I recap here why those data are relevant and what they show. First of all, one should bear in mind that split PPCs give rise to iPPC readings in Torlakian BCMS – nevertheless, this interpretation is not available for (29b). And, naïvely speaking, there is no reason why this should be the case. Hence if iPPC readings arise for split PPCs in Torlakian BCMS, what is (29b)'s special trait that blocks this interpretation? Or, asked the other way around, what would have to be the case in order to obtain an iPPC interpretation for (29b)? Well, the comitative phrase (inside the *da-* or PRO-part of the sentence) would need to have been base-generated within the DP of the (dropped) plural pronoun that occurs in the matrix. Is it a feasible assumption that at any stage of the derivation, this was the case in (29b)? I argue here that it is not. If we stick to the more common treatment of SC constructions in terms of PRO, then neither has the plural pronominal DP ever occurred within the *da-*part of the sentence, nor can the comitative phrase occur inside the restrictor of PRO itself since there is no such thing. If we follow Hornstein (1999) instead in assuming that SC constructions actually do involve syntactic movement, then still, an iPPC interpretation of (29b) is ruled out under my account – but for a slightly different reason. Under this view, we would have to say that the comitative phrase can nonetheless not have been base-generated within the plural pronoun's DP because Q-floating is a clause-bound operation. And if split PPCs involve a similar kind of movement, *s Mariju* must have originated somewhere else in the structure but within the plural pronoun's restrictor.

Nevertheless, one could object that the decisive difference between (29b) and (29a) lies elsewhere in the structure anyway. Specifically, in the fact that PRO is an anaphoric element that needs a binder, i.e. an antecedent expression from which it gets its reference. Such a binder expression can only occur somewhere higher in the syntactic structure, that is, in the matrix part of the sentence. Now the comitative phrase occurs in (29a), but not in (29b) within the matrix. The consequence is that PRO can be bound in (29a), but not in (29b) by the PPC – may it have an underlying ePPC or an underlying iPPC structure in (29a). And then, there is nothing in the structure of (29b) that would even suggest that we

are dealing with a split iPPC. I would like to make two comments on such a line of reasoning to conclude this section with: First of all, if that was indeed the reason for the difference in terms of available interpretations between (29a) and (29b), it does no harm to the theory proposed here – *s Mariju* would end up being a part of the plural pronoun's scope argument anyway. Consequently, we correctly predict only an ePPC reading to be available for (29b). It is fair to admit, however, that other analyses of PPCs would probably make similar predictions.[17] But second, it has not yet been once and for all decided whether SC constructions actually involve a PRO, or whether Hornstein's (1999) approach involving movement is more practicable in the end – although one must of course say that there are many aspects about SC in favour of the PRO view. Should the latter option nevertheless be the case, my analysis has a suitable explanation prepared. Lastly, we must not forget that some analyses of PPCs face difficulties in explaining iPPC readings of split PPCs at all.

3.2.3 Binding data

My analysis in its current form makes the following predictions regarding the binding of non-independently referential expressions such as *-self* anaphors. If the plural pronoun's restrictor is made up of a silent instance of reference to the speaker as well as of a comitative phrase including a comitative referent (i.e. in the iPPC case), we expect that the two together can co-bind an anaphor. If, on the other hand, the plural pronoun's restrictor is fully determined by SPKR and *pro$_i$* (i.e. in the ePPC case), we expect that the comitative referent cannot participate in the binding of an anaphor. I use exclusively Russian data in this section to show that these predictions are borne out. The reason for this change in the languages discussed is the following: unlike Torlakian BCMS, Russian has the possessive SELF anaphor *svoj-*. This element can be bound by antecedents of various morpho-syntactic kinds without changing its surface form. That is to say that *svoj-* can, for instance, be bound by a referential expression in the singular, or by a referential

---

[17]For example, if an iPPC had an underlying coordinative structure, then one could argue that the comitative phrase cannot be coordinated with PRO; no matter whether we are dealing with an asymmetric or symmetric kind of coordination. The precise predictions apposition-based theories would make in this respect are somewhat more difficult to calculate. But we could at least state the following: Appositives occur quite freely within a sentence in general, they need not occur strictly adjacent to their anchor expressions. A sentence-final position as the comitative phrase in (29b) occupies, however, is one of the positions that is usually readily available for appositions. Since many aspects have not been spelled out in Cable's (2017) approach, it can only be surmised that this theory (as it currently stands) would probably not preclude an iPPC reading of (29b).

expression in the plural without having to adapt its own morpho-syntactic shape – it will remain as plain *svoj-* either way. In Torlakian BCMS, pronouns such as *naše* 'our' would have to be used instead in analogous examples as the ones from Russian below. The virtue of discussing examples with *svoj-* is thus that we are on the safer side in excluding cases of mere co-reference (instead of binding).

The (un)available referential properties of *svoj-* under an iPPC and an ePPC reading of the Russian sentence in (30) show that our predictions are correct. Although *svoj-* could also be interpreted as *his* (in the sense of *Petja's*) in principle, this binding configuration is not available under the iPPC reading of (30). Namely, the comitative referent from inside the plural pronoun's restrictor is not available as a sole binder of *svoj-* – and neither is the silent instance of reference to the speaker. Rather, all referential instances from the restrictor of the plural pronoun jointly bind the SELF element here.

(30) My   s   Petej   čitaem svoju   knigu.
     we.NOM with Petja.INST read.1PL POSS.REFL book

                                    (Russian; Vassilieva & Larson 2005: 112)

    a. 'We$_j$ are reading SELF$_{j/*i/*j+i}$'s book with Petja.'   ePPC
    b. 'Petja$_i$ and I$_j$ are reading SELF$_{*j/*i/j+i}$'s book.'   iPPC

As can be seen from the respective paraphrase in (30a), the comitative referent is not at all involved in binding *svoj-* in the ePPC case. We find the same pattern in "regular" comitatives where the comitative phrase is commonly analyzed as an adjunct to VP, see (31).

(31) Mal'čik s   kotënkom ušël   v svoju   komnatu.
     boy.NOM with kitten.INST went.3SG to POSS.REFL room
     'The boy$_j$ went to SELF$_{j/*i/*j+i}$'s room with the kitten$_i$.'

                                     (Russian; Vassilieva & Larson 2005: 109)

Therefore the conclusion suggests itself that the comitative phrase adjoins the VP in ePPC structures as well.

4 Conclusion

The analysis presented in this article is based on the assumption that plural pronouns and quantifiers have some crucial semantic and syntactic properties in common. In particular, I suggested that plural pronouns select restrictor arguments just like quantifiers. Analogous to instances of universal quantification, a plural pronoun conveys that for all elements in its restrictor denotation, it holds

that *Q* (where *Q* is some predicate). Under my approach, the restrictor of a plural pronoun contains the following three ingredients: A silent instance of the plural pronoun's singular counterpart (reference to the speaker/addressee), a silent pronominal element whose reference is determined by means of a contextual assignment function $g(i)$ (reference to other(s)), and a silent operator OP_U. The function of this operator is to form a set consisting of exactly those individuals to which the two referential instances from the restrictor refer. Thus, OP_U basically has a mechanism amounting to set union.

I argued that in an iPPC structure, the comitative phrase acts as the spell-out of the more general OP_U plus $g(i)$ part which is present in the restrictor of a plural pronoun anyway by default. Moreover, I claimed that this was the case since the comitative element in these constructions has the very same semantics as OP_U – that is to say that in an iPPC the comitative phrase occurs inside the plural pronominal restrictor, i.e. within the same DP.[18] I suggested that we can account for the availability of iPPC interpretations of split PPCs by assuming that these constructions involve a kind of syntactic movement similar to the kind involved in Quantifier-floating structures. And indeed, we found an intriguing correlation: the very Slavic languages that allow iPPC readings of split PPCs also have Quantifier-floating structures. However, as already pointed out in §3.2.1, the wrong element is "floated" under my approach actually. That is, if we really want to assume that split PPCs are related to Q-floating structures, then we are faced with the question of why the restrictor gets floated in a split PPC, whereas the quantifier does in a Q-floating structure. I am aware of this problem, and of the fact that the analogy between these two cannot be 1:1 – but for the sake of explicitness, I stick with it for the time being; and leave it as an issue for further research which kind of syntactic movement could be involved in split PPCs instead, if any. An obvious candidate is so-called "Left Branch Extraction" (LBE; cf. Bošković 2008, and the references therein). But whether the generalizations made in relation to LBE in Slavic languages, and the generalizations we could derive in this article are actually compatible with each other remains to be seen. And in any case, it seems remarkable that the positions available for floated quan-

---

[18]As an anonymous reviewer mentions, this raises the question why (i)PPCs are not more widespread across the world's languages. But one of the crucial points of my analysis is indeed that the comitative element *s* and OP_U have the same semantics – and there is no salient reason to assume that a particularly large number of languages should have a comitative element whose semantics matches those of OP_U. Moreover, other languages have found different ways of expressing the same meaning as an iPPC. Just consider Icelandic which has so-called Pro[NP] constructions (cf. Sigurðsson & Wood 2020). Those Pro[NP]s superficially differ from PPCs only in that they lack a comitative element.

tifiers within a sentence coincide with the positions available for the comitative phrase in a split PPC constructions – at least in Torlakian BCMS.[19]

Other issues remained open here as well. For example, it is puzzling why an iPPC reading in Torlakian BCMS is much more salient when the plural pronoun is dropped, and *vice versa* an ePPC interpretation is much more salient when the pronoun is pronounced; see (32).

(32) a. Mi   smo   s   Mariju   otišli u muzej.
         we.NOM AUX.1PL with Maria.INST went in museum
         'We went to the museum with Maria.'   (more salient) ePPC
         'Maria and I went to the museum.'   (less prominent) iPPC

     b. *pro* s   Mariju   smo   otišli u muzej.
            with Maria.INST AUX.1PL went in museum
         'We went to the museum with Maria.'   (less prominent) ePPC
         'Maria and I went to the museum.'   (more salient) iPPC
                                                      (Torlakian BCMS)

I can only speculate here that this is a pragmatic effect. More precisely, one could assume that the pronunciation of the plural pronoun (which could otherwise be dropped) indicates that we are dealing with a fully determined instance of the plural pronoun and that the comitative phrase consequently does not occur inside its restrictor.

Moreover, it remains unclear why predicates such as *hate broccoli* or *believe in God* are felicitous with PPCs under an iPPC interpretation (but crucially not under an ePPC interpretation) in Russian (see (33a) and (33b)), but not in Torlakian BCMS, where analogous sentences are simply infelicitous no matter which reading is intended.

(33) a. My   s   Dašei   verim   v boga.
         we.NOM with Daša.INST believe.1PL in God
         Unavailable ePPC reading: 'We believe in God with Daša.'
         'Daša and I believe in God.'   iPPC
                                             (Russian; Feldman 2003)

     b. My   s   Ivanom   nenavidim brokkoli.
         we.NOM with Ivan.INST hate.1PL   broccoli
         Unavailable ePPC reading: 'We hate broccoli with Ivan.'

---

[19]But to make Torlakian BCMS not to appear very exotic in this respect: The same correspondence of available positions can also be found in other languages (belonging to different language families) that have both Quantifier-floating and (split) PPCs – namely Finnish and Fenno-Swedish; see also Butschety (2023).

Madeleine Butschety

'Ivan and I hate broccoli.'   iPPC
(Russian; Vassilieva & Larson 2005: 112)

These topics need a much longer discussion, which I have to defer to another occasion.

To conclude, based on the observed analogies between plural pronouns and (universal) quantifiers, the two readings PPCs in Slavic can give rise to were derived in terms of whether the comitative phrase occurs inside the plural pronominal restrictor (=iPPC), or outside of it as an adjunct to VP (=ePPC). Assuming that split PPCs involve a similar kind of syntactic movement as floated quantifiers, my analysis could derive in which Slavic languages (namely precisely in those that have Q-floating structures) on the one hand, and in which syntactic configurations on the other hand iPPC readings arise for split PPCs. While this analogy turned out to be not as direct as intended, the correlation between the availability of floated (universal) quantifiers and iPPC readings for split PPCs nonetheless represents a noteworthy datum. I leave it to future research to decide on whether and how these two phenomena can be put into a uniform picture.

Abbreviations

1	first person	F	feminine
2	1second person	FUT	future tense
3	third person	INST	instrumental
ADDR	addressee	NOM	nominative
AUX	auxiliary	PL	plural
CL	clitic	POSS.REFL	possessive reflexive
COMP	complementizer	SPKR	speaker
DEF	definite	SG	singular

Acknowledgments

References

Bobaljik, Jonathan. 1998. Floating quantifiers: Handle with care. *Glot International* 3. 3–10. DOI: 10.1515/9783110890952.107.

Bošković, Željko. 2004. Be careful where you float your quantifiers. *Natural Language & Linguistic Theory* 22. 681–742. DOI: 10.1007/s11049-004-2541-z.

Bošković, Željko. 2008. What will you have, DP or NP? In Emily Elfner & Martin Walkow (eds.), *NELS 37: Proceedings of the 37th annual meeting of the North East Linguistic Society*, 101–114. Amherst, MA: Booksurge Publishing.

Büring, Daniel. 2011. Pronouns. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, 971–996. Berlin: De Gruyter Mouton.

Butschety, Madeleine. 2023. *Inclusory constructions*. Graz: University of Graz. (Doctoral dissertation).

Cable, Seth. 2017. Some observations on the plural pronoun construction of Tlingit, Polish, and Russian. In Claire Halpert, Hadas Kotek & Coppe van Urk (eds.), *A pesky set: Papers for David Pesetsky*, 266–290. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris. DOI: 10.1515/9783110884166.

Daniel, Mikhail. 2000. *Tipologija associativnoj množestvennosti*. Moscow: Russian State University for Humanities. (Doctoral dissertation).

Dyła, Stefan. 1988. Quasi-comitative coordination in Polish. *Linguistics* 26. 383–414. DOI: 10.1515/ling.1988.26.3.383.

Elbourne, Paul. 2008. The interpretation of pronouns. *Language and Linguistics Compass* 2(1). 119–150. DOI: 10.1111/j.1749-818x.2007.00041.x.

Feldman, Anna. 2003. Comitative and plural pronoun constructions. In Yehuda N. Falk (ed.), *Proceedings of the 17th annual meeting of the Israel Association of Theoretical Linguistics*. https://www.iatl.org.il/?page_id=155.

Feldman, Anna & Stefan Dyła. 2008. On comitative constructions in Polish and Russian. In Lew Zybatow, Gerhild Zybatow, Luka Szucsich & Roland Meyer (eds.), *Formal description of Slavic languages: The fifth conference, Leipzig 2003*. Frankfurt am Main: Peter Lang.

Fitzpatrick, Justin M. 2006. *The syntactic and semantic roots of floating quantification*. Cambridge, MA: MIT. (Doctoral dissertation). https://dspace.mit.edu/handle/1721.1/37420.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Malden, MA: Blackwell.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30(1). 69–96. https://www.jstor.org/stable/4179050.

Madeleine Butschety

Kayne, Richard. 1981. Binding, quantifiers, clitics and control. In Frank Heny (ed.), *Binding and filtering*, 191–211. Cambridge, MA: MIT Press. DOI: 10.1515/9783111682228-005.

Ladusaw, William A. 1989. Group reference and the plural pronoun construction. In *Papers on the plural pronoun construction and comitative coordination*, 1–7. Santa Cruz, CA: UCSC Syntax Research Center.

Landau, Idan. 2003. Movement out of control. *Linguistic Inquiry* 34(3). 471–498. http://www.jstor.org/stable/4179243.

Larson, Richard & Gabriel Segal. 1995. *Knowledge of meaning: An introduction to semantic theory*. Cambridge, MA: MIT Press. DOI: 10.7551/mitpress/4076.001.0001.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretic approach. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings*, 127–147. Malden, MA: Blackwell. DOI: 10.1515/9783110852820.302.

Lipták, Anikó Kata. 2015. Identity in ellipsis: An introduction. *Lingua* 166(B). 155–171. DOI: 10.1016/j.lingua.2015.08.003.

McNally, Louise. 1993. Comitative coordination: A case study in group formation. *Natural Language & Linguistic Theory* 11. 347–379. DOI: 10.1007/BF00992917.

Merchant, Jason. 1996. Object scrambling and quantifier float in German. In Kiyomi Kusumoto (ed.), *NELS 26: Proceedings of the North East Linguistic Society*, 179–193. Amherst, MA: GLSA.

Nunberg, Geoffrey. 1993. Indexicality and deixis. *Linguistics and Philosophy* 16(1). 1–43. https://www.jstor.org/stable/25001498.

Shlonsky, Ur. 1991. Quantifiers as functional heads: A study of quantifier float in Hebrew. *Lingua* 84(2–3). 159–180. DOI: 10.1016/0024-3841(91)90069-H.

Sigurdsson, Halldór Ármann & Jim Wood. 2020. *We olaf*: pro[(x-)np] constructions in icelandic and beyond. *Glossa: a journal of general linguistics* 5. 1–16. DOI: 10.5334/gjgl.1054.

Sokolová, Jana. 2019. On comitative constructions in Slovak. *Skase Journal of Theoretical Linguistics* 16(3). 90–107.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 84(3). 425–449. https://www.jstor.org/stable/25164903.

Stokke, Andreas. 2022. Features of referential pronouns and indexical presuppositions. *Inquiry: An Interdisciplinary Journal of Philosophy* 65(8). 1083–1115. DOI: 10.1080/0020174x.2020.1801500.

Szabolcsi, Anna. 2010. *Quantification* (Research Surveys in Linguistics). Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511781681.

Trawinski, Beata. 2005. Plural comitative constructions in Polish. In Stefan Müller (ed.), *Proceedings of the 12th international conference on Head-Driven Phrase Structure Grammar*, 375–395. Stanford, CA: CSLI Publications.

Vassilieva, Maria B. & Richard K. Larson. 2001. The semantics of the plural pronoun construction. In Rachel Hastings, Brendan Jackson & Zsofia Zvolenszky (eds.), *Proceedings of SALT 11*, 449–465. Ithaca, NY: CLC Publications. https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2853/2593.

Vassilieva, Maria B. & Richard K. Larson. 2005. The semantics of the plural pronoun construction. *Natural Language Semantics* 13. 101–124. DOI: 10.1007/s11050-004-1031-5.

Vulchanova, Mila & Guiliana Giusti. 1995. Quantified noun phrase structure in Bulgarian. *University of Venice Working Papers in Linguistics* 5(1). 43–65.

Chapter 7

Equatives and two theories of negative concord

Mojmír Dočekal

Masaryk university

This article reports the results of an experiment targeting the acceptability of Czech neg-words and strong NPIs under Neg-Raising predicates and in the complement clauses of equatives. The theoretical consequences of the results are discussed and range from the support of non-standard negative concord theories to the support of non-standard degree semantics for the equative constructions.

1 Introduction

In this article, I explore expressions that are polarity-dependent. The evidence comes from Czech, a strict negative concord language. I will focus on one recent experiment in the paper.[1] The theoretical ambition of this paper is to examine the distribution of neg-words and strong NPIs in two environments: Neg-Raising predicates and equatives. The acceptability pattern of these two kinds of negative dependent expressions is challenging for standard theories of neg-words but also for the current degree theories of equatives. The data are subtle, and therefore I report results of an acceptability judgment task experiment on Czech native speakers; in this way, I add to the experimental research on Negative Polarity Items (NPIs), like Chemla et al. (2011), Gajewski (2016), Alexandropoulou et al. (2020), a.o., more specifically to the the experimental research on cross-linguistic variation in NPI licensing as found in Djärv et al. (2018), Schwarz et al. (2020), Chierchia (2019). Empirically the experimental data concern the Czech strong

---

[1] But the experiment is a continuation of many previous experimental works which incrementally changed the nature of questions and research goals reflected in the current experiment.

Mojmír Dočekal

NPIs, like *ani jeden* 'even one,' and neg-words, like *žádný* 'no' (neg-word), as exemplified in (1). Both polarity-sensitive items are in the majority of contexts interchangeable, but their meaning differs, and the experiment was focused on the environments where the meaning difference is detectable. In previous works, the acceptability of NPIs was reported as varying between speakers (Homer 2021, Mayer et al. 2018). In the neg-words research, the variation of acceptability was found to correlate with demographic factors such as age or education (Burnett et al. 2015, 2018). For these reasons, I included also demographic variables in my experimental research to see whether a more complex picture (that incorporates both grammatical and demographic factors) can explain the distribution of neg-words and strong NPIs more successfully.

(1) Petr nepotkal {ani jednoho / žádného}   studenta.
     Petr NEG.met   STRONG.NPI   NEG.WORD student
     'Petr didn't meet {even one / any} student.'

The article is structured as follows: in §1.1, I introduce the theoretical background for the experiment. In §2, I describe the experiment. In §3, I discuss the results, and in §4, I conclude the article.

1.1 Theoretical background

1.1.1 Polarity dependent expressions in equatives and under Neg-Raising predicates

The theoretical background for the experiment is the contrast between strong NPIs and neg-words in two environments, under Neg-Raising predicates and in equatives. As a baseline against the two environments, I used simple unembedded sentences with negated verbs. The baseline is important since it is the only environment where both strong NPIs and neg-words are grammatical under any theory of neg-words and NPIs. The baseline is illustrated in (1). In the current section I discuss the theoretical background for the constructions (equatives and Neg-Raising predicates), the section §1.1.3 is dedicated to the theoretical background for the licensing of strong NPIs and neg-words.

As for Neg-Raising, we can adapt any current theory of Neg-Raising, be it the presuppositional version of Gajewski (2007) or the scalar implicature version of Romoli (2013). Both share the insight that Neg-Raising predicates bear the excluded middle inference: for *believe*: $Bel(p) \lor Bel(\neg p)$, adding the negated assertion $\neg Bel(p)$ results in the deductively valid conclusion where the negation scopes in the embedded clause, $Bel(\neg p)$. Non-Neg-Raising predicates then come

without the excluded middle inference resulting in the surface interpretation of the negation. Since the scope of negation ends in the embedded clause, the negation is local, and therefore, the strong NPIs are licensed, as demonstrated with Spanish *ni un* 'not even one' in (2).

(2) No creo   que ni un solo soldado pueda lograrlo.
     not believe.1SG that not even one soldier can   achieve
     'I don't believe that not even one soldier can achieve it.'

The second environment which was tested in the experiment was equatives. The first thing to note is that the standard theory of equatives is built on the ">" analysis of comparatives (Beck 2019, Stechow 1984) where the core operation is the relation > comparing two maxima: (i) the maximum of the set of degrees from the main clause, (ii) the maximum of the set of degrees from the complement of the comparative clause, see (3) as an illustration.

(3) The dog is taller than the cat.
     a. MAX(d|the height of the dog $\geq d$) $>$ MAX(d|the height of the cat $\geq d$)

The standard theory of equatives Beck (2019), Stechow (1984), Rullmann (1995) then follows the ">" analysis of comparatives, just replacing > with $\geq$ which is in most contexts pragmatically strengthened to "=", see (4) for an illustration. Comparatives are then theoretically expected to license NPIs in their complement clauses since the degree argument is downward-monotonic, therefore if some degree $d > d'$ and there is another degree d'', such as $d' > d''$, then by transitivity $d > d''$. Intuitively: if the dog from (3) is taller (or of the same height) than the cat from (3), then he is taller than any cat smaller than the cat. The literature on comparatives (Stechow 1984, Rullmann 1995, Gajewski 2008) agrees on the empirical verification of this prediction. Weak NPIs (like English *any*) are licensed in the complement clauses of the comparative. In the case of strong NPIs, the empirical situation is less clear. Still, at least empirically, it is claimed for Germanic languages that strong NPIs appear in the complement clause of comparatives felicitously; see Hoeksema's (2012) Dutch example in (5). (5) contains the Dutch expression *ook maar* 'even,' which is taken as a standard example of a strong NPI, see Zwarts (1998).

(4) The dog is as tall as the cat.
     a. MAX(d|the height of the dog $\geq d$) $\geq$ MAX(d|the height of the cat $\geq d$)

Mojmír Dočekal

(5) Zij was beter dan ook maar iemand verwacht had.
     she AUX better than STRONG.NPI   expected AUX
     'She was better than anyone could have expected.'

Under the premise that equatives are built on the ">" analysis of comparatives, the standard theory of equatives predicts that NPIs should be licensed in the complement clause of equatives. This predictions works for English and supports the standard theories of equatives (see Stechow 1984, Beck 2019, a.o.), since NPIs in English equatives are licensed, see (6) from Seuren (1984).

(6) Paris is as quiet as ever.

The theories of Neg-Raising and equatives are general and their aim is to model the meaning of the construction. The interaction of the constructions with various classes of polarity dependent expressions is not their primary concern. Nevertheless, as discussed in this section at least for strong NPIs, the predictions of the standard theories of Neg-Raising and standard theories of equatives are clear. The strong NPIs should be licensed. But if we want to apply the predictions of the standard theories of Neg-Raising and equatives to both strong NPIs and neg-words, we have to introduce the theories of neg-words and strong NPIs. This is the topic of the next section.

1.1.2 Czech strong NPIs and neg-words

Let us introduce some background information and intuitions concerning both classes. Starting with strong NPIs (for a theoretical framework, see Gajewski 2011), Czech strong NPIs of the *ani* sort bear the unlikelihood presupposition, discussed concerning English stressed *ANY* (see Krifka 1995, a.o.), Hindi *ek bhii* (see Lahiri 1998, a.o.) or English *even one* (see Crnič 2014b, a.o.). But unlike the English or Hindi strong NPIs, the Czech strong NPIs are much more limited in distribution, requiring clause-mate negation in most of their occurrences. Nevertheless, this requirement is not obligatory, as will be demonstrated, and Czech *ani* strong NPIs can appear embedded under negated Neg-Raising predicates without any overt clause-mate negation. But in all contexts, *ani* presupposes that its prejacent (a proposition which *ani* modifies) entails all the relevant alternatives. By way of example: *ani jeden* 'even one' in (7) is acceptable since not scoring one goal entails not scoring two, three, etc. goals ($\neg\text{SCORE}(1) \vDash \neg\text{SCORE}(2 : \infty)$), the relevant alternatives but *ani deset* 'even ten' is much less acceptable since not scoring ten goals is entailed by not scoring 9, 8, etc. goals ($\neg\text{SCORE}(1 : 9) \vDash \neg\text{SCORE}(10)$) and therefore the prejacent does not entail all the relevant alternatives. In this

respect, *ani jeden* belongs to the same class of strong NPIs as English *even one*, which yield the scalar presupposition (in (7) the focus alternatives have to be less probable and entailed by the prejacent).

(7) FC Barcelona nedala   {ani jeden / #ani   deset} gól/ů.
    FC Barcelona NEG.gave even one   even ten   goal(s)
    'FC Barcelona didn't score {even one/#ten} goal(s).'

Turning now to neg-words, Czech (and generally Slavic) neg-words are similar to Italian neg-words (as *niente*, e.g., see Ladusaw 1992). In contrast to strong NPIs like *ani jeden* in (7), or English *even one*, neg-words do not bear any scalar or additive presupposition. Next, neg-words have strong syntactic requirements on their licensing, and in Czech, as in all Slavic languages, which are strict negative-concord languages (see Zeijlstra 2004, a.o.), Czech neg-words in the majority of contexts require verbal negation (in the same clause), the requirements being more strict than in case of Czech strong NPIs, see (8). Unlike strong NPIs, neg-words do not yield the scalar presupposition but their licensing is more locality constrained. Therefore Czech neg-words are degraded under negated Neg-Raising predicates (see Dočekal & Dotlačil 2016a,b for details), and (23b) for an example from the experiment.

(8) a. Petr nedal   žádný   gól.
       Petr NEG.scored NEG.word goal
       'Petr didn't score any goal.'

    b. Nikdo   {nepřišel   / #přišel}.
       NEG.WORD NEG.came   came
       'Nobody came.'

    c. *Petr neřekl, že nikdo   přišel.
       Petr NEG.said that NEG.WORD came
       'Petr didn't say that anybody came.'

The most influential current analysis of neg-words is the syntactic approach of Zeijlstra (2004), a.o. (the standard theory/Zeijlstra 2004 hereinafter). It claims for strict negative concord languages that all neg-words (and the verbal negation) carry [uNeg] feature and are checked against [iNeg] (covert) operator with the semantics of ¬. Part of this paper is dedicated to the experimental support for an alternative semantic theory of neg-words (see Ovalle & Guerzoni 2004, Kuhn 2022, a.o.; I will refer to this theory as the alternative theory/Ovalle & Guerzoni 2004), which will be explained in detail later, see §1.1.3. The empirical point concerns equatives, one of the contexts where the distribution of strong NPIs and

Mojmír Dočekal

neg-words diverge. It was noted for Polish equatives at least as early as Błaszczak (2001) that neg-words are surprisingly grammatical in them. Czech equatives are similar; they seem not to license strong (and weak) NPIs, resembling German and many other non-English equatives, see Krifka (1992), a.o. Nevertheless, neg-words are very much acceptable in the complement clauses of Czech equatives, see (9). The acceptability of neg-words in Czech equatives is also surprising against the standard theory of neg-words/Zeijlstra (2004), since there is no plausible overt or covert operator with the interpretable [iNeg] feature in the complement clause of Czech equatives. The standard theory of neg-words/Zeijlstra (2004) predicts the ungrammaticality of neg-words in Czech equatives, which is empirically wrong. Part of the experimental work reported in this paper is to test the acceptability of neg-words in Czech equatives, see (9) and to compare it with the acceptability of strong NPIs.

(9) Petr je tak vysoký jako {#ani jeden / žádný} jiný student.
Petr is so tall how STRONG.NPI NEG.WORD other student.
'Petr is as tall as any other student.'

Concerning Neg-Raising predicates, Czech (like Spanish in (2)) allows licensing of strong NPIs in the embedded clause, see (10). As for neg-words, previous experimental research reported their decreased acceptability under Neg-Raising predicates, see Dočekal & Dotlačil (2016a). Such pattern is expected in the standard theory of neg-words/Zeijlstra (2004), since the negation is syntactically localized in the root clause. In terms of the syntactic approach, the root negation bearing [iNeg] feature is too far away from eventual neg-words in the embedded clause to license them.

(10) Nechci, aby ani jeden student odešel.
NEG.want.1SG that even one student left
'I don't want even one student to leave.'

To this end, the experiment reported below scrutinizes the contrast between strong NPIs and neg-words in equatives (and under Neg-Raising predicates). First, Slavic literature observed the acceptability of neg-words in equatives, and second, NPI literature noticed the unacceptability of strong NPIs in Germanic, nevertheless, the contrast was not either experimentally researched or theoretically explained. Moreover, equatives are one of the environments where the contrast between Czech neg-words and strong NPIs is most robust, but still, there seems to be a speaker variation involved. In simple terms, some speakers treat *ani* as a

neg-word and therefore do not accept it as much in equatives, unlike the speakers who use *ani* as the strong NPI. In a bit broader picture, the speaker variation resembles the variation of English NPIs vs. negative quantifiers, e.g., as studied first by functional linguists, see Tottie (1991), a.o., and more recently in the formal syntactic tradition, see Burnett et al. (2015, 2018), a.o. Burnett et al. (2015, 2018) show that the formal constraints explain the English speaker variation with higher success than historical and social factors (discovered in the functional tradition before). According to Burnett et al. (2018), the English negative quantifiers are replaced by NPIs in lower syntactic domains.[2] This process overrules any demographic factors, like age or education. In a similar vein, Burnett et al. (2015) describe the variable negative concord in Québec French as explainable by the interplay of grammatical and demographic factors where the first type of factors is decisive.

The speaker variation mentioned above is an intriguing and hard-to-pin-down phenomenon and one of the reasons for using experimental methods since it certainly resists any simple intuition-based methods for data collecting. The emerging picture is that speaker variation concerning negation, negative concord, negative quantifiers, and NPIs comes both from social and grammatical sources, and only experimental work can give some reasonable answers as to their respective strength. In this respect, the experimental work reported below is the first tiny step in explaining Slavic neg-words vs. NPIs speaker variation due to the possible interplay between demographic and grammatical factors.

The following section introduces the licensing conditions for strong NPIs and neg-words in a formal way. At the end of the section the predictions of the theories of neg-words, strong NPIs, Neg-Raising and equatives are summarized.

1.1.3 Assumptions concerning licensing of (strong) NPIs

Let us assume a standard approach to NPIs and strong NPIs licensing. For the general framework, the so-called *even*-theory of NPIs licensing is naturally the most attractive candidate, see Krifka (1995), Lahiri (1998), Crnič (2014b), a.o., since *ani* bears the unlikelihood presupposition similar to English *even*. And for strong NPIs, let us follow Gajewski's formalization of strong NPIs (Gajewski 2011). According to Gajewski (2011), strong NPIs are licensed in downward-entailing (DE)

---

[2]Consider a contrast like *There were no jobs to be had* – higher syntactic domain vs. *I can't have any form of gluten*, where in the first sentence, the negative quantifier is used, while in the second sentence, an NPI occurs. Examples come from Burnett et al. (2018), where it is claimed that while this constraint is soft in contemporary English, it is a hard one in the Scandinavian language family.

environments. But the downward entailments are checked both in Truth Conditions (TC), the at-issue part of the meaning, and in the non-at-issue meaning, presuppositions and implicatures being the most pertinent non-at-issue meaning components. Weak NPIs, on the other hand, require DE environments only in the TC part of the meaning. The conditions for weak and strong NPIs are summarized in (11).

(11) An NPI is licensed in the environment γ
$[_{\alpha}exh[_{\beta}...[_{\gamma} \text{ NPI }]...]]$:
   a. the environment γ is DE in β weak NPIs
   b. the environment γ is DE in α strong NPIs

The standard exhaustifier from (11) is the formalization of *only*-kind of focus operator which works very well for weak NPIs like unstressed English *any*. But for other weak or strong NPIs with the unlikelihood presupposition meaning core, another kind of exhaustifier, a covert counterpart of English *even* was proposed (see Crnič 2011, 2014a). The same mechanism is used in formal approaches to focus particles, see Panizza & Sudo (2020). The *even* exhaustifier, like its overt version, then comes with two presuppositions. The first is scalar, demonstrated in (12a) – the sentence is acceptable in such contexts where a dancing Pope is very unlikely (compatible with the actual world). The second is additive, exemplified with (12b). The sentence is true if two, three, ...cats will make the Pope happy as well. The placement of focus determines the nature of alternatives used in presuppositions. Let us follow the formalization of both presuppositions by Panizza & Sudo (2020), see (13). For monotonic scales, likelihood from (13) translates into entailment (after Crnič 2011), therefore the predictions of traditional downward entailing approaches like (Ladusaw 1992) and *even*-theories of NPIs collapse for downward monotonic contexts.

(12)   a. Even the Pope$_F$ danced.
      b. Even one$_F$ cat will make the Pope happy.

(13) 'Even ϕ' presupposes:
    a. that ϕ is relatively unlikely to be true among Alt(ϕ); and
    b. that there is $\psi \in$ Alt(ϕ) that is not entailed by ϕ and is true.

Assuming this standard approach to strong NPIs, its predictions are clear for simple negated sentences like (14) where *even* associates with the weak scalar item (the numeral *one*). The scalar presupposition of *even* has to scope over negation,

schematically [$_{even}$¬[one student arrived]] and since this logical form entails all other alternatives ([$_{even}$¬[n students arrived]], where $n > 1$), the prejacent is both strongest and the least likely from the alternatives and the scalar presupposition (see (13)) is fulfilled. In the experiment, I used baseline sentences of similar form for both strong NPIs and neg-words. Both expressions were unsurprisingly well accepted in the baseline.

(14)   Even one student didn't arrive.

For Neg-Raising predicates like *want* from (15), then the schematic scope configuration in the embedded clause is covert (*even*) $> ¬ > [_{\alpha}...$ one ...] – the scalar presupposition of *even* is fulfilled. Moreover, the licensing condition for strong NPIs (see (11)) requires that the local domain (α in (15)) is DE, meaning after we factor in all non-at-issue meaning components, which is the case for Neg-Raising predicates.

(15)   The director doesn't want [$_{\alpha}$ even one student to depart].

As for equatives, they are theoretically expected to license NPIs which seems to be the case for English weak NPIs as illustrated by (6), repeated for convenience below as (16).

(16)   Paris is as quiet as ever.

But it was noticed before that this does not hold cross-linguistically, see Krifka (1992) for German and Penka (2016) for German and Romance languages. But at least in the comparative/equative ">" theories, if comparatives license strong NPIs, the expectation is that equatives will behave similarly. Turning now to Slavic equatives, there are many factors at play here, though. First, Slavic equatives are different from English equatives, and their morpho-syntax is very similar to correlatives (as German and Romance equatives). And since it is known at least from Jacobson (1995) that correlatives are bad licensors of NPIs, the expectation is that both weak and strong NPIs will be much worse in Slavic equatives (compared to Germanic languages).[3] To summarize, the standard degree theory of equatives – the standard degree theory/Rullmann (1995) hereinafter – (and assuming the standard theory of NPIs licensing introduced above) predicts that NPIs (weak and strong) should be licensed in equatives. I was unable to locate

---

[3]To address this issue, another experiment targeting both weak and strong NPIs in comparatives and equatives is in preparation.

any scholarly discourse pertaining to strong NPIs and equatives, but the following example (17) from Chisholm (2010) can be seen as an empirical approval of the standard degree theory prediction for English. In (17), there is a strong NPI *until recently* in the equative standard; the verb *changed* is telic. Therefore the strong NPI should be licensed by the DE logical properties of the English equative, which seems to be the case. But as will be demonstrated in the next section, the situation is quite different in Czech, which confirms the observations concerning German and Romance equatives (Krifka 1992, Penka 2016) and in alternative theories of equatives (the alternative degree theories/Penka 2016 furthermore).

(17)   Under the party system in Canada cabinets changed as often as, until recently, they did in France.

1.1.4 Assumptions concerning licensing of neg-words

As for the licensing of neg-words, I will now introduce the syntactic theory of Negative Concord (NC) developed by Zeijlstra (2004), Penka (2007), Zeijlstra (2022) in detail. The standard theory/Zeijlstra (2004) is the syntactic tool for dealing with negative concord both in strict and non-strict negative concord languages. Since Czech (like all Slavic languages) is an example of strict negative concord, I will focus on the part of the theory that deals with strict NC. The basic assumption for strict NC languages in Penka/Zeijlstra's syntactic theory is that all morphologically negated words come without semantic negation. Negwords and sentential negation carry a so-called uninterpretable [uNeg] feature, which is in agreement with the logical operator (propositional negation) that has an interpretable [iNeg] feature. Sentential negation is a signal of propositional negation, but propositional negation is located higher in the syntactic tree than sentential negation. The syntactic theory then treats neg-words as indefinites, and their negation is purely syntactical (the uninterpretable feature). The purpose of the uninterpretable feature is then to signal the presence of the propositional negation operator. Let us illustrate the mechanism used in the syntactic theory on a Czech example in (18a). The sentence contains three morphological negations, but according to the syntactic theory, none of them bears semantic force, which is delegated to the abstract logical Operator with the semantics of classical propositional negation, see (18a-i). The final logical form is in (18b).

(18)   a.   Nikdo   neviděl nic.
            NEG.person NEG.saw NEG.thing
            'Nobody saw anything.'

        i. $\text{Op}_{\neg[\text{iNeg}]}[\text{Nikdo}_{\neg[\text{uNeg}]} \text{ neviděl}_{\neg[\text{uNeg}]} \text{ nic}_{\neg[\text{uNeg}]}.]$

        ii. $[\![\text{nikdo}]\!] = \lambda P \exists x [\text{PERSON}(x) \wedge P(x)]$

        iii. $[\![\text{nic}]\!] = \lambda P \exists x [\text{THING}(x) \wedge P(x)]$

        iv. $[\![\text{neviděl}]\!] = \lambda y.\lambda x.\text{SEE}(x, y)$

    b. $\neg \exists x \exists y [\text{PERSON}(x) \wedge \text{THING}(y) \wedge \text{SEE}(x, y)]$

The syntactic theory is well equipped to deal with the locality constraints on negative concord and, of course, easily explains the baseline kind of example like (18a), where the neg-words and verbal negation appear in a root clause. The logical Operator has to be local, around the level of TP projection (of the clause where neg-words or verbal negation appears).

As for Neg-Raising, the predictions of the syntactic theory are the following: since the inferential process by which the scope of negation ends on the embedded predicate (schematically: $\neg$ NegRaisingVerb [Predicate] $\rightsquigarrow$ NegRaisingVerb [$\neg$ Predicate]) is pragmatic in nature, the excluded middle inference is (depending on theory) treated either as a presupposition or as an implicature. The valid scope of the invisible Operator is the root sentence (Op$_\neg$[NegRaisingVerb [Predicate]]). Therefore, neg-words in the embedded sentence are too far away for the agreement between the uninterpretable feature and the interpretable feature of the Operator.

Nevertheless, in the case of equatives, the standard theory/Zeijlstra (2004) simply predicts the ungrammaticality of neg-words (in the case of the positive main predicate), which is empirically wrong. Let us start with some empirical observations. According to Sketch Engine (Kilgarriff et al. 2014), in their csTenTen19 (the most representative Czech corpus in Sketch Engine), there are 28 occurrences of neg-words in the standard clause of equatives.[4] One example sentence from the query is in (19). This is in contrast to strong NPI: in Sketch Engine, there is no occurrence of strong NPI *ani* in the standard clause of equatives. This asymmetry is also verified by the intuitions of native speakers, as will be reported in the experiment. The empirical inadequacy of the syntactic theory follows from the standard theory of equatives as the $\geq$ relation between two maxima of two sets of degrees – there is no place for negation in the semantics of equatives, neither in the standard theory nor in the alternative theories of equatives (see Penka 2016, a.o.). And for this reason, I will now introduce the alternative, non-standard theory of neg-words.

---

[4]The CQL used for the search was: `[lemma="tak"]` `[tag="k2.*"]` `[lemma="jak[o]?"]` `[lemma="žádný"]`.

(19) Ve zbarvení   je pstruh obecný tak variabilní jako žádná   naše ryba.
      in coloration is trout   brown as variable   like NEG.WORD our fish
      'The brown trout is as variable in coloration as any of our fish.'

The alternative theory of neg-words was formulated in Ovalle & Guerzoni (2004),
and a modern reformulation can be found in Kuhn (2022). It shares some as-
sumptions with the syntactic theory, though. First, both theories agree on the
indefinite description status of neg-words. Therefore neg-words denote sortally
existential quantifiers like in (20a) in the alternative theory too. The negative
force, which in the syntactic theory is carried by the covert Op (the bearer of the
classical logical semantic of ¬), is in the semantic/pragmatic theory reformulated
as a presupposition of empty reference in the original version, see (20b). Or in
the dynamic reformulation as a test on the cardinality of discourse referents like
in (20c). In this article, we can abstract away from the formal implementations
and work with the core assumption: the emptiness of reference is a presuppo-
sition with the usual projection properties of presuppositions. One of the main
differences concerns the interpretation of verbal negation in strict negative con-
cord languages though. While in the standard theory/Zeijlstra (2004), the verbal
negation is just an agreement negation with the active covert logical operator,
which carries the logical negation, in the alternative theory, the verbal negation
has its semantic interpretation, the classical propositional logic ¬.

(20) a. $\llbracket\text{neg-word}\rrbracket = \lambda P.\exists x[\text{SORT}(x) \wedge P(x)]$ TC
      b. $\llbracket\text{neg-word}\rrbracket = \neg\exists x[\text{SORT}(x) \wedge P(x)]$ non-at-issue
      c. after Kuhn (2022): $\wedge 0_x$ ...postsupposition (highest scope)

The original version of the semantic/pragmatic theory/Ovalle & Guerzoni (2004)
does not come with any locality constraints on the neg-words licensing, which
is a problematic assumption since negative concord is, in most cases, limited to
the clause-internal dependency between neg-words and verbal negation. This is
also one of the reasons why the syntactic approach is so successful and remains
the standard theory of neg-words today. Kuhn (2022) improves in many aspects
over the original version of the semantic/pragmatic theory, one of them is the
delimitation of the emptiness of reference presupposition in terms of previous
contexts and also in tying it to discourse referents and therefore making the pre-
supposition more specific. But most importantly, Kuhn (2022) brings some syn-
tactic constraints into the game. He formalizes neg-words' syntax via split-scope
around their licensor (prototypically verbal negation). Since the split scope is re-
alized via quantifier raising, some locality constraints on the neg-word emerge.

More specifically, Kuhn's (2022) empirical claim is that the locality constraints on neg-words licensing should correspond to the locality of quantifier raising in the particular language and construction. Whether this is the right theoretical solution is a separate question, which is not answerable in this article. Still, it is definitely a step in the right direction, including some form of syntactic sensitivity for locality into the semantic/pragmatic theory.

Let us go through the predictions the semantic/pragmatic theory/Ovalle & Guerzoni (2004) makes concerning the baseline (simple root sentences with negated verbs), Neg-Raising sentences, and equatives. For the first environment, the predictions of the standard syntactic approach/Zeijlstra (2004) do not differ from the alternative one/Ovalle & Guerzoni (2004). Both approaches agree on the indefinite and positive at-issue meaning of neg-words. The syntactic theory delegates the negative property into the uninterpretable features; the alternative theory explains the negative force as a presupposition. In simple cases, like (18a), both theories predict grammaticality (either via feature checking or by the verification of the emptiness of the reference presupposition).[5]

The predictions of the standard theory/Zeijlstra (2004) concerning Neg-Raising were already introduced. The alternative theory/Ovalle & Guerzoni (2004) requires quantifier rising of the neg-word over its licensor (negation) in syntax, but since the scope of negation in Neg-Raising predicates ends on the embedded

---

[5]More interesting is how both theories account for the sentences like Spanish (i), where the neg-word c-commands a positive predicate. Such configurations are ungrammatical in strict negative concord languages, though. The syntactic approach does not have a straightforward answer for the ungrammaticality of such [Neg-word positive-V] sentences since both neg-words and verbal negation are posited to bear uninterpretable features, so it is not clear why one such feature is not enough to signal the covert Op$_{\neg}$. A way out is offered by Penka (2007) in (ii), but as she herself admits, the principle is not anything else than restating the problem. For the alternative theory, the answer for the ungrammaticality of (i) in the strict NC languages is straightforward: the presupposition of neg-word clashes with the assertion of the sentence, leading to a contradiction. Nevertheless, the alternative theory has to use more machinery to account for non-strict negative concord languages like Spanish, exemplified in (i). The solution, in a nutshell, lies in the accommodation of the emptiness of the reference presupposition, which can happen in specific circumstances. The technical details and extensive discussion can be found in Kuhn (2022).

(i) Nadie   vino.
     NEG.WORD came
     'Nobody came.'

(ii) Principle for the expression of negation:
      Mark sentential negation on the finite verb, unless this results in a different meaning.

verb (but in the pragmatic part of the derivation), the alternative theory can predict somehow decreased acceptability of neg-words. Moreover, the emptiness of reference presupposition can be relativized to the belief or other possible worlds. Nevertheless, a full comparison of both theories with respect to Neg-Raising would have to take into account also non-Neg-Raising predicates and islands. Such configurations were not tested in the current experiment, though.

Finally, concerning the equatives, only the alternative theory of neg-words can reasonably explain why neg-words are licensed in the standard clause of equatives. First, the emptiness of reference presupposition can be satisfied in equative sentences like (19): it would require that no other fish (with the exception of brown trout) has the particular degree (on the scale of coloration) which is compatible with the truth conditions of the equative. Moreover, the split scope part of the mechanics would need to quantifier rise the neg-word over the given operator (MAX), and also, the dynamic properties of equatives would have to be checked off. Precise derivation of this must wait for future work, but the alternative approach has at least a good chance to derive the empirical asymmetry: strict negative concord languages seem to allow the neg-words in the standard clauses of equatives but do not allow (strong) NPIs there. There is a couple of other environments studied before where such licensing of neg-words goes beyond negation: the complement of prepositions like *without*, licensing of Spanish neg-words under verbs like *forbid, doubt* and *deny* (see Herburger 2001).

I will end this section via recapitulation of the predictions. As it is clear, the predictions are very much theory dependent, and for many patterns, the non-standard theories (either in the polarity or in the degree theories) are more promising than the established ones. Table 1 represents the predictions for three conditions: BAS(eline), N(eg-)R(aising) and EQ(uatives). We can expect that baseline will be acceptable for all speakers. Non-standard theories of neg-words/Ovalle & Guerzoni (2004) predict acceptance neg-words in Neg-Raising predicates. Alternative neg-word theories/Ovalle & Guerzoni (2004) and alternative degree theories of equatives/Penka (2016) predict acceptance of neg-words in equatives and rejection of strong NPIs in equatives. On the other hand, standard syntactic and degree theories predict no neg word licensing in equatives (rejection) and licensing of strong NPIs (acceptance) due to the downward entailing environment of the equative clause. And likewise for neg-raising predicates.

1.2 Research questions

We will tackle two questions. The first question in (21) is the main empirical question behind the experiment and, more generally, the search for the distinction be-

Table 1: Expected acceptability (Czech speakers)

Condition	BAS	NR	EQ
strong NPIs (standard NPIs theories)	High	High	High
neg-words (standard neg-words theories/ Zeijlstra 2004)	High	Low	Low
strong NPIs (non-standard equative theories/ Penka 2016)	High	High	Low
neg-words (non-standard neg-words theories/Ovalle & Guerzoni 2004)	High	Low	High

tween Czech strong NPIs and neg-words focused on one particular environment. The question is theoretically important since the current standard theories of equatives (like Stechow 1984, Beck 2019) build upon the analysis of *as*-clause of the equatives as downward-monotonic, therefore predicting at least grammaticality of weak NPIs and unacceptability of negation, negative quantifiers (and neg-words in languages with negative concord). This is the empirical pattern of English, but as suggested above, exactly the opposite is true for Slavic (as well as for German and other non-English) equatives. The experiment also scrutinizes the Neg-Raising. For Neg-Raising the acceptability pattern is expected to be reversed (compared to equatives): strong NPIs should be more acceptable than neg-words. But since both standard/Zeijlstra (2004) and alternative theory of neg-words/Ovalle & Guerzoni (2004) predict the same pattern in case of Neg-Raising, the first research question is focused on equatives only. Theoretical consequence of the positive answer to the first research question is an empirical support for non-standard theories of neg-words and equatives.

(21)   Question 1: Are Czech equatives acceptable with neg-words and
          non-acceptable with strong NPIs?

The second question concerns the factors of the variation in the acceptability of strong NPIs. As introduced above, previous works on variation in polarity-sensitive expressions revealed that both grammatical and demographic factors could play various roles in the speaker variation of the negative dependent expressions. Burnett et al. (2015) convincingly show that next to grammatical (syntactic factors), a proper analysis of variation should control for age and education level, since these demographic factors explain some portion of the speaker variation for negative concord (absence or presence of negative concord in Montréal

Mojmír Dočekal

French in case of Burnett et al. 2015). Since the variation in speakers and their interpretation of strong NPIs was detected in previous research (Dočekal & Dotlačil 2017, Dočekal 2020), I included age, reading time (as a measure of education level or aspiration), and region as demographic questions in my experiment. The second research question in (22) phrases exactly this research agenda targeting demographic factors. And as it will be shown, the experimental data give us precise enough results to give some answers to both questions.

(22)   Question 2: Is speaker variation of Czech strong NPIs caused by grammatical or demographic factors?

2 Experiment

The experiment aimed at answering the two research questions, (21) and (22): to test acceptability of neg-words and strong NPIs in Neg-Raising and in equatives. Next, the speaker variation was tested as well.

2.1 Methods

2.1.1 Participants & fillers

The experiment was run online on the L-Rex platform (Starschenko & Wierzba 2023). The participants were students of Masaryk University (Brno) and Charles University (Prague), and the majority of the students received credit for their participation. 105 participants filled out the experiment. The experiment included practice items to help subjects familiarize themselves with the acceptability judgment task, which was then used in the experiment itself. The experiment also included 64 fillers, half of them grammatical Czech sentences and half clearly ungrammatical sentences. Both halves of the fillers were complexity-wise similar to the items; the ungrammatical fillers included unlicensed anaphors and neg-words unlicensed by constituent negation, a.o. The exclusion rate was 66% success. 82 of the participants passed the fillers, and their data points were included in the analysis.[6]

---

[6]The criterion was whether the subject was more than 66% successful in fillers or not. One of the two anonymous reviewers asked about the details of the exclusion rate and also why the most used standard rate, 75%, was not used. The exclusion criterion was measured as follows: for each participant the difference between answers to good and bad fillers was computed – since the scale is 7-point, the difference was in the interval 0 to 6. The percentual success was then computed on the difference scale for each subject. As a sanity check, I ran the analysis with 75% exclusion rate during the revisions of the article. The descriptive and inferential statistics

2.1.2 Materials & procedure

Each questionnaire consisted of 64 items, and there were 48 randomized lists generated from the items by L-Rex. The questionnaire started with three demographic-related questions: (i) the age of the participant, (ii) the region of the participant during their first language acquisition, and (iii) their daily reading time (explained as reading time of books, journals, not looking at the screen of phones, etc.). Each participant filled out 128 trials (half items, half fillers) in the acceptability part of the experiment, which is the part reported in the present article.[7]

The experiment consisted of two parts: (i) acceptability judgment task where sentences were judged without context; (ii) acceptability judgment task where sentences were judged against a probability/scalarity manipulated context (see footnote 7). In both parts, participants judged the acceptability of sentences on a 1 to 7-point Likert scale (1 the worst – the least acceptable, 7 the best – the most acceptable). And in both parts, all conditions were crossed with two conditions: (i) NEG-WORDS, (ii) STRONG NPIS.

An example item from the experiment is in (23). As demonstrated by (23), there were three conditions: (i) BAS(ELINE), (23a), (ii) Neg-Raising, NR, (23b), (iii) equative, EQ, (23c). All three conditions were crossed with two types of negative polarity expression: (i) neg-words *žádný*, Z, (ii) strong NPIs *ani*, A. Therefore the experiment was 3x2 design, mnemonic for crossed conditions for baseline are BASA for strong NPIs, BASZ for neg-words, e.g.

(23) a. V království nezůstal   {žádný   / ani jeden} zloděj.
          in kingdom NEG.remained NEG.WORD NPI   thief
          'No thief remained in the kingdom.'

      b. Král nechce,   aby v království zůstal   {žádný   / ani jeden}
          King NEG.wants that in kingdom   remained NEG.WORD NPI
          zloděj.
          thief
          'The king doesn't want any thief to remain in the kingdom.'

      c. Zloděj ze   souostroví Qwghlm je tak šikovný jako {žádný   /
          thief   from archipelago Qwghlm is so clever   how NEG.WORD

---

remained the same as in the original analysis modulo changes in the second digit to the right of the decimal point; also the strength of the effects remained the same.

[7]The other part of the experiment was an acceptability judgment task with a probability/scalarity manipulated. The results of the second part are not reported in the present article due to space reasons.

Mojmír Dočekal

ani jeden} zloděj.
NPI   thief
'The thief from the Qwghlm archipelago is as clever as any other
thief.'

2.2 Predictions

All the discussed theories predict that neg-words and strong NPIs will be well
accepted in the baseline condition, BAS. The condition is present in the experi-
ment to check how much worse the other two conditions will be compared to
the baseline.

On the contrary, the standard theory of neg-words/Zeijlstra (2004) and the
alternative theory of neg-words/Ovalle & Guerzoni (2004) differ in their predic-
tions for equatives, EQ: the standard theory predicts neg-words to be not accept-
able in EQ while the alternative theory is compatible with their acceptability in
EQ. The standard theory of equatives/Rullmann (1995) predicts the acceptability
of strong NPIs in EQ while the alternative theory of equatives/Penka (2016) pre-
dicts their unacceptability.

Finally, both theories of neg-words predict the decreased acceptability of neg-
words in NR and much higher acceptability of strong NPIs in NR.

Concerning the speaker variation, there are no theory-specific predictions. But
since the speaker variation was observed in previous research, it is expected that
the variation will be observed in the current experiment too. The demographic
factors were included in the experiment to test whether the variation is caused
by grammatical or demographic factors.

2.3 Results

The descriptive statistics results can be seen in the graph of acceptance, including
error bars, in Figure 1. [8] As can be seen with the naked eye, both expressions are
nearly at the ceiling in BAS, but their acceptability in the two other conditions is
reversed. While neg-words are much more acceptable in equatives (EQZ), strong
NPIs are preferred in Neg-Raising contexts (NRA).[9]

---

[8]The graph of acceptance in Figure 1 uses the standard Cartesian coordinate system with the
y-axis origin at 0. That does not mean that the response scale was 0 to 7 but is simply the
default behavior of the GGPLOT2 R package, see Wickham (2016).

[9]One of the two anonymous reviewers asked whether unacceptable fillers were judged worse
than neg-words (or strong NPIs) in Neg-Raising contexts. The answer is yes, the ungrammati-
cal fillers median acceptability was 1 while the median acceptability of neg-words and strong
NPIs was 2.
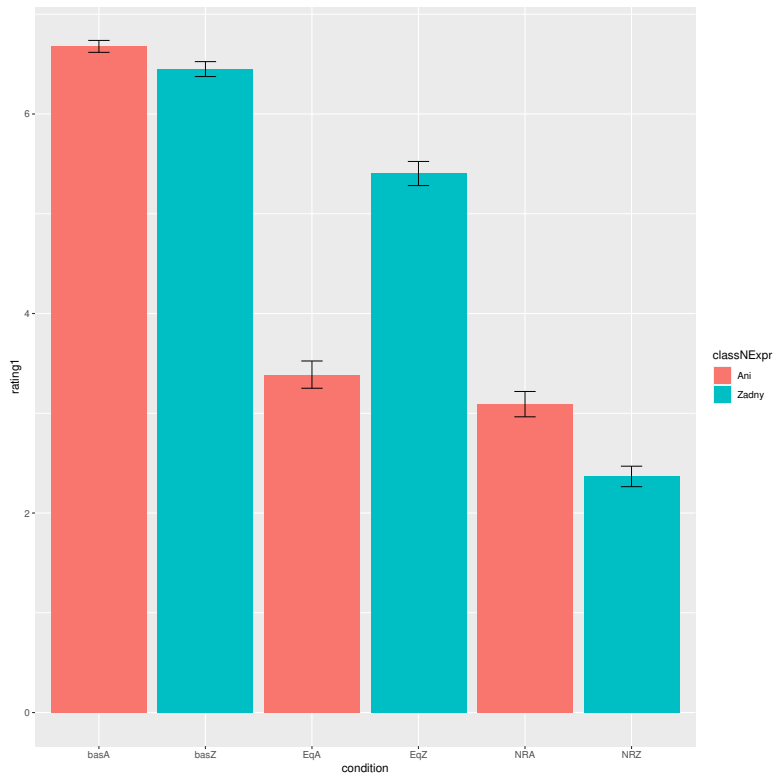
Figure 1: Graph of acceptance (+error bars)
conditions: ʙᴀseline, ᴇǫuative, NegRaising
epxressions: ᴀni (strong NPI), zádný (neg-word)

2.3.1 Inferential statistics

The Bayesian hierarchical random-effects model with default priors was fit using
the R package ʀsᴛᴀɴᴀʀᴍ (Goodrich et al. 2022): the dependent variable was the
subject's response; the independent variables were: (i) environment (ʙᴀs, ᴇǫ, ɴʀ),
(ii) type of the polarity-dependent expression (ᴀ, ᴢ), and their interaction; the ref-
erence level was ʙᴀs, ᴀ. The baseline was selected as the condition which should
be uncontroversially accepted by speakers, which indeed was the case since no
main effect was positive against the baseline. The model included random effects
for both subject and item intercepts.

The model was fit to the data and we found that (i) baseline was very well
accepted (Intercept = 6.67, 95% C(redibility) I(nterval)= [6.38, 6.95]), there is no
distinction between neg-words and strong NPIs in it and sine qua non, both ex-

pressions are acceptable to the same extent (posterior main effect in the form of median and 95% CI: $\hat{\mu} = -0.20$, CI $= [-0.47, 0.08]$), (ii) neg-words were much better accepted in equatives than strong NPIs (the positive interaction of EQ by z: $\hat{\mu} = 2.18$, CI $= [1.81, 2.58]$ – against the reference level), (iii) strong NPIs were preferred in Neg-Raising (the negative interaction of NR by z: $\hat{\mu} = -0.53$, CI $= [-0.91, -0.14]$ – against the reference level).[10] The results are also supported by the results of R(egion) O(f) P(ractical) E(quivalence), ROPE: only z is not significant, since it is 23% in ROPE. For all medians, confidence intervals, and ROPE percents, see Table 2, all percents of ROPE are computed for the interval $[-0.10, 0.10]$. Medians, 95% credibility intervals, and ROPE are also visually represented by the graph in Figure 2. Notice that in Figure 2, is (as it is usual in Bayesian modeling) the reference level condition (BAS, A) coded as 0 of the x-axis and each condition (or interaction of condition) has its own y-line (with the distribution and median); the credibility of a condition can be visually inspected via observation of the condition including 0 (clearly z, e.g.) or differing from it either positively (the interaction between EQ by z, e.g.) or negatively (NR, e.g.).

2.3.2 Demographic factors

Next, three Bayesian generalized mixed linear models were fitted to detect the effects of demographic factors inhibiting or prohibiting acceptability. This was important since previous work (see Burnett et al. 2015 and Burnett et al. 2018, a.o.) revealed that both grammatical and demographic factors are at play when negative polarity variation is linguistically studied. As a reminder, the experiment included three demographic questions: region, age, and daily reading time. The last factor was used as a proxy for investigating the educational level. The selection of factors was influenced by the previous work on variation in negative dependent expressions: Burnett et al. (2015) have shown that both age, education level, and location of the speakers can have an impact on the variation. And since in the previous experimental work it was revealed that there are idiolects of Czech speakers interpreting the strong NPI *ani* (Dočekal & Dotlačil

---

[10]As one of the two anonymous reviewers correctly points out, NR (whether it includes strong NPI or neg-word) is pretty badly acceptable – see the row of NR in Table 2. I agree, but this seems to be the case generally; the results of my experiment resonate with the experimental finding from Dočekal & Dotlačil (2016a), where the strong negative effect between the acceptable baseline and tested Neg-Raising was observed. The reasons for this negative effect are unclear, but see Alexandropoulou et al. (2020) for some other environments where including NPIs leads to a strong decrease in acceptability even if such an effect is theoretically unexpected.

Table 2: Bayesian model and its posterior distribution for the experiment

Parameter			
	Median	CI	% in ROPE
Intercept	6.67	$[6.38, 6.95]$	0%
EQ	-3.27	$[-3.53, -3.00]$	0%
NR	-3.57	$[-3.86, -3.30]$	0%
Z	-0.20	$[-0.47, 0.08]$	23.08%
EQ:Z	2.18	$[1.81, 2.58]$	0%
NR:Z	-0.53	$[-0.91, -0.14]$	0%
Random effects			
	Name	SD	
subject	Intercept	0.57	
item	Intercept	0.35	

2017, Dočekal 2020), I included the three mentioned demographic factors to detect whether the variation can be traced down to some extralinguistic sources eventually. Nevertheless, it has to be said that the pool of subjects was rather homogeneous, consisting mainly of university students. Therefore, at least the education level results should be taken with a grain of salt and mainly as a first step in the general description of polarity items variation in Slavic languages. Also, for this reason, I move the inferential statistics into footnotes and describe the main outcomes in terms of descriptive statistics.

Let us start with AGE. Descriptively, AGE ranged from 19 to 71, with a median = 23, mean = 25.59, and sd = 9.47. The age was first z-transformed and then plugged in as the third interaction variable in the Bayesian model (next to the two conditions, z and BAS/NR/EQ environment).[11] But the model did not confirm any effect on overall acceptability or any particular age-related inhibition or prohibition of construction or negative dependent expression.

Next, REGION was more varied than AGE, where the data points from the first to the third quantile were in the range of 21 to 25. But since I did not control for the specificity of the values entered into the form, the answers ranged from

---

[11]The model revealed that the acceptability overall was not affected by age at all (main effect of AGE: $\hat{\mu} = 0.01$, CI = $[-0.25, 0.28]$, ROPE: 58.00% for the $[-0.10, 0.10]$ interval). There was also no significant interaction with any single or pair of conditions. The lowest ROPE was 30.55% for the three-way interaction between EQ:Z:AGE. All other interactions had an even bigger portion in ROPE and were also less significant.

Mojmír Dočekal
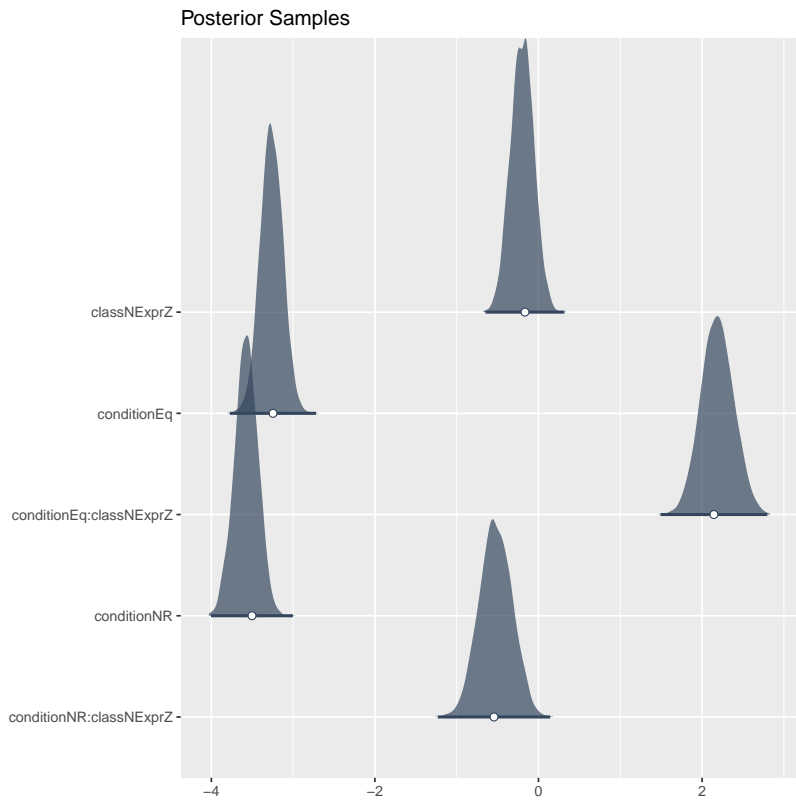
Posterior Samples

Figure 2: Graph of posterior samples with ROPE (for the experiment)
conditions: BAseline, EQuative, NegRaising
expressions: Ani (strong NPI), zádný (neg-word)

city-specific to region-specific. For this reason, I aggregated all the answers into the discrete factor with two levels: MORAVIAN, NONMORAVIAN. 67% of subjects entered as their region NONMORAVIAN, the remaining 33% identified themselves as being from Moravia. Again the factor REGION was used as the third interaction variable in the Bayesian model.[12] Overall, the REGION did not increase or decrease acceptability, but there is some anecdotal evidence for higher acceptance of Neg-Raising in the Czech part of the population. Nevertheless, the interaction effect is

---

[12]The main effect of the region was not credible ($\hat{\mu} = 0.33$, CI = $[-0.18, 0.88]$, ROPE: 14.79% for the $[-0.10, 0.10]$ interval) but this time there was very weak evidence coming from interactions. Namely, there seemed to be a slight tendency for higher acceptance of Neg-Raising in the non-Moravian part of the Czech Republic (the interaction NR:MORAVIAN: $\hat{\mu} = -0.61$, CI = $[-1.28, 0.04]$, ROPE: 3.89% for the $[-0.10, 0.10]$ interval). All other interactions with REGION were less significant.

so weak that I doubt there is any genuine linguistic Neg-Raising isogloss between Czech dialects.

The last demographic factor was reading time. As hinted above, the factor was used to get information about education or education aspirations. The answers (converted to hours) ranged from 0 to 10 hours, with 1 hour as the median, 1.43 hours as the mean, and the range of first and third quantiles being 1 hour and 2 hours, respectively. Similarly to AGE, data points are centered around the mean with a small standard deviation, 1.26, and few outliers. As in the case of REGION, I recorded the continuous variable as a factor READINGTIME with two levels: OVER1HOUR,UNDER1HOUR dividing the sample according to the median value of reading time. The result was two nearly proportional halves: 52% of the subjects claimed that their daily reading time is under 60 minutes, and the remaining 48% entered that they read more than one hour. The third demographic factor (READINGTIME) was plugged into the Bayesian model as an independent (interaction) variable.[13] The modeling results show that there is some weak evidence for the positive correlation between reading time and the acceptance of the Neg-Raising construction: subjects who claimed to read more were more accepting of the Neg-Raising construction. Such a tendency is intuitively plausible but does not say anything linguistically important about the constructions and polarity-dependent expressions tested in the experiment.

Let us summarize: the design of the experiment and three demographic questions did not reveal any important information concerning the demography-related variation in polarity constructions of Czech speakers. Two weak effects can be interpreted as clues about region and education-level variation concerning Neg-Raising. Still, there seems to be nothing significant in the variation of *ani* vs. *žádný* in the studied constructions. So, whatever speaker variation (in the usage of *ani*) we will discuss further, it seems not to be related to age, region, or education level as revealed by the sample of the experiment (in this respect, the results of the experiment are different from the previous work on speaker variation in polarity dependent expressions, like Burnett et al. 2015, 2018).

---

[13]And again, as with two previous demographic factors, the main effect of READINGTIME was not credible ($\hat{\mu} = -0.13$, CI $= [-0.63, 0.39]$, ROPE: 28.89% for the $[-0.10, 0.10]$ interval). And similarly to REGION, there was one weakly credible interaction: subjects claiming to read more than average (over 60 minutes daily) were accepting Neg-Raising constructions more (NR:OVER1HOUR interaction: $\hat{\mu} = 0.66$, CI $= [0.04, 1.27]$, ROPE: 1.24% for the $[-0.10, 0.10]$ interval). All other interactions were much less credible.

Mojmír Dočekal

2.3.3 Correlations

The Bayesian model revealed that both non-baseline environments (NR and EQ) were less accepted by speakers, but there was no difference between *ani* and *žádný* in terms of main effects. Nevertheless, speakers accepted in equatives much more neg-words than strong NPIs (the strong and only one positive interaction effect). But speakers are also inclined to reject neg-words in Neg-Raising against strong NPIs (the negative interaction effect between z and NR). The theoretical consequences of these findings will be discussed below, but let us turn to another kind of variation, this time not demographic.

The first important thing to note is that all speakers agreed on their high acceptance of baseline, and in this condition, they accepted neg-words and strong NPIs indistinguishably. But this acceptance of both polarity expressions diverged in the two other conditions. Namely, some speakers rate *ani* high in equatives (unlike the main thrust of speakers, recall that the strong positive interaction between z and EQ) but also reject it in NegRaising (again going against the overall acceptance of strong NPIs there: the negative interaction between z and NR). And vice versa, subjects who reject strong NPIs in equatives (behaving according to the negative interaction effect) accept strong NPIs in Neg-Raising (again verifying the negative interaction effect). In both cases, we observe a negative correlation between the acceptability of neg-words/strong NPIs in the two environments, equatives, and Neg-Raising. One way to understand this reversed correlation is to assume that the first kind of speakers (those who accept *ani* in equatives) treats *ani* more like neg-word and not strong NPI. The rest of the sample (majority, in fact) treats *ani* as strong NPI and therefore accepts it under Neg-Raisers and rejects it in equatives.

Therefore, post-hoc correlation statistics were run and are reported below. Notice, though, that the correlation is post-hoc in the sense of interpreting subclusters of speakers (let us say idiolects in linguistic terms) but not in the sense of avoiding the Type I error (mistakenly rejecting the null hypothesis) since the null hypothesis is not important part of Bayesian statistics, instead the bulk of the inference statistics in Bayesian framework is posterior distribution which represents the probability of the parameters of a model given the data and which was reported here in §2.3.1. The motivation for the correlation analysis comes both from the observed variation introduced above and from the previous work on Czech strong NPIs (Dočekal & Dotlačil 2017, Dočekal 2020) where it was observed that there are idiolects of Czech speakers concerning their strong NPIs interpretation.

The way the correlations were checked statistically is the following. First,

the acceptance of conditions was z-transformed (by subject). And then, such z-transformed variables were checked for correlations across conditions. And indeed, there is a strong negative correlation between the acceptability of *ani* in equatives and its acceptability under Neg-Raising predicates (Pearson's product-moment correlation: $t = −5.93$, $p < 0.001$). The correlation graph is in Figure 3. This means that we can identify two groups of speakers: (i) speakers who accept *ani* under Neg-Raisers and reject it with equatives (top left section in the Figure 3), (ii) speakers who accept *ani* in equatives and reject it under Neg-Raisers (the bottom right part of Figure 3). But crucially, no speakers are accepting both conditions (the empty top right corner) nor speakers who would reject both conditions (the empty space in the bottom left part). And also, there is no correlation between the acceptability of *ani* acceptability in the baseline and equatives, just pure noise as can be seen in Figure 4. This correlation of *ani* between NR and EQ resonates the previous work (Dočekal & Dotlačil 2017) where similar correlations were found (for *ani*) in the case of probability manipulated conditions and Neg-Raising.
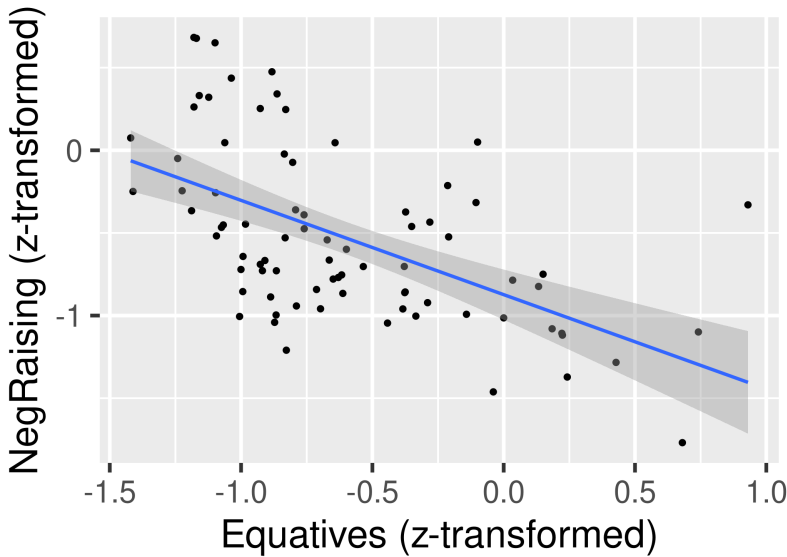
Figure 3: Correlations between Equatives and Neg-Raising for *ani*

Mojmír Dočekal
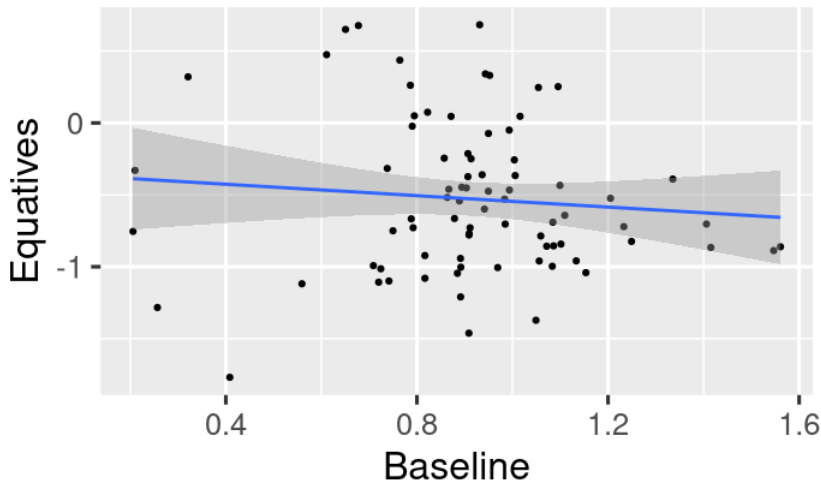
Figure 4: Correlations between Equatives and Baseline for *ani*

2.3.4 Discussion

The reported experimental results suggest that Czech strong NPIs are (for most speakers) accepted under Neg-Raising predicates (NR was credible negative effect but the negative interaction effect NR by z shows that participants prefer strong NPIs in NR) and rejected in equatives (the main effect of EQ is negatively strong but there is a very strong positive interaction EQ by z indicating robust preference for neg-words in equatives). For Czech neg-words, the opposite is true: most speakers accept them in equatives but reject them under Neg-Raising predicates. In the case of *ani* (strong NPI), there is a speaker variation, and some subset of speakers treats it more like a neg-word; nothing similar was found for neg-words. The correlation discussed in §2.3.3 suggest that there are two kinds of speakers: (i) speakers accepting *ani* in NR and rejecting it in EQ (strong NPI treatment of *ani*); (ii) speakers accepting *ani* in EQ and rejecting it in NR (the speakers who use *ani* more like neg-word). Moreover, the speaker variation does not seem to be derivable from demographic factors (or at least not from the demographic factors controlled in the experiment).

3 Theoretical consequences

The nature of this article is mainly experimental. For this reason and the obvious constraints of space, the consequences of the analyzed experimental data will be

discussed only to a limited extent.

3.1 Application: strong NPIs

Now, let us demonstrate how the approaches introduced in §1.1 can be applied to the data gathered in the experiment. First, concerning the strong NPIs (Czech *ani*), I will show how the *even*-approach fits the Czech data from the experiment, starting with the baseline. For a reminder, the Table 2 summarizes predictions of the theories for the experiment. The results of the experiment confirm more the non-standard theories of neg-words/Ovalle & Guerzoni (2004) and non-standard theories of equatives/Penka (2016).

As discussed above, *ani* comes with the *even*-presuppositions, namely the scalar and the additive. In the formal terms, this is cached as an association of *ani* with covert *even* scoping at the propositional level; see a schematic representation of the baseline in (24) and its logical form in (24a). Since the sentence is negated, the entailment between numerals is reversed by negation: $\neg(\llbracket$ one thief $\rrbracket ...) \vDash \neg(\llbracket$ two thieves $\rrbracket ...)$. The alternatives to the prejacent come from the focused numeral since *ani* can associate with nouns, clauses, etc., but as usually in Slavic languages, it associates mostly with its sister node (the numeral in (24)). Because the entailment is reversed, the *even*-approach predictions agree with the downward-entailing explanation. Moreover, since *ani* is a strong NPI, it requires DE/*even*-presuppositions to be satisfied both in truth conditions (configurationally β in (24a)) but also at the level of non-at-issue meaning (where the exhaustifier, silent *even*, scopes: α in (24a)). Next, we have to check both presuppositions as schematically formalized in (24d) and (24e), which is also fulfilled. Theoretical prediction of the standard account/Gajewski (2011) then agrees with the high baseline acceptability of *ani*: even the inferential statistic baseline median intercept was 6.67 on the 7-point Likert scale.

(24)   Ani $[\text{one}]_F$ thief neg-remained in the kingdom.

    a. $[_\alpha$ *(even)* $[_\beta \neg [_\gamma$ ani one thief remained in the kingdom $]]$ $]$

    b. TC (in β) DE: ☒

    c. non-at-issue (in α) DE: ☒

    d. scalar presupposition of (even): $\rightarrow \neg$(two thieves remained), $\neg$(three thieves remained), ...: ☒

    e. additive presupposition: $\neg$(two thieves remained) $\vee \neg$(three thieves remained), ...: ☒

Continuing now to the Neg-Raising condition. Neg-words were in NR less accepted than strong NPIs; although the effect was not particularly strong, it was still significant. The theories of Neg-Raising (Gajewski 2007, Romoli 2013) predict that for the Neg-Raising predicates the scope of root negation ends (via presupposition or implicature calculation) in the embedded clause. Schematically we can formalize the important ingredients of the LF for the experiment NR conditions as in (25). From the point of view of the standard theory of strong NPIs and the *even*-presuppositions, the logical form of the embedded clause is the same as in the case of the baseline, therefore the explanation of the NR acceptability is the same as in the baseline case. What differs is the actual acceptance by speakers (much lower in the case of Neg-Raising than in the baseline), but as discussed in footnote 6, in this respect, our experiment nearly replicates the previous findings in case Neg-Raising and NPIs generally. The second difference concerns the diverging acceptability of strong NPIs and neg-words in Neg-Raising contexts (both kinds of expressions were indistinguishably well accepted in the baseline). Still, this point will be discussed in §3.2.

(25)   The king does ~~not~~ want $[_{\alpha}$(*even*) $[_{\beta}\neg$ [ani [one]$_F$ thief remained in the kingdom]]].

Finally, the standard theory of equatives predicts that strong NPIs can be licensed in the standard clause. This was argued before to be wrong for German and Romance languages (Krifka 1992, Penka 2016) and also clashes with the intuition of Czech speakers for whom strong NPIs were much less accepted than neg-words in the standard clauses of equatives. A proper investigation of Slavic equatives must wait for future work but let us take the first steps in this direction. If we take seriously the morpho-syntax of Slavic equatives and follow the theoretical hints from Penka (2016), it is possible to model the equative conditions from the experiment as in (26). First, since the morpho-syntax of Czech (and Slavic, Romance, German) equatives is built upon the correlatives, anaphoric *so* (Czech *tak* 'so', see (23c) for full glosses) picks up the referent of the definite degree description (index 2). The definitive degree description is yielded by the MAX_{inf} operator, similar to free relatives. And since free relatives are known not to license NPIs (Jacobson 1995), the correlative standard equative clause is expected not to be a good environment for NPIs. Therefore we can predict, if we assume this non-standard but well-motivated theory of equatives, why in Czech equatives, strong NPIs are not accepted: since β is most probably not DE, even if the presuppositions of silent *even* in α were satisfied (MAX_{inf} makes the standard of equatives most probably non-monotonic), the licensing of strong NPIs is not satisfied. The much better acceptance of neg-words in equatives will be discussed in §3.2.

(26) The thief from the Qwghlm archipelago is so^2 clever $[_\alpha$ (even) $[_\beta MAX_{inf}$ how$_1$ ani [one]$_F$ thief t_1 ~~clever is.~~]]^2

3.2 Neg-words

In this section, I will present the application of theories introduced in §1.1.4 to the results of the experiment.

The baseline can be explained easily both in the standard theory of neg-words/ Zeijlstra (2004) and in the alternative theory of neg-words/Ovalle & Guerzoni (2004). I discuss just the alternative theory explanation here. In a negative sentence, like schematic (27), the truth conditions (indefinite descriptions) of the neg-word and its presupposition agree: the indefinite description is under the scope of negation, and the presupposition of the emptiness of the discourse referent is compatible with the truth conditions. But in a positive minimal pair sentence, like (28), the existential quantification over the discourse referent and the presupposition of its reference emptiness would clash. Therefore, the positive minimal pair sentence is predicted to be unacceptable. This is exactly what we observe in the experiment: the baseline is very well accepted, but the positive minimal pair would be rejected.

(27) neg-word thief neg-remained in the kingdom.
    a. $[\neg[\exists x[\text{THIEF}(x) \wedge \text{REMAINED}(x)]\,]]\,\wedge 0_x$

(28) neg-word thief remained in the kingdom.
    a. $[\exists x[\text{THIEF}(x) \wedge \text{REMAINED}(x)]\,]\,\wedge 0_x$ $\qquad\qquad\qquad\bot$

Turning now to Neg-Raising, neg-words were less accepted in NR than strong NPIs. From the perspective of the alternative theory of neg-words/Ovalle & Guerzoni (2004), we should expect that this would align with the locality constraints: the neg-words have to quantifier raise over its licensors, creating a split-scope logical form. But since the root negation ends in the embedded clause in the case of Neg-Raising predicates, as discussed above, the schematic logical form for Neg-Raising conditions can be rendered as (29). This reasoning explains the difference between relatively well licensing (both for neg-words and strong NPIs) Neg-Raising predicates and much worse non-Neg-Raising predicates (verbs of causation or communication) in such configurations like (29) – see Dočekal & Dotlačil (2016a) for experimental data and analysis. Recall that experimental results showed a slight preference for strong NPIs in this condition. The difference can be captured as follows: since the split scope relies on syntactic mechanisms

and the negation ends in the embedded clause in the pragmatic part of the derivation, there is a slight timing issue which maybe can be coerced. A similar kind of explanation can be retold in the standard approach/Zeijlstra (2004), too: the syntactic licensing of features should proceed before the pragmatic mechanisms like presuppositions and implicatures calculation. But in both kinds of explanation, a lot hinges upon the assumptions about the architecture of grammar and many other assumptions. Moreover, it is not totally clear how to linguistically interpret the effects from the experiment – as Table 2 shows, both kinds of expressions (strong NPIs and neg-words) lead to a dramatic acceptability decrease (−3.57), but there is the slight preference for strong NPIs (the negative interaction of neg-words with Neg-Raising: −0.53).

(29) The king wants ~~not~~ that [neg-word thief$_2$ ¬ [t_2 remained in the kingdom]].

Finally, concerning the equatives: (i) as tested in the experiment, neg-words are, but strong NPIs are not acceptable in the complement clause of the Czech equatives, (ii) adding to this, verbal negation is not acceptable either – see (30). Especially the high acceptability of neg-words is surprising from the perspective of English since negative quantifiers are distinctly odd in this position (in comparatives, but as discussed in 1.1.3, the expectations are – in the standard theories – similar for comparatives and equatives), see (31) from Gajewski (2008).

(30) Petr je tak chytrý jak {nikdo   jiný / *Marie ne / *ani   jeden}.
      Petr is so smart as   NEG.WORD else Mary not STRONG.NPI
      'Petr is as smart as anyone.'

(31) *Mary is taller than no boys are.

The ambition of this article is not to solve the above-mentioned theoretical puzzles. But let us at least indicate where a possible solution can be. The experimental results show that Czech neg-words are very much accepted in the complement clause of equatives, while strong NPIs are degraded there. Moreover, intuitions and preliminary results from the follow-up experiment suggest that weak NPIs are not acceptable in the equatives either, following the German and Romance data discussed in Krifka (1992), Penka (2016). One possible explanation is that Czech complement clauses of equatives are not downward-monotonic in either truth conditions or non-at-issue meaning. But for some reason, the emptiness of the neg-words' discourse referents presupposition can be easily satisfied in this environment. Consider (32): the emptiness of discourse referents presupposition here would be that there is no such thief with the degree of cleverness *d*

which would exceed the degree of Qwghlm archipelago thief. Such presupposition is plausible, and more generally, it can be said that Czech equatives are one of the rare environments where the neg-words can be licensed by expressions not-including negation (similarly to Spanish verbs like *forbid, doubt* or *deny*).

(32)   The thief from the Qwghlm archipelago is so^2 clever [MAX_{inf} how_1 neg-word thief t_1 ~~clever is.~~]^2

This can be compatible with Penka's (2016) suggestion to replace the MAX operator in analyzing English equatives with a different relation on the degrees, MAX_{inf} discussed in §3.1. But so far, I consider the evidence to be non-conclusive regarding the monotonic properties of Czech complement clauses of equatives; they are not downward entailing for sure, but the resulting two possibilities: upward entailing or non-monotonic, are still open. From a theoretical standpoint, I agree with Penka (2016) that current degree theories of equatives do not hold the cross-linguistic water. And in the same direction, it is clear that a purely syntactic approach to neg-words faces big trouble when posed with the acceptability of neg-words in equatives. No matter how the cross-linguistically feasible degree theory of equatives will look like (e.g., using MAX_{inf} as suggested by Penka 2016), there is clearly no room for sentential negation operator in its version for Slavic (and Romance) equatives since then the weak or strong NPIs would be admissible there, contrary to facts. Concluding this section, merging the non-standard theory of neg-words/Ovalle & Guerzoni (2004) and the non-standard theory of equatives/Penka (2016) is promising as a theoretical explanation for the high acceptability of neg-words in the standards of Czech equatives and the very low acceptance of strong NPIs there.

4 Summary

The findings of the current study provide answer to research question 1, repeated below as (33). The experiment confirmed the robust acceptability of neg-words in the standard clause of equatives. This can be explained deriving from the neg-words presupposition relativized to the set of degrees introduced in the main clause if we follow the alternative theory of neg-words/Ovalle & Guerzoni (2004). Neg-words in examples like (30) are well accepted since, in this configuration, the presupposition does not require total emptiness of reference, just the emptiness of reference for such discourse referents whose degree would exceed the degree of the subject. Next, the strong NPI unacceptability in Czech equatives is a direct consequence of the Czech equatives complement clauses not being downward

monotonic. The sub-answers to research question 1 are in (34). The results of the experiment bring empirical support for alternative theories of neg-words/Ovalle & Guerzoni (2004) and alternative theories of equatives/Penka (2016).

(33) Question 1: Are Czech equatives acceptable with neg-words and non-acceptable with strong NPIs?

(34) The non-standard theories of negative concord and equatives give promising answers:

   a. The semantic/pragmatic theory of neg-words allows the presupposition of discourse emptiness to be satisfied (relativized to degrees of the main clause).
   b. The complement clause of the Czech equatives is not downward entailing.

Now, research question 2 is repeated below as (35). Some speaker variation was observed. Recall that for some speakers *ani* behaved more like a neg-word. Nevertheless, it is not likely that the variation can be related to demographic factors such as age, region, or daily reading time. But there is one way to theoretically explain the variation: we can assume that both neg-words and strong NPIs in Czech come with presuppositions, the emptiness of discourse referents for neg-words, and the scalar presupposition for strong NPIs. Then the flux from strong NPIs to neg-words can be theoretically cashed out as follows: speakers substitute the scalar presupposition with the emptiness of discourse referents presupposition. Speculatively, we can try to explain the one-way direction in terms of economy: the scalar presupposition needs a covert exhaustifier, but the emptiness presupposition does not. Therefore it is less costly and more attractive for speakers who oscillate between the two presuppositions. For this reason, there is no speaker variation concerning neg-words: adopting the scalar presupposition would mean a less economical logical representation. Why the flux is unrelated to the demographic factors is an issue for future work. The answers are summarized in (36).

(35) Question 2: Is speaker variation of Czech strong NPIs caused by grammatical or demographic factors?

(36) The speaker variation is explainable as shifting from the scalar to the emptiness of the DR presupposition (in case of *ani jeden* 'even one').

   a. Social factors don't seem to play a role in this shift.

Let us end this section with some open questions. The first of them concerns the locality constraints on neg-words licensing. The alternative theory of neg-words/Ovalle & Guerzoni (2004) predicts that the neg-word locality should approximate the quantifier raising. Only syntactic islands (such as relative clauses) should be hard limits for both neg-words licensing and quantifier raising. The syntactic literature on the topic of Slavic quantifier raising seems to argue for a possibility of overt movement (out of non-island clauses) but obligatory reconstruction (see Neeleman & Titov 2009, a.o.). Still, the experimental research in this direction seems limited to mono-clausal conditions (see, e.g., Ionin & Luchkina 2018). So there is space for future research in this direction, and only then can we conclude whether the locality constraints between quantifier raising and neg-words licensing coincide. Another open question concerns the cross-linguistic variation in neg-words licensing: in Romance languages, neg-words are licensed in *before*-clauses and under *doubt*-type predicates; in Slavic languages, this is not the case. The alternative theory of neg-words/Ovalle & Guerzoni (2004) of neg-words predicts that this should follow from the different presupposition projection properties in the two types of languages. Whether this is true remains again a question for future work.

Abbreviations

1	first person	NEG	negation
AUX	auxiliary	NPI	Negative Polarity Item
EQ	equative	NR	Neg-Raising
COMP	comparative	SG	singular

Acknowledgments

References

Alexandropoulou, Stavroula, Lisa Bylinina & Rick Nouwen. 2020. Is there *any* licensing in non-DE contexts? An experimental study. In Michael Franke, Nikola Kompa, Mingya Liu, Jutta L. Mueller & Juliane Schwab (eds.), *Proceedings of Sinn und Bedeutung 24*, 35–47. University of Konstanz: Gesellschaft für Semantik. DOI: 10.18148/sub/2020.v24i1.851.

Mojmír Dočekal

Beck, Sigrid. 2019. Comparison constructions. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: Lexical Structures and Adjectives*, 415–476. Berlin, Boston: De Gruyter Mouton. DOI: 10.1515/9783110626391-013.

Błaszczak, Joanna. 2001. *Investigation into the interaction between the indefinites and negation*, vol. 51 (Studia Grammatica). Berlin: Akademie Verlag.

Burnett, Heather, Hilda Koopman & Sali A. Tagliamonte. 2018. Structural explanations in syntactic variation: The evolution of English negative and polarity indefinites. *Language Variation and Change* 30(1). 83–107. DOI: 10.1017/s0954394517000266.

Burnett, Heather, Mireille Tremblay & Hélène Blondeau. 2015. The variable grammar of negative concord in Montréal French. *University of Pennsylvania Working Papers in Linguistics* 21(2). 3. https://repository.upenn.edu/handle/20.500.14332/45068.

Chemla, Emmanuel, Vincent Homer & Daniel Rothschild. 2011. Modularity and intuitions in formal semantics: The case of polarity items. *Linguistics and Philosophy* 34(6). 537–570. DOI: 10.1007/s10988-012-9106-0.

Chierchia, Gennaro. 2019. Factivity meets polarity: On two differences between Italian versus English factives. In Daniel Altshuler & Jessica Rett (eds.), *The semantics of plurals, focus, degrees, and times*, 111–134. Kluwer: Springer. DOI: 10.1007/978-3-030-04438-1_6.

Chisholm, Hugh (ed.). 2010. *Encyclopaedia Britannica*, vol. 5. Illinois: Project Gutenberg. https://www.gutenberg.org/files/33052/33052-h/33052-h.htm.

Crnič, Luka. 2011. *Getting even.* Cambridge, MA: MIT. (Doctoral dissertation).

Crnič, Luka. 2014a. Against a Dogma on NPI licensing. In Luka Crnič & Uli Sauerland (eds.), *The art and craft of semantics: A Festschrift for Irene Heim*, vol. 1, 117–145. Cambridge, MA: MIT Press. https://semanticsarchive.net/Archive/jZiNmM4N/Crnic.pdf.

Crnič, Luka. 2014b. Non-monotonicity in NPI licensing. *Natural Language Semantics* 22(2). 169–217. DOI: 10.1007/s11050-014-9104-6.

Djärv, Kajsa, Jérémy Zehr & Florian Schwarz. 2018. Cognitive vs. emotive factives: An experimental differentiation. In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of Sinn und Bedeutung*, vol. 21, 367–386. University of Konstanz: Gesellschaft für Semantik. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/143.

Dočekal, Mojmír. 2020. N-words and NPIs: Between syntax, semantics, and experiments. In Franc Marušič, Petra Mišmaš & Rok Žaucer (eds.), *Advances in formal Slavic linguistics 2017*, vol. 3, 15–49. Berlin: Language Science Press. DOI: 10.5281/zenodo.3763740.

Dočekal, Mojmír & Jakub Dotlačil. 2017. *Strong NPIs vs. n-words: acceptability experiment in Czech*. Presentation at Sinn und Bedeutung 22. Berlin. https://sinnundbedeutung22.wordpress.com/wp-content/uploads/2017/09/all_abstracts.pdf.

Dočekal, Mojmír & Jakub Dotlačil. 2016a. Experimental evidence for neg-raising in Slavic. *Linguistica* 56(1). 93. DOI: 10.4312/linguistica.56.1.93-109.

Dočekal, Mojmír & Jakub Dotlačil. 2016b. When is not-believing believing that not? In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of Sinn und Bedeutung*, vol. 21, 387–398. University of Konstanz: Gesellschaft für Semantik. https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/144.

Gajewski, Jon. 2008. More on quantifiers in comparative clauses. In Tova Friedman & Satoshi Ito (eds.), *Semantics and linguistic theory*, vol. 18, 340–357. New York: Linguistic Society of America. DOI: 10.3765/salt.v0i0.2494.

Gajewski, Jon. 2016. Another look at NPIs in definite descriptions: An experimental approach. In Pierre Larrivée & Chungmin Lee (eds.), *Negation and polarity: Experimental perspectives*, 307–327. Kluwer: Springer. DOI: 10.1007/978-3-319-17464-8_13.

Gajewski, Jon R. 2011. Licensing strong NPIs. *Natural Language Semantics* 19(2). 109–148. DOI: 10.1007/s11050-010-9067-1.

Gajewski, Jon Robert. 2007. Neg-raising and polarity. *Linguistics and Philosophy* 30(3). 289–328. DOI: 10.1007/s10988-007-9020-z.

Goodrich, Ben, Jonah Gabry, Imad Ali & Sam Brilleman. 2022. *Rstanarm: Bayesian applied regression modeling via Stan*. R package version 2.21.3. https://mc-stan.org/rstanarm/.

Herburger, Elena. 2001. The negative concord puzzle revisited. *Natural Language semantics* 9(3). 289–333. https://www.jstor.org/stable/23748812.

Hoeksema, Jack. 2012. On the natural history of negative polarity items. *Linguistic Analysis* 38(1). 3–33.

Homer, Vincent. 2021. Domains of polarity items. *Journal of Semantics* 38(1). 1–48. DOI: https://doi.org/10.1093/jos/ffaa006.

Ionin, Tania & Tatiana Luchkina. 2018. Focus on Russian scope: An experimental investigation of the relationship between quantifier scope, prosody, and information structure. *Linguistic Inquiry* 49(4). 741–779. DOI: 10.1162/ling_a_00288.

Jacobson, Pauline. 1995. On the Quantificational Force of Free Relatives. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara Partee (eds.), 451–486. Dordrecht: Kluwer. DOI: 10.1007/978-94-017-2817-1_15.

Mojmír Dočekal

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý & Vít Suchomel. 2014. The Sketch Engine: ten years on. *Lexicography* 1. 7–36. DOI: 10.1007/s40607-014-0009-9.

Krifka, Manfred. 1992. Some remarks on polarity items. In Dietmar Zaefferer (ed.), *Semantic universals and universal semantics*, 150–189. Berlin: Foris. DOI: 10.1515/9783110870527-009.

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25(3-4). 209–257.

Kuhn, Jeremy. 2022. The dynamics of negative concord. *Linguistics and Philosophy* 45(1). 153–198. DOI: 10.1007/s10988-020-09318-3.

Ladusaw, William A. 1992. Expressing negation. In Chris Barker & David Dowty (eds.), *Semantics and linguistic theory*, vol. 2, 237–260. Ohio State University: Linguistic Society of America. DOI: 10.3765/salt.v2i0.3030.

Lahiri, Utpal. 1998. Focus and negative polarity in Hindi. *Natural Language Semantics* 6(1). 57–123. DOI: 10.1023/A:1008211808250.

Mayer, Erika, Susanne Wurmbrand & Jon Sprouse. 2018. An experimental investigation of NPI licensing under DE flip-flop. In Sherry Hucklebridge & Max Nelson (eds.), *NELS 48: Proceedings of the Forty-Eighth Annual Meeting of the North East Linguistic Society*. Amherst, MA: GLSA: University of Massachusetts.

Neeleman, Ad & Elena Titov. 2009. Focus, contrast, and stress in Russian. *Linguistic Inquiry* 40(3). 514–524. DOI: 10.1162/ling.2009.40.3.514.

Ovalle, Luis Alonso & Elena Guerzoni. 2004. Double negatives, negative concord and metalinguistic negation. In Anne Pycha Mary Androni Erin Debenport & Keiko Yoshimura (eds.), *Proceedings from the main session of the thirty-eighth meeting of the Chicago Linguistic Society: Volume 38-1*, 15–31. Chicago, IL: Chicago Linguistic Society Publications.

Panizza, Daniele & Yasutada Sudo. 2020. Minimal sufficiency with covert *even*. *Glossa: a journal of general linguistics* 5(1). 98. DOI: 10.5334/gjgl.1118.

Penka, Doris. 2007. *Negative indefinites*. Tübingen: Universität Tübingen. (Doctoral dissertation). DOI: 10.1093/acprof:oso/9780199567263.001.0001.

Penka, Doris. 2016. *Degree equatives – the same as comparatives*. Presentation at the Workshop Equative Constructions, Universität zu Köln. https://ling.sprachwiss.uni-konstanz.de/pages/home/penka/Pubs/equatives_koeln.pdf.

Romoli, Jacopo. 2013. A scalar implicature-based approach to neg-raising. *Linguistics and Philosophy* 36(4). 291–353. DOI: 10.1007/s10988-013-9136-2.

Rullmann, Hotze. 1995. *Maximality in the semantics of Wh-constructions*. Amherst, MA: UMass. (Doctoral dissertation).

Schwarz, Florian, Kajsa Djärv & Jérémy Zehr. 2020. Do Italian factives entail their presupposition? yes, but… In Rajesh Bhatt, Ilaria Frana & Paula Menéndez-Benito (eds.), *Making worlds accessible: Essays in honor of Angelika Kratzer*, 150–166. Amherst, MA: ScholarWorks at UMass Amherst. https://scholarworks.umass.edu/ak_festsite_schrift/1/.

Seuren, Pieter A.M. 1984. The comparative revisited. *Journal of Semantics* 3(1). 109–141. DOI: 10.1093/jos/3.1-2.109.

Starschenko, Alexej & Marta Wierzba. 2023. *L-Rex Linguistic rating experiments [software], version 1.0.2.* GNU General Public License v3.0. https://github.com/2e2a/l-rex/.

Stechow, Arnim von. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3(1-2). 1–77. DOI: 10.1093/jos/3.1-2.1.

Tottie, Gunnel. 1991. *Negation in English speech and writing: A study in variation*, vol. 4 (Quantitative analyses of linguistic structure). University of California: Academic Press.

Wickham, Hadley. 2016. *Ggplot2: elegant graphics for data analysis.* New York: Springer-Verlag. https://ggplot2.tidyverse.org.

Zeijlstra, Hedde. 2004. *Sentential negation and negative concord.* Amsterdam: University of Amsterdam. (Doctoral dissertation).

Zeijlstra, Hedde. 2022. Negation and negative dependencies. In Hedde Zeijlstra (ed.), vol. 5, chap. Negative concord is syntactic agreement, 45–76. Oxford: Oxford University Press. DOI: 10.1093/oso/9780198833239.003.0003.

Zwarts, Frans. 1998. Three types of polarity. In Fritz Hamm & Erhard Hinrichs (eds.), *Plurality and quantification*, 177–238. Dordrecht: Springer. DOI: 10.1007/978-94-017-2706-8_5.

Chapter 8

The meaning of Czech response particles

Kateřina Hrdinková & Radek Šimík

Charles University

This article deals with the semantics and interpretation of Czech response particles *ano* 'yes' and *ne* 'no'. Based on two experiments involving responses to negative polar questions, we argue that *ano* 'yes' encodes the relative feature [AGREE] and *ne* 'no' encodes the absolute feature [−], adopting the parlance of Roelofsen & Farkas's (2015) feature model. This contrasts with the proposal of Gruet-Skrabalova (2016), who argues, following previous work on English, that both of the Czech response particles are ambiguous between a relative and an absolute reading. We also find some tentative evidence for context affecting the interpretation of response particles, in line with the predictions of Krifka (2013).

1 Introduction

Response particles like *yes* and *no* are a common way to respond to polar questions. They exhibit anaphoric behavior in that their interpretation crucially depends on previous context and, more specifically, on the form and interpretation of the polar question they respond to. While responses to affirmative questions are largely unproblematic, responses to negative questions give rise to ambiguities (Kramer & Rawlins 2011, Espinal & Tubau 2019, Roelofsen & Farkas 2015, Krifka 2013; etc.); see (1) and (2), respectively. (The translations in (2B) are tentative and will be rectified in view of the experimental results.)

(1) A: Zalil   Petr květiny?
         watered Petr flowers
         'Has Petr watered flowers?'

Kateřina Hrdinková & Radek Šimík

> B: Ano. (= Zalil.) / Ne. (= Nezalil.)
> yes watered no NEG.watered
> 'Yes. (= He has.) / No. (= He hasn't.)'

(2) A: Nezalil Petr květiny?
NEG.watered Petr flowers
'Hasn't Peter watered flowers?'

> B: Ano. (= Zalil / Nezalil.) / Ne. (= Zalil / Nezalil.)
> yes watered NEG.watered no watered NEG.watered
> 'Yes. (= He has. / He hasn't.) / No. (= He has. / He hasn't.)'
>
> (translations tentative; to be rectified)

If the polar question is negative, as in (2), both *ano* 'yes' and *ne* 'no' can in principle correspond to a positive or a negative answer. They can, however, differ in naturalness and likelihood. To give an example from German, Claus et al. (2017) found out that it is more natural to confirm negative questions by *ja* 'yes' than by *nein* 'no'.

Using a version of the truth-value judgment task, we investigate the meaning of the two Czech response particles – *ano* 'yes' and *ne* 'no', hoping to contribute to related recent literature on Slavic languages (e.g. Gruet-Skrabalova 2016, Esipova 2021, Geist & Repp 2023). A more specific goal is to evaluate the adequacy of two types of existing accounts of response particle meaning: the feature model of Roelofsen & Farkas (2015), in which response particles have a lexically specified range of meanings, and the saliency account of Krifka (2013), in which the meaning is expected to be more context-dependent. We also discuss our results in view of Gruet-Skrabalova's (2016) analysis of Czech response particles, which is couched in the feature model. We conclude that our data primarily support a particular version of the feature model, though not the one proposed by Gruet-Skrabalova (2016). More specifically, we see a very clear tendency for *ano* 'yes' to express agreement (the feature [AGREE]) with its antecedent, be it positive or negative, and *ne* 'no' to express a negative proposition (the feature [−]), independently of the polarity of the antecedent. What counts as the "antecedent" is crucially modulated by the interrogative strategy used: negative polar questions with an interrogative syntax (verb-first) primarily contribute a positive antecedent (i.e., the negation is, by hypothesis, "pleonastic"), while negative polar questions with a declarative syntax (non-verb-first) contribute a negative antecedent (negation is semantic/propositional). Even though the feature model appears to be most suitable for modelling our results, we also observe – in a subset of our data – a statistically significant result predicted by Krifka's (2013) saliency theory.

The article is structured as follows. §2 briefly introduces the two approaches under consideration – the feature model (Roelofsen & Farkas 2015) and the saliency theory (Krifka 2013). We also discuss Gruet-Skrabalova's (2016) particular application of the feature model on Czech. §3 reports on the experiments we have conducted: experiment 1, in which we investigated responses to negative polar questions with interrogative syntax (V1), and experiment 2, in which we looked at negative polar questions with declarative syntax. In §4 we discuss the results and propose a new implementation of the feature model which is consistent with the results. Finally, §5 concludes the paper.

2 Approaches to response particle meaning

2.1 Feature model

2.1.1 Roelofsen & Farkas (2015)

The influential FEATURE MODEL of Roelofsen & Farkas (2015) is based on the assumption that a response particle like 'yes' or 'no' has a lexically specified range of meanings, defined in terms of two types of features – absolute and relative polarity features. The ABSOLUTE FEATURES [+] and [−] correspond to the polarity of the response. The relative features [AGREE] and [REVERSE] indicate a relation between the response and its propositional antecedent (derived from a polar question or an assertion that antecedes the response): the former expresses agreement with the polarity of the antecedent, the latter reverses its polarity. The semantics of the features is presuppositional (see Roelofsen & Farkas 2015: 385f. for details). In the lexicon, a response particle can either be specified for a single feature or for a combination of features. Additional complexity may arise in the process of feature realization (spellout), where Roelofsen & Farkas (2015) assume that a feature bundle can be realized by a particle which matches only its proper subset.[1]

Table 1 shows the assumed lexical entries and corresponding realization patterns of the English particles *yes* and *no* and the German particle *doch*. By hypothesis, the English particles encode single features, but are lexically ambiguous – they either encode the respective absolute or relative features. If a feature bundle is generated in the syntax (and interpreted in the semantics), it is realized by a particle whose lexical makeup matches a proper subset of that bundle. In two cases – [AGREE, −] and [REVERSE, +] – both *yes* and *no* provide a good match,

---

[1]For a recent experimental evaluation of Roelofsen & Farkas's (2015) model, see Maldonado & Culbertson (2023).

giving rise to an ambiguity which must be resolved pragmatically.[2] An example of a bundle-encoding particle is German *doch*, which responds to negative antecedents and at the same time reverses their polarity, whence [REVERSE, +].

Table 1: Feature bundles in the feature model

	Lexically encoded by	Realized by
[+]	*yes*	*yes*
[−]	*no*	*no*
[AGREE]	*yes*	*yes*
[REVERSE]	*no*	*no*
[AGREE, +]	n.a.	*yes*
[AGREE, −]	n.a.	*yes* or *no*
[REVERSE, +]	*doch*	*yes* or *no* / *doch*
[REVERSE, −]	n.a.	*no*

Thus, the English *yes* can signal that the answer has a positive polarity [+] or it can agree with its propositional antecedent [AGREE]. In contrast, *no* can serve either to signal negative polarity [−], or to reverse the polarity of its antecedent [REVERSE]. These double properties of *yes* and *no*, according to Roelofsen & Farkas (2015: 383), explain why response particles are generally clear after a positive question / positive statement (3), while a double interpretation is possible after a negative question / negative statement (4).

(3) Amy left.   (positive antecedent)
    Agreement: Yes, she did. / *No, she did.
    Reversion: *Yes, she didn't. / No, she didn't.

(4) Amy didn't leave.   (negative antecedent)
    Agreement: Yes, she didn't. / No, she didn't.
    Reversion: Yes, she did. / No, she did.

---

[2]Roelofsen & Farkas (2015) employ a set of additional markedness-based rules which nudge the likelihood in one or the other way.

2.1.2 The feature model applied to Czech: Gruet-Skrabalova (2016)

Gruet-Skrabalova (2016) adopts Roelofsen & Farkas's (2015) feature model and adapts it to Czech. Gruet-Skrabalova assumes that Czech response particles *ano* 'yes' and *ne* 'no' exhibit the same ambiguity as the English particles *yes* and *no*, i.e., they can either realize the absolute features ([+] and [−], respectively) or the relative features ([AGREE] and [REVERSE], respectively).

Gruet-Skrabalova (2016) further modulates her analysis relative to the form of the question which antecedes the response. She assumes that in interrogative questions, i.e., questions with the verb in clause-initial position (V1), the polarity is neutralized. Response particles used in reaction to V1 questions therefore realize their absolute features. This is illustrated for the case of negative V1 questions in (5), where there is no ambiguity in the response: *ano* 'yes' indicates positive polarity and *ne* 'no' negative polarity. In declarative questions, i.e., questions with the verb in a different than initial position (non-V1), the polarity is salient and the response particles realize their relative features. As a result, response particles are also not ambiguous in this case, but have opposite truth-conditions; see (6).

(5) Negative interrogative question (with V1)

   A: Nenapsala Jitka esej?
           NEG.wrote Jitka essay
           'Hasn't Jitka written an essay?'

   B: Ano. (= Napsala.) / Ne. (= Nenapsala.)
           yes   wrote   no   NEG.wrote
           'Yes. (= She has.) / No. (= She hasn't.)'

(6) Negative declarative question (with non-V1)

   A: Jitka esej   nenapsala?
           Jitka essay NEG.wrote
           'Hasn't Jitka written an essay?'

   B: Ano. (= Nenapsala.) / Ne. (= Napsala.)
           yes   NEG.wrote   no   wrote
           'Yes. (= She hasn't.) / No. (= She has.)'

2.2 Saliency account: Krifka (2013)

Krifka's (2013) SALIENCY ACCOUNT takes an additional factor into account, namely the role of contextual and more generally pragmatic considerations, co-determining which proposition is selected as the antecedent for the response particle.

Kateřina Hrdinková & Radek Šimík

In Krifka's (2013) theory, response particles are propositional anaphors, not unlike pronouns.[3] If a response particle is preceded by a question which contains negation, there are in principle two possible antecedents for the response particle: either the negative proposition or the negation's prejacent, i.e., the corresponding positive proposition. This is illustrated in (7) (adapted from Krifka 2013: 14).[4]

(7) [$_{ActP}$ did REQUEST [$_{NegP}$ Ede not [$_{TP}$ t$_{Ede}$ t$_{did}$ steal the cookie]]]?
        ↪ d ↪ d'

Response particles used in reaction to a question like (7) can thus be interpreted as in (8), capturing the ambiguity discussed above.

(8) a. *Yes.* ⇝ ASSERT(d') ≈ *Yes, he did!*   (rejecting accent, with clause)
     b. *Yes.* ⇝ ASSERT(d) ≈ *Yes, he didn't.*   (natural, but with clause)
     c. *No.* ⇝ ASSERT($\neg d'$) ≈ *No (he didn't).*   (natural, clause not necessary)
     d. *No.* ⇝ ASSERT($\neg d$) ≈ *Well, he did!*   (rejecting accent, with clause)

What is of interest to us is how the antecedent of the response particle is selected, i.e., whether the response particle denotes d (the negative proposition) or d' (the positive proposition). Krifka (2013) assumes that the saliency of the propositions – and hence the likelihood of their antecedent status – can be modulated contextually. Example (9) (adapted from Krifka 2013: 14) and the matching example (10) (created by us) illustrate this point. These examples differ in the question under discussion put on the table by A: in (9), the issue is negatively defined, in (10), the issue is positively defined. Although the default antecedent for the response particles in both cases will be the negative proposition asserted by B, the context is assumed to modulate the availability of the positive antecedent, which leads to an increased likelihood of the truth-conditionally opposite responses in (10).[5]

---

[3]In this respect, Roelofsen & Farkas's (2015) and Krifka's (2013) theories are similar. Both crucially build on an analogy with pronouns – the former via pronominal-like presuppositions (not discussed here), the latter via the anaphoric potential of pronouns.

[4]For questions with high negation (*Didn't Ede steal the cookie?*), Krifka (2013: 14) assumes only one possible antecedent, namely the positive d'. The negation in this case is applied outside of the scope of the proposition, making it unavailable for anaphoric pickup.

[5]An anonymous reviewer points out that the positive interpretation ('He climbed it') of response (10A$_1$) might be contingent on *yes* being pronounced with a specific intonation. This is indeed what Goodhue & Wagner (2018) confirmed experimentally; they call the fall rise intonation used in these cases "contradiction contour", following Liberman & Sag (1974).

(9) Negative context (italicized)

A: Which of the mountains on this list *did Reinhold Messner not climb*?

B: Well, let's see… He did not climb Mountain Cotopaxi in Ecuador.

A$_1$: Yes.
Likely: 'He didn't climb it.'
Unlikely: 'He climbed it.'

A$_2$: No.
Likely: 'He climbed it.'
Unlikely: 'He didn't climb it.'

(10) Positive context (italicized)

A: Which of the mountains on this list *did Reinhold Messner climb*?

B: Well, let's see…He did not climb Mountain Cotopaxi in Ecuador.

A$_1$: Yes.
Likely: 'He didn't climb it.'
More likely than in (9): 'He climbed it.'

A$_2$: No.
Likely: 'He climbed it.'
More likely than in (9): 'He didn't climb it.'

One of our experiments (experiment 1) will tap not only into the basic meaning of Czech response particles, which can be formulated in terms of the feature model, but also into the influence of the context in determining the antecedent of the response particles.

3 Experiments

The aim of our experiments was to find out the preferred meaning of *ano* 'yes' and *ne* 'no' in response to polar questions and test the above-mentioned approaches, in particular Gruet-Skrabalova's (2016) version of the feature-based analysis and Krifka's (2013) idea that the choice of the response particle antecedent is context-dependent. The experimental design was inspired by Kramer & Rawlins (2012) and Claus et al. (2017), who investigate the meaning of response particles in English and German.

Our results suggest a relative ([AGREE]-based) semantics for *ano* 'yes' and absolute ([−]-based) semantics for *ne* 'no'. This can be easily modeled using the feature model. The particular predictions of Gruet-Skrabalova (2016) were, however,

not borne out: we do not see evidence for relative ([REVERSE]-based) semantics for *ne* 'no'. In addition, in a particular corner of our data, we see a pattern which is predicted by Krifka's (2013) saliency account.

We used 2 experiments combined in a single setup, so that each of the two experiments provided fillers for the other one. This setup makes it possible to draw inferences cross-experimentally. The more complex and powered experiment 1 uses a $2 \times 2 \times 2$ design and investigates responses to syntactically interrogative negative polar questions. Experiment 2 uses a 2×2 design and investigates responses to syntactically declarative negative polar questions.

We first describe aspects common to the two experiments (see §3.1) and then turn to the individual experiments (§3.2–§3.3).

3.1 Aspects common to both experiments

As detailed in Table 2, our experimental setup consisted of two experiments with 16 and 8 items, respectively, and an additional set of 16 filler items, giving a total of 40 items. The number matches the number of stimuli seen by each participant.

Table 2: Overall experimental setup

Experiment 1	16
Experiment 2	8
Fillers	16
Total	40

3.1.1 Task, procedure, and dependent variables

The participants were exposed to written stimuli which consisted of a short narrative (a few sentences) followed by a short dialogue between two people (A and B for ease of reference), in which A opens the dialogue by an assertion associated with the narrative, B asks a relevant polar question, and A responds by saying either 'yes' or 'no'. The narrative and the dialogues contained mildly colloquial elements, in order to simulate an informal setting – a dialogue between two friends. The participant's task was twofold: (i) to determine whether A's 'yes'/'no' response is consistent with the information provided in the narrative (i.e., a truth-value judgment task) and (ii) to rate the naturalness of that response

given the preceding narrative and dialogue (on a scale from 1 = completely unnatural to 7 = completely natural). In this paper, we analyze the responses from task (i) and leave (ii) aside. This is mainly because there was a strong correlation between the two in the sense that responses which were judged as consistent with the information provided were also rated as natural and conversely – responses judged as inconsistent were rated as unnatural.

3.1.2 Participants

Data from 66 adult native speakers of Czech (43 women, 23 men) entered the analyses. We used convenient sampling, recruiting participants from an extended social network of the first author. Most of the participants (44) were 18–29 years old and most (43) had university education. All participants were informed about the purpose of the experiment and all gave informed consent to participate in the experiment and the subsequent processing and anonymous publication of the collected data.

The analyzed sample consists of participants who have passed a preset quality measurement, namely scoring in the expected way on variable (i) in at least 75 % of the 8 filler items, where the relation between the information provided (in the narrative) and the 'yes'/'no' response was particularly transparent.

3.1.3 Software and administration

The experiments were prepared and administered using the L-Rex software (Starschenko & Wierzba 2023). The stimuli from each experiment were distributed on lists using the Latin Square design, so that one participant saw only a single stimulus from each item. The lists from each experiment were then combined and the order of presentation was pseudo-randomized in such a way that two stimuli from a single experiment never directly followed one another and two stimuli from a single condition were always interspersed by at least one stimulus from a different condition.

The experiment was distributed online by sending a link. Participants took part at their own personal computers and most of them needed 25–40 minutes to complete the experiment.

3.2 Experiment 1: Syntactically interrogative negative polar questions

3.2.1 Design and manipulated variables

This experiment focuses on the most complex and problematic case: responses to negative polar questions which are syntactically interrogative, which means that

Kateřina Hrdinková & Radek Šimík

the finite verb is located in the clause-initial position (V1 for short); see (11). Since negation is obligatorily prefixed to the verb in Czech, there is no reliable formal difference between high and low negation (cf. Ladd 1981) and its semantic correlate outer (extra-propositional) vs. inner (propositional) negation (AnderBois 2019, Goodhue 2022). Yet there is a general consensus that negation in V1 polar questions in Czech corresponds to high negation, which is either pleonastic (expletive) or applies at an illocutionary level (Repp 2013; for a discussion on Czech, see Staňková 2023, Staňková & Šimík to appear, and the references cited therein).

(11) Neprodala Jitka ty   staré boty?
      NEG.sold   Jitka DEM old   shoes
      'Didn't Jitka / Did Jitka not sell the old shoes?'

In a factorial $2 \times 2 \times 2$ design, we manipulated three variables (all within items and subjects): INFORMATION, CONTEXT, and RESPONSE. The INFORMATION variable, with values positive (i_pos) and negative (i_neg) is manipulated in the lead-in narrative and fixes the factual state of affairs relative to which the participant judges the consistency of the response. The CONTEXT variable, likewise with values positive (c_pos) and negative (c_neg), was manipulated in the first utterance of the dialogue, which is then followed by the polar question. Finally, the RESPONSE variable has the values *ano* (yes) and *ne* (no) and is manipulated in the final utterance of the dialogue. For purposes of visualization and statistical analysis, we have found it useful to include an auxiliary variable, namely the ACCORDANCE between INFORMATION and RESPONSE, yielding the value accord for the cases where the positive information is matched by a yes response and negative information by a no response, and the value discord where this is not so. Table 3 provides an overview of all the 8 unique conditions of experiment 1.

3.2.2 Materials

Example (12) provides an example of an item (particularly item 14) in all eight conditions. The values of the INFORMATION variable is set in small caps, the CONTEXT in italics, and the RESPONSE in boldface. The parts that remained constant across the manipulations – including the negative polar question – are set in ordinary typeface. The value of the INFORMATION variable was located in the lead-in narrative, specifically in the position indicated by […].[6]

---

[6]All the experiment materials, results, and outputs of statistical models are available at Open Science Framework under the following link: https://doi.org/10.17605/OSF.IO/9VXJS.

Table 3: Conditions in the factorial design of Experiment 1

	INFORMATION	CONTEXT	RESPONSE	ACCORDANCE
a	i_pos	c_pos	yes	accord
b	i_pos	c_pos	no	discord
c	i_pos	c_neg	yes	accord
d	i_pos	c_neg	no	discord
e	i_neg	c_pos	yes	discord
f	i_neg	c_pos	no	accord
g	i_neg	c_neg	yes	discord
h	i_neg	c_neg	no	accord

(12) Eva a   Lída se   zúčastnily vánočního plesu ve svém rodném městě.
Eva and Lída REFL took.part Christmas ball in their birth   town
Lída, která se   velmi zajímá o společenský život ve svém rodišti, […]
Lída who REFL a.lot interest in social   life in her   birthplace
Když se   po nějaké době potkají, probírají spolu   proběhlý ples.
when REFL after some time meet   discuss together passed   ball
'Eva and Lída attended a Christmas ball in their hometown. Lída, who is
interested in the social life in her hometown very much, […] When Eva
and Lída meet after a while, they discuss the ball together.'

   a. ví,   že   HLAVNÍ ORGANIZÁTORKOU BYLA JEJICH BÝVALÁ
      knows that main   organizer   was their former
      SPOLUŽAČKA A   ZKUŠENÁ   ORGANIZÁTORKA PLESŮ ALICE.
      classmate   and experienced organizer   balls Alice
      'knows that the main organizer was their former classmate and an
      experienced ball organizer Alice.'

   Lída: *Ten ples se   jim moc povedl.*
         DEM ball REFL them much worked.out
         'The ball worked out really well.'

   Eva: Neorganizovala ho Alice?
        NEG.organized it Alice
        'Didn't Alice organize it?'

   Lída: **Ano.**
         yes
         'Yes.'

Kateřina Hrdinková & Radek Šimík

  b. ví,   že   ʜʟᴀᴠɴɪ́ ᴏʀɢᴀɴɪᴢᴀ́ᴛᴏʀᴋᴏᴜ ʙʏʟᴀ ᴊᴇᴊɪᴄʜ ʙʏ́ᴠᴀʟᴀ́
     knows that main   organizer   was their former
     ꜱᴘᴏʟᴜᴢ̌ᴀᴄ̌ᴋᴀ ᴀ   ᴢᴋᴜꜱ̌ᴇɴᴀ́   ᴏʀɢᴀɴɪᴢᴀ́ᴛᴏʀᴋᴀ ᴘʟᴇꜱᴜ̊ Aʟɪᴄᴇ.
     classmate   and experienced organizer   balls Alice
     'knows that the main organizer was their former classmate and an
     experienced ball organizer Alice.'

     Lída: *Ten ples se   jim   moc   povedl.*
         ᴅᴇᴍ ball ʀᴇꜰʟ them much worked.out
         'The ball worked out really well.'

     Eva: Neorganizovala ho Alice?
         ɴᴇɢ.organized it Alice
         'Didn't Alice organize it?'

     Lída: **Ne.**
         no
         'No.'

  c. ví,   že   ʜʟᴀᴠɴɪ́ ᴏʀɢᴀɴɪᴢᴀ́ᴛᴏʀᴋᴏᴜ ʙʏʟᴀ ᴊᴇᴊɪᴄʜ ʙʏ́ᴠᴀʟᴀ́
     knows that main   organizer   was their former
     ꜱᴘᴏʟᴜᴢ̌ᴀᴄ̌ᴋᴀ ᴀ   ᴢᴋᴜꜱ̌ᴇɴᴀ́   ᴏʀɢᴀɴɪᴢᴀ́ᴛᴏʀᴋᴀ ᴘʟᴇꜱᴜ̊ Aʟɪᴄᴇ.
     classmate   and experienced organizer   balls Alice
     'knows that the main organizer was their former classmate and an
     experienced ball organizer Alice.'

     Lída: *Ten ples se   jim   moc   nepovedl.*
         ᴅᴇᴍ ball ʀᴇꜰʟ them much ɴᴇɢ.worked.out
         'The ball worked out really well.'

     Eva: Neorganizovala ho Alice?
         ɴᴇɢ.organized it Alice
         'Didn't Alice organize it?'

     Lída: **Ano.**
         yes
         'Yes.'

  d. ví,   že   ʜʟᴀᴠɴɪ́ ᴏʀɢᴀɴɪᴢᴀ́ᴛᴏʀᴋᴏᴜ ʙʏʟᴀ ᴊᴇᴊɪᴄʜ ʙʏ́ᴠᴀʟᴀ́
     knows that main   organizer   was their former
     ꜱᴘᴏʟᴜᴢ̌ᴀᴄ̌ᴋᴀ ᴀ   ᴢᴋᴜꜱ̌ᴇɴᴀ́   ᴏʀɢᴀɴɪᴢᴀ́ᴛᴏʀᴋᴀ ᴘʟᴇꜱᴜ̊ Aʟɪᴄᴇ.
     classmate   and experienced organizer   balls Alice
     'knows that the main organizer was their former classmate and an
     experienced ball organizer Alice.'

Lída: *Ten ples se   jim   moc   nepovedl.*
   DEM ball REFL them much NEG.worked.out
   'The ball worked out really well.'

Eva: Neorganizovala ho Alice?
   NEG.organized   it Alice
   'Didn't Alice organize it?'

Lída: **Ne.**
   no
   'No.'

e. VÍ,   ŽE   JEJICH BÝVALÁ SPOLUŽAČKA A   ZKUŠENÁ
   knows that their   former classmate   and experienced
   ORGANIZÁTORKA PLESŮ ALICE SE   TENTOKRÁT NA ORGANIZACI
   organizer   balls Alice REFL this.time   in organization
   NEPODÍLELA.
   NEG.was.involved
   'knows that their former classmate and an experienced ball organizer
   Alice wasn't involved in the organization this time.'

Lída: *Ten ples se   jim   moc   povedl.*
   DEM ball REFL them much worked.out
   'The ball worked out really well.'

Eva: Neorganizovala ho Alice?
   NEG.organized   it Alice
   'Didn't Alice organize it?'

Lída: **Ano.**
   yes
   'Yes.'

f. VÍ,   ŽE   JEJICH BÝVALÁ SPOLUŽAČKA A   ZKUŠENÁ
   knows that their   former classmate   and experienced
   ORGANIZÁTORKA PLESŮ ALICE SE   TENTOKRÁT NA ORGANIZACI
   organizer   balls Alice REFL this.time   in organization
   NEPODÍLELA.
   NEG.was.involved
   'knows that their former classmate and an experienced ball organizer
   Alice wasn't involved in the organization this time.'

Lída: *Ten ples se   jim   moc   povedl.*
   DEM ball REFL them much worked.out
   'The ball worked out really well.'

Eva: Neorganizovala ho Alice?
NEG.organized it Alice
'Didn't Alice organize it?'

Lída: **Ne.**
no
'No.'

g. VÍ, ŽE JEJICH BÝVALÁ SPOLUŽAČKA A ZKUŠENÁ
knows that their former classmate and experienced
ORGANIZÁTORKA PLESŮ ALICE SE TENTOKRÁT NA ORGANIZACI
organizer balls Alice REFL this.time in organization
NEPODÍLELA.
NEG.was.involved
'knows that their former classmate and an experienced ball organizer
Alice wasn't involved in the organization this time.'

Lída: *Ten ples se jim moc nepovedl.*
DEM ball REFL them much NEG.worked.out
'The ball worked out really well.'

Eva: Neorganizovala ho Alice?
NEG.organized it Alice
'Didn't Alice organize it?'

Lída: **Ano.**
yes
'Yes.'

h. VÍ, ŽE JEJICH BÝVALÁ SPOLUŽAČKA A ZKUŠENÁ
knows that their former classmate and experienced
ORGANIZÁTORKA PLESŮ ALICE SE TENTOKRÁT NA ORGANIZACI
organizer balls Alice REFL this.time in organization
NEPODÍLELA.
NEG.was.involved
'knows that their former classmate and an experienced ball organizer
Alice wasn't involved in the organization this time.'

Lída: *Ten ples se jim moc nepovedl.*
DEM ball REFL them much NEG.worked.out
'The ball worked out really well.'

Eva: Neorganizovala ho Alice?
NEG.organized it Alice
'Didn't Alice organize it?'

> Lída: **Ne.**
>   no
>   'No.'

16 items like (12) were created, meaning that each participant was exposed to each unique condition twice (following the Latin Square distribution we used; see §3.1). This number – admittedly not great judging by current standards (Häussler & Juzek 2017) – resulted from a compromise between statistical power considerations and the significant cognitive load imposed by the task on the participants. The introductory narrative was always presented in stylistically neutral language and the dialogues occasionally contained colloquial expressions. The information at issue (above: whether Alice organized the ball) is known to the first dialogue participant (above: Lída), but not to the second one (above: Eva). The first participant makes a claim relevant to the information at issue (above: how the ball worked out), but does not reveal its value. The first participant's utterance stands in a particular relation to the information: it makes it more likely or less likely. The second participant asks a question about the information, followed by a response of the first participant.

Consider (12a) for illustration. In this condition it is the case that Lída (the first dialogue participant) knows that Alice organized the ball. Also, as the narrative implies, both Lída and Eva (the second dialogue participant) are aware that Alice is a good ball organizer. Lída's first utterance in the dialogue – that the ball worked out really well – implies that the ball was organized by Alice (a case of evidential bias). Eva then asks a polar question, in order to verify or falsify the implication. Lída responds *ano* 'yes'. Setting the naturalness rating aside, the participant had the option to either say that Lída's response is consistent with the information provided (i.e., the response is true), or to say that Lída's response is not consistent with the information provided (i.e., the response is false). In the former case, we assume that the participant either interprets the response absolutely (feature [+] / positive polarity = 'Alice organized the ball.') or relatively (feature [AGREE] / agreement with the antecedent 'Alice organized the ball' – made available by the prejacent of the negative polar question asked). In the latter case, the participant interpreted Lída's response as involving the relative feature [AGREE], agreeing with the negative antecedent – 'Alice didn't organize the ball' – made available by the prejacent of the polar questions, including the negation. Whether the positive or the negative form of the antecedent is more salient (and hence whether the response is considered as true or false, provided it is interpreted relatively) is, by hypothesis (Krifka 2013), co-determined by the context – Lída's first utterance (more technically: evidential bias).

3.2.3 Results

Figure 1 shows the raw results of experiment 1, in particular of the ratings of consistency between the RESPONSE (yes vs. no) and the INFORMATION provided (i_pos vs. i_neg). Figure 2 provides the corresponding 95% confidence intervals (computed with the emmeans function of the emmeans package of R; Lenth 2024). The results are visualized – and also statistically analyzed – using the auxiliary ACCORDANCE variable. The value accord combined with the value i_pos equals the value yes, combined with the value i_neg equals no, and conversely for discord. The values of the response variable (yes vs. no) are indicated in the top left corner of each of the four panes for clarity. The reason for using ACCORDANCE rather than RESPONSE is that from the perspective of the results, the levels of the former variable form more of a natural class than the levels of the latter variable. The results are thus easier to evaluate and interpret.[7]

Looking at the dependent variable, we note that consistency value 1 indicates that the participant considered the response to be consistent with the information provided, or, in other words, true relative to the information provided. Consistency value 0 indicates a judgment of falsity.

To give a particular example, the top right panel shows that a no response after negative information (corresponding to accord / conditions f/g of our design; see Table 3 and (12f/g)) was considered consistent with the information in about 78 % of the cases; on the other hand, the bottom left panel shows that a no response after positive information (corresponding to discord / conditions b/d) was considered consistent with the information in only about 15 % of the cases.

We fitted a generalized linear mixed model, using the glmer function of the lme4 package (Bates et al. 2015) of the R software (R Core Team 2021), to estimate the effect of INFORMATION, CONTEXT, ACCORDANCE, and their mutual interactions on the consistency rating. We included random intercepts and slopes for both items and participants; INFORMATION and CONTEXT were sum-coded, ACCORDANCE treatment-coded (using accord as the reference level).[8] The model confirms the naked-eye-visible effect of ACCORDANCE: responses which were in accord with the information (the top row in Figure 1) were rated as true much more often than responses which were in discord with the information

---

[7]See the discussion (§4) for a visualization using the RESPONSE variable.

[8]The particular formula used was: CONSISTENCY RATING ~ INFORMATION * CONTEXT * ACCORDANCE + (1 + INFORMATION + CONTEXT + ACCORDANCE | participant) + (1 + INFORMATION + CONTEXT + ACCORDANCE | item). Treatment coding was used for ACCORDANCE because it has a natural reference level (accord) at which we expected high consistency (as compared to the discord level). Such a clear relationship was absent in the other factors, for which reason we applied sum coding to them.

yes   no

no   yes
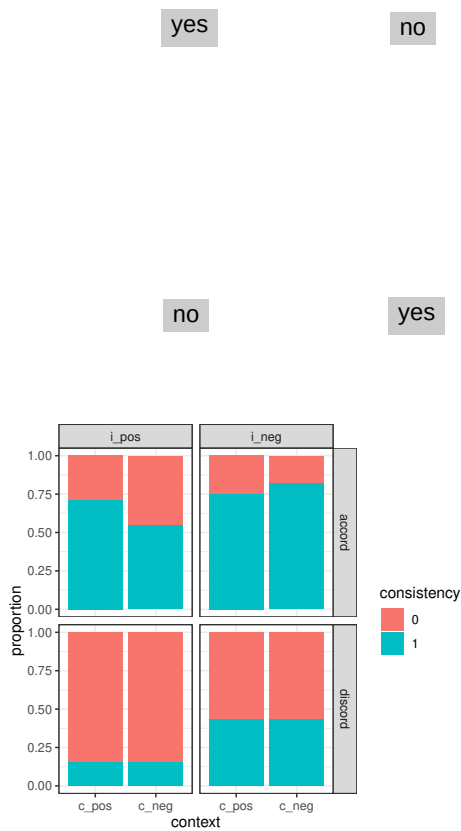
Figure 1: Experiment 1: Response–information consistency ratings
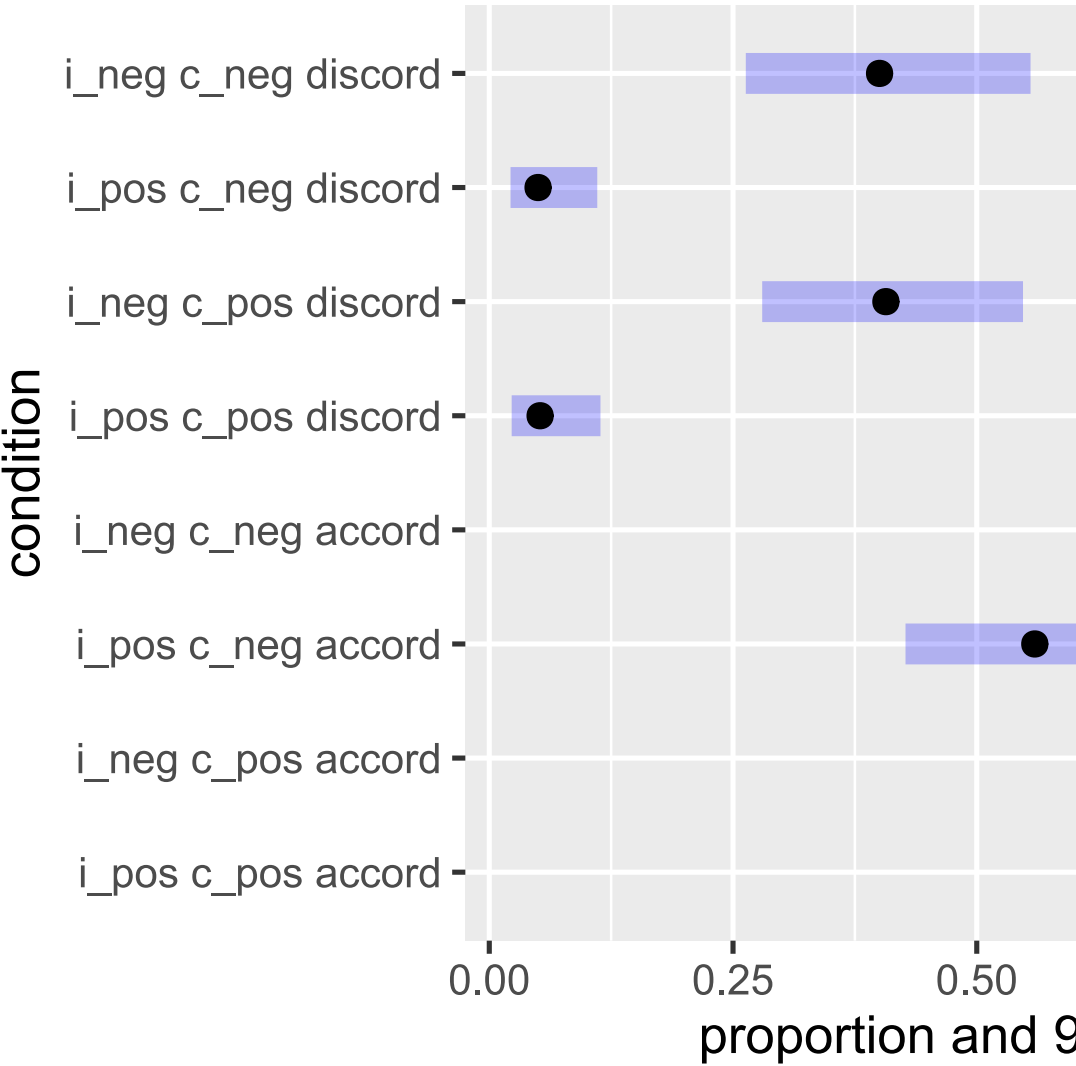
Figure 2: Experiment 1: 95% confidence intervals of consistency ratings

($z = -6.580$, $p < .001$). For instance, a positive response after positive information (see in (12a)) was rated as more consistent than after negative information (see (12e)). In addition, the model revealed a significant main effect of INFORMATION ($z = -3.660$, $p < .001$; not easily interpretable), an interaction between INFORMATION and ACCORDANCE ($z = -3.849$, $p < .001$), between INFORMATION and CONTEXT ($z = 3.627$, $p < .001$), and a three-way interaction between all factors ($z = -2.399$, $p = .016$). The interaction between INFORMATION and ACCORDANCE indicates that the effect of INFORMATION is more pronounced in the discord level of ACCORDANCE. The interaction between INFORMATION and CONTEXT is only visible if the response was in accord with the information (the top pane in Figure 1), which is also indicated by the significant three-way interaction. In order to see the effect CONTEXT in a clearer way, we fitted a model on the accord data subset, including random intercepts and slopes for both items and participants and sum-coding for both predictors – INFORMATION and CONTEXT.[9] This model confirmed the aforementioned interaction ($z = 3.584$, $p < .001$) and a further statistical analysis (nesting CONTEXT within the levels of INFORMATION) revealed that its source is both in i_pos and i_neg: if the information was positive, responses were rated as true more often if the context was also positive, as in (12a) as opposed to (12c) (simple effect of CONTEXT within i_pos; $z = 1.999$, $p = .046$), and if the information was negative, responses were rated as true more often if context was also negative, as in (12h) as opposed to (12f) (simple effect of CONTEXT within i_neg; $z = -3.389$, $p < .001$).[10]

3.2.4 Discussion

Experiment 1 clearly reveals that particle responses to negative syntactically interrogative questions are judged to be true if their polarity is in accordance with the polarity of the information provided: *ano* 'yes' is judged as true if the information is positive and *ne* 'no' is judged as true if the information is negative. This result is consistent with the absolute feature analysis, under which *ano* 'yes' encodes [+] and *ne* 'no' encodes [−]. If the negation is pleonastic, as is commonly assumed for interrogative (V1) questions (Gruet-Skrabalova 2016), then the results also follow under the relative feature analysis, under which *ano* encodes [AGREE] and *ne* [REVERSE], because in both cases the antecedent is positive.

---

[9]The formula used was: CONSISTENCY RATING ~ INFORMATION * CONTEXT + (1 + INFORMATION * CONTEXT | participant) + (1 + INFORMATION * CONTEXT | item).

[10]The formula used for the last model was: CONSISTENCY RATING ~ INFORMATION / CONTEXT + (1 + INFORMATION * CONTEXT | participant) + (1 + INFORMATION * CONTEXT | item).

Kateřina Hrdinková & Radek Šimík

That said, we should also note that there is a difference in the behavior of *ano* 'yes' and *ne* 'no'. While the effect of ACCORDANCE is very clear for *ne* (numerical difference of 63 % between accord and discord), it is much less pronounced for *ano* (numerical difference of 20 %). The ratings for *ano* 'yes' are closer to chance in both accord and discord, indicating a greater degree of uncertainty in the consistency ratings. This pattern would be expected under the conjunction of the following two premises: the negative polar question makes the negative proposition available as an antecedent (i.e., negation is not pleonastic) and *ano* 'yes' encodes [AGREE], i.e., its semantics is relative and agrees either with the positive antecedent ('yes, she did') or with a negative antecedent ('yes, she didn't'). The fact that agreement with the positive antecedent is judged as true significantly more often than agreement with the negative antecedent would then reflect on the relative availability of the two antecedent types. This explanation would further be consistent with the fact that the availability of the positive antecedent is modulated by the context (simple effect of CONTEXT), in line with Krifka (2013): if the context is positive, the positive antecedent is available more (71 %) than if the context is negative (56 %). What is unexpected is that there is no analogous simple effect of CONTEXT if *ano* 'yes' is in discord with the information provided, i.e., if the information provided is negative. In this latter case, we could expect the negative context to make the negative antecedent more accessible and hence increase the consistency judgment relative to the positive context condition. This expectation is not met: *ano* 'yes' is judged as consistent in 43 % of the cases irrespective of the value of the CONTEXT variable.

Turning to the interpretation of *ne* 'no', the overall results are consistent with the absolute semantics ([−]). What is unexpected under this view, however, is the simple effect of context in the accord condition, i.e., that *ne* 'no' is judged true in more cases if the context is negative (79 %) than if it is positive (75 %). While this effect is numerically smaller than for 'yes' responses, it is statistically stronger. A [REVERSE]-based semantics would have a handle on this effect, but would leave the very low consistency ratings in the discord condition unexplained.

3.3 Experiment 2: Syntactically declarative polar questions

3.3.1 Design and manipulated variables

This experiment focuses on responses to negative syntactically declarative polar questions, i.e., questions in which the verb is located after the subject (V2); see (13).

(13) Jitka neprodala ty   staré boty?
      Jitka NEG.sold   DEM old   shoes
      'Jitka didn't sell the old shoes?'

We only used two crossed factors in this experiment – INFORMATION and RE-
SPONSE. Context was always negative because negative declarative questions
only sound natural in contexts indicating negative evidential bias (Gunlogson
2002, Staňková 2023) and our primary interest was the interpretation (not so
much naturalness) of response particles. The overview of the individual condi-
tions is provided in Table 4. The materials were parallel to those in experiment
1; we do not include a token set here in the interest of space.

Table 4: Conditions in the factorial design of Experiment 2

	INFORMATION	RESPONSE	ACCORDANCE
a	i_pos	yes	accord
b	i_pos	no	discord
c	i_neg	yes	discord
d	i_neg	no	accord

3.3.2 Results

Figure 3 shows the raw results of experiment 2, in particular the ratings of con-
sistency between the RESPONSE (yes vs. no) and the INFORMATION provided (i_pos
vs. i_neg). The values of the auxiliary ACCORDANCE variable are provided as labels,
for completeness. Figure 4 provides the corresponding 95% confidence intervals.
We fitted a generalized linear mixed model to estimate the effect of INFORMA-
TION, RESPONSE, and their interaction on the consistency rating. Both factors
were sum-coded. Random intercepts and slopes for both items and participants
were included.[11] The model confirms the visually clear main effect of INFORMA-
TION: responses are judged as more consistent with negative information (i_neg)
than with positive information (i_pos) ($z = -9.191$, $p < .001$). Furthermore, there
is an interaction between INFORMATION and RESPONSE: the effect of INFORMA-
TION is more pronounced with *ne* than with *ano* ($z = 4.091$, $p < .001$).

---

[11]The formula used was: CONSISTENCY RATING ~ INFORMATION * RESPONSE + (1 + INFORMATION
+ RESPONSE | participant) + (1 + INFORMATION + RESPONSE | item).

Kateřina Hrdinková & Radek Šimík

accord   discord   discord   accord
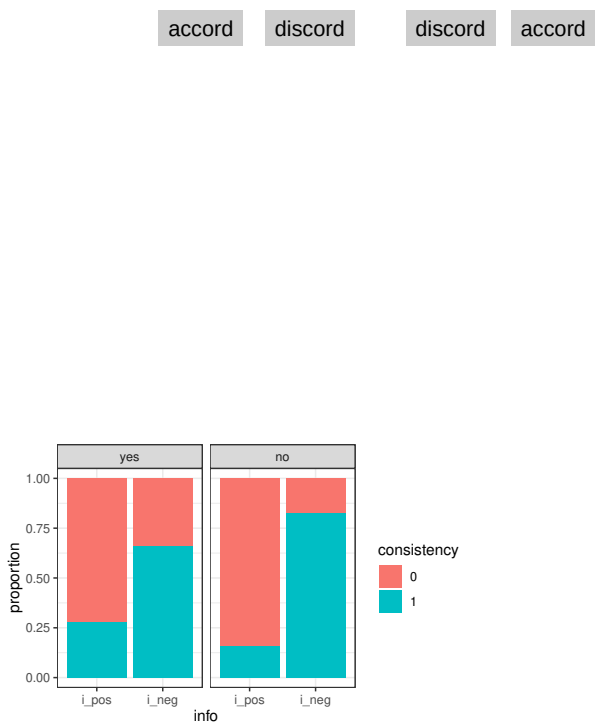
Figure 3: Experiment 2: Response–information consistency ratings
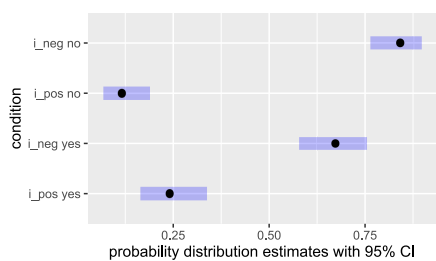
Figure 4: Experiment 2: 95% confidence intervals of consistency ratings

Kateřina Hrdinková & Radek Šimík

3.3.3 Discussion

The most significant result of experiment 2 is that both *ano* 'yes' and *ne* 'no' have the same truth conditions when responding to negative declarative polar questions: they are both judged as true if the information provided is negative (main effect of INFORMATION). This result follows on the premise that (i) *ano* 'yes' has relative semantics (encodes [AGREE]) and agrees with the negative antecedent ('yes, she didn't'), and, (ii) *ne* 'no' has absolute semantics (encodes [−]) and negative polarity ('(no,) she didn't').

The fact that the effect of INFORMATION is stronger for negative than for positive responses is consistent with this view. The relative semantics of *ano* 'yes' leaves some room for uncertainty as to which antecedent functions as the particle's prejacent. While the declarative form of the negative question makes the negative antecedent highly salient (making the 'yes, she didn't' interpretation true in 66 % of the cases in the i_neg/discord condition and 72 % in the i_pos/accord condition), the positive antecedent can also be accessed, at least when compared to the corresponding interpretations in the no condition (the 'yes, she did' interpretation is judged as true in 28 % of the cases in the i_pos/accord condition and 34 % in i_neg/discord condition). Compared to that, the hypothetical 'no, she did' interpretation is rather exceptional (only evident in about 17 % of the no responses overall).

4 General discussion

The results of our two experiments lend solid support to (i) relative, [AGREE]-based semantics of *ano* 'yes' and (ii) absolute, [−]-based semantics of *ne* 'no'. The proposed lexical encoding and the corresponding realization possibilities are represented in Table 5, an updated version of Table 1. The Czech particles *ano* 'yes' and *ne* 'no' are framed for clarity. The last column indicates where the evidence for the realization (im)possibilities stems from and whether the evidence is positive (judgment of truth) or negative (judgment of falsity).

In order to aid the discussion visually, we insert Figure 5, which includes data from both experiments: the top pane visualizes results of experiment 1 (aggregating over both levels of the CONTEXT variable), in which the question was interrogative (verb-first), and the bottom pane visualizes the results of experiment 2, in which the question was declarative (non-verb-first). For ease of reference, we label the individual stacked bars by capital letters.

Let us go through Table 5 step-by-step. If *ano* encoded [+] alone, we would expect the yes response in experiment 2 to be judged as consistent with the posi-
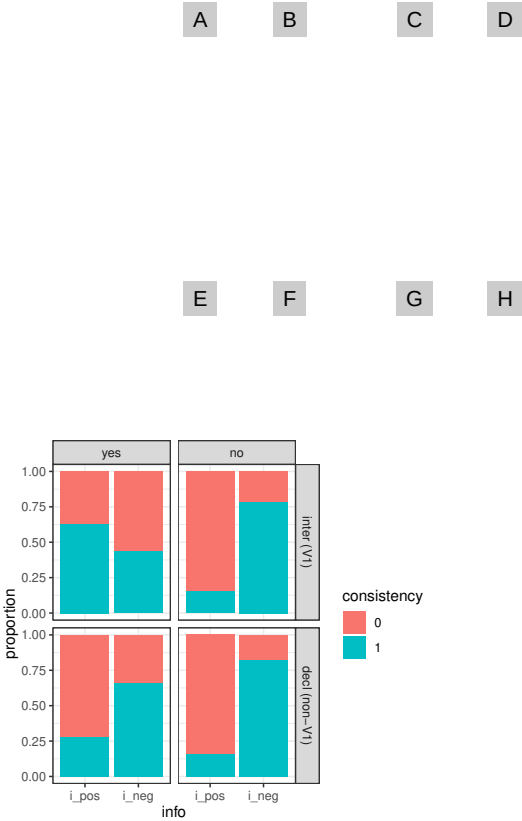
Figure 5: Both experiments: Response–information consistency ratings

Kateřina Hrdinková & Radek Šimík

Table 5: Feature bundles in the feature model (updated)

	Lexically encoded by	Realized by	(Positive/Negative) evidence from
[+]	*yes*	*yes*, ~~*ano*~~	exp 2 (neg)
[−]	*no*, *ne*	*no*, *ne*	exp 1, 2 (pos)
[AGREE]	*yes*, *ano*	*yes*, *ano*	exp 1, 2 (pos)
[REVERSE]	*no*	*no*, ~~*ne*~~	exp 1, 2 (neg)
[AGREE, +]	n.a.	*yes*, *ano*	exp 1 (pos)
[AGREE, −]	n.a.	*yes* or *no*, *ano* or *ne*	exp 1, 2 (pos)
[REVERSE, +]	*doch*	*yes* or *no*/*doch*, ~~*ano*~~ / ~~*ne*~~	exp 1, 2 (neg)
[REVERSE, −]	n.a.	*no*, *ne*	exp 1 (pos)

tive information (bar E in Figure 5). The fact that it cannot "ignore" the negative antecedent (i.e., agrees with it; bar F), strongly supports its relative (rather than absolute) semantics.

The assumption that *ne* encodes the absolute feature [−] is supported by the stability of its consistence with the negative information, independently of the question type preceding it (bars D and H).

That *ano* encodes [AGREE] is witnessed primarily by the differential behavior of this particle in experiment 1 and experiment 2. In the former, *ano* is judged more consistent with positive information (bar A, vs. B), and in the latter, *ano* is judged more consistent with negative information (bar F, vs. E). This follows if *ano* agrees with its antecedent and if interrogative (V1) questions make the positive antecedent more salient (cf. pleonastic negation), while declarative (non-V1) questions make the negative antecedent more salient. The fact that the effect of the INFORMATION variable is less pronounced in the yes condition, as compared to the no condition, is – or so we hypothesize – also consistent with the relative vs. absolute semantics of *ano* vs. *ne*, respectively. While the absolute semantics of *ne* remains largely intact to the polarity of its antecedent (being sensitive merely to the polarity-free prejacent), the relative semantics of *ano* leaves room for pragmatic and contextual considerations as to which antecedent – whether positive or negative – is selected as the prejacent of *ano*, which in turn leads to a greater variance in the consistency judgments and their overall centering around chance. This is especially evident in the results of experiment 1 (see A vs. B), where we also observed the effect of the CONTEXT variable predicted by Krifka (2013): posi-

tive context (as compared to negative context) supports the selection of a positive prejacent (see §3.2.3).

The fact that *ne* does not encode [REVERSE] is supported by the results of both experiments, but especially of experiment 2: while *ano* switched its truth conditions between experiment 1 and experiment 2, the truth conditions of *ne* remain stable. This clearly indicates that the differential availability of the two polar antecedents in these experiments had no effect on the meaning of *ne*, militating against its relative semantics. What supports the relative semantics, and is unexpected under our analysis, is that the consistence of the *ne* response with the negative information (bar D) is modulated by context: the consistence is higher if the context is negative – a mirror image of what happens in A. This effect is numerically small (only about 4 %), but statistically significant.

Let us now turn to the realization of the four logically possible feature combinations. The [AGREE, +] bundle is realized by *ano*, which, by the subset principle employed in the feature model, spells out [AGREE] (leaving [+] unrealized). This case is instantiated by bar A, where agreement is with a positive antecedent (supported by the tendentially pleonastic nature of the negation) and where the polarity of the response is, accordingly, positive ('yes, she did'). The [REVERSE, −] bundle is realized by *ne*, which spells out the subset [−]. This case is instantiated by bar D, where the polarity of the response is negative ('she didn't') and is reversed as compared to the primarily positive polarity of the antecedent. The [AGREE, −] bundle can in principle be realized in two ways – either by *ano,* which spells out [AGREE], or by *ne*, which spells out [−]. That precisely this is the case is witnessed by the identical truth conditions of the two particles in experiment 2 or, more specifically, by the analogous consistency ratings in bars F and H. In this case, the response agrees with the negative antecedent (≈ *ano*) and thus conveys a response of negative polarity (≈ *ne*). The most problematic case is represented by the bundle [REVERSE, +], which finds no suitable match in the lexical meanings of *ano* or *ne.* This scenario is represented by bars E and G and, as is evident from the consistency ratings, neither *ano*, nor *ne* are capable of reliably conveying it. It follows that a response with positive polarity reacting to a clearly negative antecedent (contributed in experiment 2 by the negative declarative question) cannot be expressed by a standalone particle in Czech. Instead, a more complex structure is warranted, such as a fragment (elliptical) response containing a verb explicitly specified for polarity (Gruet-Skrabalova 2015, 2016) or the positive particle *ano* 'yes' preceded by *ale* 'but'; see (14). We hypothesize that the particle *ale* 'but' reverses the salience of the two polar alternatives, making the positive one, which is otherwise only latently present, more salient and hence available as an antecedent of the relative particle *ano* 'yes'.

Kateřina Hrdinková & Radek Šimík

(14) A: Jitka neprodala ty   staré boty?
          Jitka NEG.sold   DEM old   shoes
          'Did Jitka not sell the old shoes?'

      B_1: #Ano. / #Ne.
               yes   no
              Intended: 'She sold the old shoes.'

      B_2: Prodala.
              sold
              'She did.'

      B_3: Ale ano.
              but yes
              'She did.'

This latter point brings us to a discussion of the predictions made by Gruet-Skrabalova (2016). Gruet-Skrabalova proposes that both of the Czech response particles can realize both the relative features and the absolute features. Our experimental results do not lend support to this claim. More particularly, we see only little evidence for [REVERSE] being realized by *ne* 'no' or for [+] being realized by *ano* 'yes'. The infelicity of the response ($14B_1$), which reflects our experimental results, is an example of this.

Our experimental results and the analysis we offer bear implications for the interpretation of Czech negative polar questions. Negative polar interrogatives (V1) are often considered to contain pleonastic negation, i.e., a negation which does not contribute propositional negation (see Staňková 2023 and the references therein). Such questions can thus be expected to only contribute positive propositions as antecedents available for anaphoric pick-up by the relative particle *ano* 'yes'. Counter to this expectation, we see that the negative proposition is not completely unavailable. In experiment 1, *ano* 'yes' is considered to be consistent with negative information in 43 % of the cases (bar B), a proportion which is hardly negligible (esp. when compared to no+i_pos condition; see bar C). We take this to indicate that negation attached to a fronted verb in polar interrogatives is not necessarily pleonastic; it can either be marginally read as propositional negation or contributes an illocutionary negation (called FALSUM by Repp 2013), which can (marginally) participate in forming an antecedent – possibly a speech act – which can in turn function as the prejacent of *ano* 'yes'. While a more detailed investigation of the interactions between the semantics of negative polar interrogatives and the semantics of polar responses is still missing, the experimental results reported in Staňková (2023) are consistent with the view just suggested.

Negative declarative questions, on the other hand, primarily contribute a negative antecedent, witnessed by the high consistence of *ano* 'yes' with negative information in experiment 2 (bar F). While this is, *prima facie*, an expected result, we also know from Staňková's (2023) results that negative declarative questions readily contribute not only inner negation (licensing negative concord items), but also outer ("pleonastic") negation (compatible with positive polarity items). If this is the case, we would expect the positive proposition to be more readily available for anaphoric pickup by *ano* 'yes'. Yet this is only possible in 28 % of the cases (bar E). Admittedly, however, the salience of the positive proposition is reduced by two factors in our experimental design (of experiment 2): the absence of any polarity item indicating outer negation and the contextual negative evidence (bias). It is an open issue whether the manipulation of these factors would have an impact on the availability of the positive interpretation of the particle *ano* (matched by an increased consistence in what would correspond to the E bar).

5 Conclusion

Our paper contributes the first experimental data pertaining to the semantics and interpretation of the two Czech polar response particles – *ano* 'yes' and *ne* 'no'. Building on the feature model of Roelofsen & Farkas (2015) and based on the results of our two experiments, we have argued that *ano* 'yes' lexically encodes the relative feature [AGREE] and *ne* 'no' encodes the absolute feature [−]. This stands in contrast to what has been proposed for Czech by Gruet-Skrabalova (2016) or for English by Roelofsen & Farkas (2015), namely that response particles are ambiguous between the relative and the absolute meaning. In addition, the results of experiment 1 reveal tentative evidence in favor of Krifka's (2013) proposal that context can affect the choice of the antecedent for relative response particles. More specifically, we saw that the relative particle *ano* 'yes' is resolved to a positive antecedent more often in case it is preceded by a positive context, as compared to a negative context. What is puzzling is that an inverse effect is observed for the particle *ne* 'no', which otherwise exhibits a pattern consistent with absolute lexical semantics (which in turn should be insensitive to contextual manipulations). The effect is numerically much smaller, but statistically stronger.

Finally, we have drawn some implications for the semantics of polar questions. Counter to the common assumption that negation on the fronted verb in interrogative questions is pleonastic (e.g., Gruet-Skrabalova 2016), i.e. not interpreted, we have seen some tentative evidence for the availability of a negative structure

Kateřina Hrdinková & Radek Šimík

being contributed by such questions. Whether it is a negative proposition or a negative speech act (as assumed e.g. by Staňková 2023) remains an open question. Likewise, it remains open how negative declarative questions (non-V1) in which negation is interpreted as outer negation (Staňková 2023) are responded to. The prediction is that the positive interpretation of *ano* 'yes' should be available to a greater extent in these cases.

Abbreviations

DEM demonstrative   REFL reflexive
NEG negation

Acknowledgments

References

AnderBois, Scott. 2019. Negation, alternatives, and negative polar questions in American English. In Klaus von Heusinger, Edgar Onea & Malte Zimmermann (eds.), *Questions in discourse – Volume 1: Semantics* (Current Research in the Semantics / Pragmatics Interface 35), 118–171. Leiden: Brill. DOI: 10.1163/9789004378308_004.

Bates, Douglas, Martin Maechler, 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.

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(19). 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. DOI: 10.1353/jsl.2021.a923057.

Espinal, M. Teresa & Susagna Tubau. 2019. Response systems: The syntax and semantics of fragment answers and response particles. *Annual Review of Linguistics* 5. 261–287. DOI: 10.1146/annurev-linguistics-011718-012613.

Geist, Ljudmila & Sophie Repp. 2023. Responding to negative biased questions in Russian. In Petr Biskup, Marcel Börner, Olav Mueller-Reichau & Iuliia Shcherbina (eds.), *Advances in formal Slavic linguistics 2021* (Open Slavic Linguistics 8), 145–177. Berlin: Language Science Press. DOI: 10.5281/zenodo.10123641.

Goodhue, Daniel. 2022. Isn't there more than one way to bias a polar question? *Natural Language Semantics* 30(4). 379–413. DOI: 10.1007/s11050-022-09198-2.

Goodhue, Daniel & Michael Wagner. 2018. Intonation, *yes* and *no. Glossa: a journal of general linguistics* 3(1). 5. 1–45. DOI: 10.5334/gjgl.210.

Gruet-Skrabalova, Hana. 2015. Verbs and particles in minimal answers to *yes-no* questions in Czech. 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*, 197–215. Frankfurt am Main: Peter Lang.

Gruet-Skrabalova, Hana. 2016. Yes or no, or how to answer a negative question. *Linguistica* 56(1). 127–142. DOI: 10.4312/linguistica.56.1.127-142.

Gunlogson, Christine. 2002. Declarative questions. In Brendan Jackson (ed.), *SALT 12: Proceedings from the 12th conference on Semantics and Linguistic Theory*, 124–143. Ithaca, NY: CLC Publications. DOI: 10.3765/salt.v12i0.2860.

Häussler, Jana & Tom Juzek. 2017. Hot topics surrounding acceptability judgment tasks. In Sam Featherston, Robin Hörnig, Reinhild Steinberg, Birgit Umbreit & Jennifer Wallis (eds.), *Proceedings of Linguistic Evidence 2016: Empirical, theoretical, and computational perspectives*. Tübingen: University of Tübingen. DOI: 10.15496/publikation-19039.

Kramer, Ruth & Kyle Rawlins. 2011. Polarity particles: An ellipsis account. In Suzi Lima, Kevin Mullin & Brian W. Smith (eds.), *NELS 39: Proceedings of the Thirty-Ninth Annual Meeting of the North East Lingusitic Society*, vol. 2, 479–492. Amherst, MA: GLSA Publications.

Kramer, Ruth & Kyle Rawlins. 2012. An ellipsis approach to answer particles in positive and negative contexts. Presented in the Workshop on the Syntax of Answers to Polar Questions (Newcastle University).

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.

Ladd, Robert D. 1981. A first look at the semantics and pragmatics of negative questions and tag questions. In Roberta Hendrick, Carrie Masek & Mary Frances Miller (eds.), *CLS 17: Papers from the seventeenth regional meeting of the Chicago Linguistics Society*, 164–171. Chicago, IL: Chicago Linguistic Society.

Kateřina Hrdinková & Radek Šimík

Lenth, Russell V. 2024. emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.10.2.090002. https://rvlenth.github.io/emmeans/.

Liberman, Mark & Ivan Sag. 1974. Prosodic form and discourse function. In *CLS 10: Papers from the tenth regional meeting of the Chicago Linguistic Society*, 416–427. Chicago, IL: Chicago Linguistic Society.

Maldonado, Mora & Jennifer Culbertson. 2023. *You say yes, I say no*: Investigating the link between meaning and form in response particles. *Glossa: a journal of general linguistics* 8(1). 1–41. DOI: 10.16995/glossa.9185.

R Core Team. 2021. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. https://www.r-project.org.

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. Brill. DOI: 10.1163/9789004183988_008.

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.

Staňková, Anna. 2023. *The expression of speaker's bias in Czech polar questions*. Praha: Charles University. (MA thesis).

Staňková, Anna & Radek Šimík. to appear. Negation in Czech polar questions. *Journal of Slavic Linguistics*.

Starschenko, Alexej & Marta Wierzba. 2023. L-Rex Linguistic rating experiments [software], version 1.0.2. GNU General Public License v3.0. https://github.com/2e2a/l-rex/.

Chapter 9

ABA in Russian adjectives, subextraction, and Nanosyntax

Daniar Kasenov

HSE University / Lomonosov Moscow State University

This paper's core focus is the ABA pattern exhibited by a number of Russian adjectives in their degree paradigms (positive, comparative, superlative). While the surface pattern seems to be a counterexample to the *ABA generalization about adjectival degree paradigms stated in Bobaljik 2012, a more involved exploration of Russian adjectival morphology shows that there are more classes of Russian adjectives that are problematic for the contemporary syntactic approaches to morphology (DM, Nanosyntax) given Bobaljik's containment hypothesis. This paper provides a description of these patterns in Russian adjectival morphology and provides an analysis for all the problematic classes in the framework of Nanosyntax, making use of two recent technical developments in the theory (Movement-Containing Trees of Blix 2022 and the subextraction Spell-Out algorithm of Caha & Taraldsen Medová 2022, 2023).

1 Introduction

This paper is concerned with an apparent counterexample to Bobaljik's ABA generalization in the domain of degree morphology (Bobaljik 2012). According to Bobaljik's cross-linguistic study, there is no language that has an adjective which has a suppletive stem α in positive and superlative and a suppletive stem β in the comparative. To give an example, a logically possible language English′, in which the adjective 'bad' has the forms *bad* 'bad.POS', *worse* 'bad.CMPR', *baddest* 'bad.SPRL', is impossible according to the generalization that Bobaljik draws from his typological study. The attested and unattested patterns of Bobaljik's three-cell paradigm are summarized in Table 1.

Daniar Kasenov

Table 1: (Un)attested suppletion patterns in adjectival paradigms

	POS	CMPR	SPRL	
AAA	*pretty*	*pretti-er*	*pretti-est*	English
ABB	*bad*	*worse*	*worst*	English
ABC	*bon-us*	*mel-ior*	*optim-us*	Latin
AAB	*bad*	*badd-er*	*worst*	Unattested!
ABA	*bad*	*worse*	*badd-est*	Unattested!

More generally, *ABA phenomena (such as the suppletion patterns in adjectives) are

> morphological patterns in which, given some arrangement of the relevant forms in a structured sequence, the first and third [forms] may share some property "A" only if the middle member shares that property as well. If the middle member is distinct from the first, then the third member of the sequence must also be distinct. (Bobaljik & Sauerland 2018: 1–2)

This work follows the tradition of examining *ABA phenomena (patterns of suppletion, syncretism, and other morphological properties) through the lens of the theories of morphology which assume the Single Engine Hypothesis (Marantz 2001), the idea that all complex expressions (including words) are built by syntax. Among others, these theories include the realizational approaches of Distributed Morphology (DM, Halle & Marantz 1994 and subsequent work) and Nanosyntax (Starke 2009 and subsequent work). In such approaches, *ABA phenomena are often understood structurally: cases of *ABA are due to the complex internal structure of examined wordforms, in which one form contains the other. Such analyses have been proposed for adjectival suppletion (Bobaljik 2012), case syncretism (Caha 2009), reflexive pronominal paradigms (Middleton 2021), numeral morphology (Sudo & Nevins 2022), and many other phenomena. Bobaljik himself has accounted for the ban on ABA via the containment structure provided in Figure 1.

Note, however, that containment structures do not rule out an ABA pattern by themselves. Such structures rule them out in conjunction with the widespread conception of (morphologically-conditioned) allomorphy phenomena in Distributed Morphology (the framework of choice in Bobaljik 2012) as contextual allomorphy (see Bonet & Harbour 2012 and Gouskova & Bobaljik 2020 for an
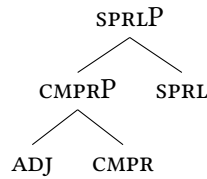
Figure 1: Containment structure for degree morphology

overview). The core logic is as follows: in DM, morphological forms are the results of Vocabulary Insertion rules which map syntactic objects onto morpho-phonological strings. Assuming that, allomorphy is understood as the same syntactic object being referenced by several mapping rules which differ by the contexts of their application. For example, the two allomorphs of the root of the English adjective *bad* are the results of two distinct insertion rules, which differ by their contexts (the *worse*-rule applied in the context of a CMPR node), as shown by the Vocabulary Insertion rules in (1).

(1) Vocabulary Insertion rules for $\sqrt{\text{BAD}}$
   a. $\sqrt{\text{BAD}} \leftrightarrow worse$ /__] CMPR]
   b. $\sqrt{\text{BAD}} \leftrightarrow bad$

In the absence of a specified rule for superlative forms, the containment structure ensures that the rule that applies in the context of the CMPR node also applies in the superlative form. Thus, the only allowed way for the ABA pattern to arise is accidental homophony: there should be two distinct insertion rules which just so happen to have the same morpho-phonological string as the result (like the Vocabulary Insertion rules in 2). Accidental homophonies are thought to be a rare occasion and definitely unlikely to hold across many "lexemes" (see Bobaljik & Sauerland 2018 for further discussion).

(2) Vocabulary Insertion rules for an ABA pattern *bad-worse-baddest*
   a. $\sqrt{\text{BAD}} \leftrightarrow bad$ /__] CMPR] SPRL]
   b. $\sqrt{\text{BAD}} \leftrightarrow worse$ /__] CMPR]
   c. $\sqrt{\text{BAD}} \leftrightarrow bad$

Given this background, the core data in this paper comes from a certain class of Russian adjectives, which exhibits an ABA pattern with respect to the presence of the augment affix *-(o)k*.[1] An example of such an adjective and its paradigm

---

[1] I follow Vanden Wyngaerd et al. (2020) in using the term AUGMENT for this sort of adjectival affix in Slavic languages.

Daniar Kasenov

is given in Table 2. Note that the alternation between *-(o)k* and *-(o)č* variants of the augment affix is due to morpho-phonological processes of palatalization. Without going into the depths of this phenomenon (see Blumenfeld 2003 and see Halle 1959 for a thorough treatment), it is relevant for our purposes that certain affixes turn segments /k g t s/ to /č ž č š/, respectively.

Table 2: *ABA-violating paradigm of Russian adjective *vysokij* 'high'

POS	CMPR	SPRL
vys-ok-ij	*vyš-e*	*vys-oč-aj-š-ij*
high-AUG-AGR	high-AGR	high-AUG-CMPR-SPRL-AGR
'high'	'higher'	'highest'

The problem is clear: Bobaljik's *ABA generalization rules out such patterns and yet they are found in these adjectives. While one could consider the observed surface pattern to be a reason to abandon Bobaljik's structure, this paper aims to provide an analysis of Russian *ABA-violating adjectives that does not abandon the containment structure for degree morphology.

This paper's goals are two-fold. The first goal is to provide a thorough examination of Russian adjectival morphology and to pinpoint the problems it poses for contemporary generative approaches to morphology. The second goal is to resolve said problems following existing Nanosyntax work on degree morphology (Caha et al. 2019, Vanden Wyngaerd et al. 2020, Caha & Taraldsen Medová 2023). The technical solution will be based upon two novel ideas in the Nanosyntax literature: the movement-containing trees (MCTs) of Blix (2022) and the subextraction spell-out algorithm of Caha & Taraldsen Medová (2022, 2023).

The paper is structured as follows. §2 examines the data and argues that there are three distinct classes of Russian adjectives whose degree paradigms are problematic for a Bobaljik-style approach. §3 introduces the theoretical framework of the analysis to come, namely, Nanosyntax. §4 presents my own solution to the puzzles posed by Russian adjectival morphology while introducing unfamiliar technical elements (Movement-Containing Trees and the subextraction spell-out algorithm) and showing which parts of the data require them. §5 concludes the paper.

2 The landscape of Russian adjectival morphology

This section presents the main patterns found in Russian adjectival morphology and discusses the existing allomorphs of the comparative and the superlative affixes, the periphrastic forms and three classes of adjectives that are problematic from the viewpoint of a simple containment structure for the degree morphology (Bobaljik 2012).

2.1 Basic adjectival morphology of Russian

Let us consider a basic Russian adjective with a basic degree paradigm. The adjective *glupyj* 'dumb' is an exemplar. As shown in Table 3, Russian degree morphology shows a straightforward containment of the comparative form *glup-ej-* in the superlative *glup-ej-š-*, once we consider the affix *-e* of the comparative form to be a φ-deficient agreement affix (or something else but crucially something irrelevant to the degree morphology).

Table 3: The basic degree paradigm of the adjective *glupyj* 'dumb'

POS	CMPR	SPRL
glup-yj	*glup-ej-e*	*glup-ej-š-ij*
dumb-AGR	dumb-CMPR-AGR	dumb-CMPR-SPRL-AGR
'dumb'	'dumber'	'dumbest'

The paradigm in Table 3 shows that the comparative affix is *-ej-* while the superlative affix is *-š-*. From here, I will refer to them as such, even if they arise in a form other than comparative or superlative, respectively. Most Russian adjectives have a similar paradigm, some of which are given in Table 4.

However, the presented synthetic paradigm is not the only way of forming Russian comparatives and superlatives and, for the sake of completeness of the overview of Russian adjectival morphology, I should introduce the other morphological strategies as well. The first thing to mention are the analytic forms *bolee*+ADJ and *samyj*+ADJ, as shown in Table 5 for the adjective *glupyj* 'dumb'. Since these are outside the scope of this paper, I refer the reader to Matushansky (2002) for discussion of analytic comparatives and Goncharov (2015) for discussion of analytic superlatives. However, I will make an important observation that some speakers of Russian outright reject synthetic superlative forms and tend to prefer the analytic form across the board. The interaction between analytic and

Daniar Kasenov

Table 4: More regular adjectives

POS	CMPR	SPRL
Paradigm for *umnyj* 'smart'		
umn-yj	*umn-ej-e*	*umn-ej-š-ij*
smart-AGR	smart-CMPR-AGR	smart-CMPR-SPRL-AGR
'smart'	'smarter'	'smartest'
Paradigm for *važnyj* 'important'		
važn-yj	*važn-ej-e*	*važn-ej-š-ij*
important-AGR	important-CMPR-AGR	important-CMPR-SPRL-AGR
'important'	'more important'	'most important'
Paradigm for *krasivyj* 'pretty'		
krasiv-yj	*krasiv-ej-e*	*krasiv-ej-š-ij*
pretty-AGR	pretty-CMPR-AGR	pretty-CMPR-SPRL-AGR
'pretty'	'prettier'	'prettiest'

synthetic comparative is more intricate and appears to be conditioned by many factors, including syllabic length of the adjective (see Kosheleva 2016 for discussion). Given this preference, some speakers may find the forms presented later in the text to be dubious. I put this difference in idiolects aside and leave them for a further sociolinguistic exploration.

Table 5: The analytic paradigm of the adjective *glupyj* 'dumb'

POS	CMPR	SPRL
glup-yj	*bolee glup-yj*	*samyj glup-yj*
dumb-AGR	more dumb-AGR	most dumb-AGR
'dumb'	'dumber'	'dumbest'

Another thing to note are the *nai*-superlatives, which consist of the prefix *nai*-and the synthetic superlative form, exemplified for the adjective *glupyj* 'dumb' in Table 6. These superlatives seem to be in free variation with regular synthetic superlatives, but some speakers consider them a "more marked" form conveying a focus on the degree. As far as I am aware, *nai*-superlatives never present suppletion/allomorphy patterns distinct from synthetic comparatives, hence, they will not be discussed in this paper in detail.

Table 6: Russian *nai*-superlatives

POS	CMPR	SPRL
glup-yj	*glup-ej-š-ij*	*nai-glup-ej-š-ij*
dumb-AGR	dumb-CMPR-SPRL-AGR	*nai*-dumb-CMPR-SPRL-AGR
'dumb'	'dumber'	'dumbest'

To round up this short section, for our purposes it is important that, in the regular case, the Russian comparative affix is *-ej-*, the Russian superlative affix is *-š-*, and the superlative form contains the comparative affix – while the comparative affix is *-ej*, the superlative is ADJ-*ej*-*š*-AGR (with *-ij* 'AGR.M.SG' being the used concord affix throughout the paper). These observations are perfectly in line with the theory of adjectival degree morphology laid out in Bobaljik 2012, according to which the superlative form is built on top of the comparative form (the containment hypothesis), as discussed in the introduction. In light of the accordance of the data of basic Russian adjectives with Bobaljik's theory, the next subsection is devoted to showing the adjectives which deviate from the basic pattern of comparatives being formed with *-ej-* and superlatives being formed with *-š-* on top of the comparative form.

2.2 Three problematic classes of adjectives

Exemplars of the three puzzling classes are the adjectives *strog-ij* 'strict', *rez-k-ij* 'harsh', and *vys-ok-ij* 'high'. Let us go through these adjectives one by one. The paradigm of the adjective *strog-ij* 'strict' (shown in Table 7) presents the following puzzle: despite it forming a zero-comparative with no overt comparative affix, the comparative affix *-aj-* appears in addition to the superlative affix *-š-* in the superlative form. Recall that alternations like *strog-*/*strož-* are due to the palatalization phenomena (Blumenfeld 2003) and are not relevant for the present study's focus on the morphologically-conditioned allomorphy in Russian adjectives.

One could argue that *-ajš-* should not be decomposed and rather be treated as an allomorph of the superlative affix for adjectives which form a zero-comparative. However, the adjective *krut-oj* 'cool' and similar ones (the paradigms of which are shown in Table 8) provide circumstantial evidence against such a hypothesis: *krut-oj* forms a zero-comparative *kruč-e* 'cooler' and a superlative form *krut-ej-š-ij* 'coolest', which suggests that the *-aj-* found in *strož-aj-š-ij* 'strictest' is

Daniar Kasenov

Table 7: The degree paradigm of the adjective *strog-ij* 'strict'

POS	CMPR	SPRL
strog-ij	*strož-e*	*strož-aj-š-ij*
strict-AGR	strict-AGR	strict-CMPR-SPRL-AGR
'strict'	'stricter'	'strictest'

the allomorph of *-ej-*, the comparative affix in Russian. And again, the *krut-*/*kruč-*, *čist-*/*čišč-* and *bogat-*/*bogač-* alternations are morpho-phonological in nature and are thus irrelevant to the morphosyntactically-conditioned allomorphy patterns discussed in this paper.

Table 8: Zero comparatives with *-ej-*

POS	CMPR	SPRL
The degree paradigm of the adjective *krut-oj* 'cool'		
krut-oj	*kruč-e*	*krut-ej-š-ij*
cool-AGR	cool-AGR	cool-CMPR-SPRL-AGR
'cool'	'cooler'	'coolest'
The degree paradigm of the adjective *čistyj* 'clean'		
čist-yj	*čišč-e*	*čist-ej-š-ij*
clean-AGR	clean-AGR	clean-CMPR-SPRL-AGR
'clean'	'cleaner'	'cleanest'
The degree paradigm of the adjective *bogatyj* 'rich'		
bogat-yj	*bogač-e*	*bogat-ej-š-ij*
rich-AGR	rich-AGR	rich-CMPR-SPRL-AGR
'rich'	'richer'	'richest'

Taking both *strog-ij* 'strict' and *krut-oj* 'cool' into account, a puzzling picture emerges: while these adjectives form zero-comparatives (without the comparative affix *-ej-*/*-aj-*), the comparative affix emerges in the superlative form. While one could argue that we are dealing with an affix *-ejš-*/*-ajš-*, such an analysis misses a clear parallel to the regular adjectives like *glup-yj* 'dumb' in the superlative form. In what follows, I assume that the *-ej-*/*-aj-* in the superlative form is the same morphological entity (= result of the same insertion rule) as the *-ej-*

found in the comparative forms of regular adjectives. Furthermore, I make the assumption that the alternation between *-ej-* and *-aj-* is morpho-phonological in nature, which is supported by the observation that *-aj-* allomorph is only found after /k/-, /g/-, and /x/-final adjectival stems (which are transformed into /č ž š/, respectively). Of course, this argument predicts that /k/-, /g/-, and /x/-final adjectives form their comparative forms with *-aj-* but such adjectives always form zero-comparatives, so the prediction cannot be tested.

Assuming that the overtness of the comparative affix is the default option (*-ej-*/*-aj-* is the default allomorph), the pattern of zero-comparatives presents a non-trivial problem for a theory like Bobaljik's. In order to account for the covertness of CMPR in the comparative form *strože* 'stricter', one has to posit a VI rule like (3b) which expones CMPR as a zero in the context of adjectives like *krutoj* 'cool' and *strogij* 'strict', but then posit a more specified rule like (3a) which expones CMPR as the default allomorph *-ej-*/*-aj-* since without such a rule there would be no way for the CMPR to be exponed in the superlative form.

(3) Zero-comparatives require accidental homophony of CMPR
    a. CMPR $\leftrightarrow$ *-ej-*/*-aj-* /X]__] SPRL] where $X \in \{\sqrt{\text{STRICT}}, \sqrt{\text{COOL}}, ...\}$
    b. CMPR $\leftrightarrow \varnothing$ /X]__ where $X \in \{\sqrt{\text{STRICT}}, \sqrt{\text{COOL}}, ...\}$
    c. CMPR $\leftrightarrow$ *-ej-*/*-aj-*

Given that having a zero-comparative is a property of multiple lexical items, the accidental homophony solution appears dubious and, hence, I consider the pattern to be problematic for a straightforward DM approach to the presented data. In addition to that, there are two other problematic classes of adjectives left to be presented in this section, the first of which is exemplified by the adjective *rez-k-ij* 'harsh'. I dub this class AUGMENT ADJECTIVES, borrowing the term for the *-(o)k-* affix from Vanden Wyngaerd et al. (2020). Example paradigms of augment adjectives are provided in Table 9, the alternation between *-(o)k-* and *-(o)č-* is due to palatalization phenomena and the presence of *o* is conditioned by stress (cf. *výsokij* and *΄redkij*) and, hence, both alternations are ignored for present purposes.

On the surface, the augment adjectives present the very same pattern as zero-comparative adjectives: in the context of some syntactic nodes (be it $\sqrt{\text{STRICT}}$ or the augment *-(o)k-*) the CMPR node is zero but is exponed as its default form once the SPRL node enters the structure. The problem posed by augment adjectives is thus the same posed by zero-comparatives, which raises the question of whether it is even sensible to draw a distinction between the two classes. However, foreshadowing my analysis, I will pursue the analytic strategy of deriving the patterns as portmanteaux – hence, the distinction between a root- and augment-triggered zero-comparative will prove useful in the later sections.

Daniar Kasenov

Table 9: Augment adjectives

POS	CMPR	SPRL
The degree paradigm of the adjective *rezkij* 'harsh'		
rez-k-ij	*rez-č-e*	*rez-č-aj-š-ij*
harsh-AUG-AGR	harsh-AUG-AGR	harsh-AUG-CMPR-SPRL-AGR
'harsh'	'harsher'	'harshest'
The degree paradigm of the adjective *žutkij* 'eerie'		
žut-k-ij	*žut-č-e*	*žut-č-aj-š-ij*
eerie-AUG-AGR	eerie-AUG-AGR	eerie-AUG-CMPR-SPRL-AGR
'eerie'	'eerier'	'eeriest'
The degree paradigm of the adjective *žarkij* 'hot'		
žar-k-ij	*žar-č-e*	*žar-č-aj-š-ij*
hot-AUG-AGR	hot-AUG-AGR	hot-AUG-CMPR-SPRL-AGR
'hot'	'hotter'	'hottest'
The degree paradigm of the adjective *gromkij* 'loud'		
grom-k-ij	*grom-č-e*	*grom-č-aj-š-ij*
loud-AUG-AGR	loud-AUG-AGR	loud-AUG-CMPR-SPRL-AGR
'loud'	'louder'	'loudest'

Now, consider the final class of adjectives: the *ABA-violating adjectives like *vys-ok-ij* (already mentioned in the introduction) in Table 10. Descriptively, the pattern is that the augment is not present in the comparative form but is present in the positive and superlative forms, which fits the ABA pattern as formulated by Bobaljik & Sauerland (2018), and is, thus, highly problematic for a theory that adheres to the containment hypothesis of Bobaljik 2012, which was put forward in order to exclude ABA patterns in degree morphology of adjectives.

Given the observations about the zero-comparative adjectives and the augment adjectives, however, we can decompose the *ABA-violating pattern into the combination of the observations about zero-comparative adjectives and augment adjectives in the following way. The ABA pattern consists of (i) the CMPR node being zero-exponed in comparative form only in the context of the augment (augment-adjectives pattern); (ii) the node adjacent to the adjectival root being zero-exponed in the comparative form only (zero-comparatives pattern). I believe that decomposing the ABA pattern into two distinct and attested patterns in Russian allows for a more grounded analysis (even though the two phenomena

Table 10: *ABA-violating adjectives

POS	CMPR	SPRL
Degree paradigm of the adjective *vys-ok-ij* 'high'		
vys-ok-ij	*vyš-e*	*vys-oč-aj-š-ij*
high-AUG-AGR	high-AGR	high-AUG-CMPR-SPRL-AGR
'high'	'higher'	'highest'
Degree paradigm of the adjective *red-k-ij* 'rare'		
red-k-ij	*rež-e*	*red-č-aj-š-ij*
rare-AUG-AGR	rare-AGR	rare-AUG-CMPR-SPRL-AGR
'rare'	'rarer'	'rarest'
Degree paradigm of the adjective *šyr-ok-ij* 'wide'		
šyr-ok-ij	*šyr-e*	*šyr-oč-aj-š-ij*
wide-AUG-AGR	wide-AGR	wide-AUG-CMPR-SPRL-AGR
'wide'	'wider'	'widest'
Degree paradigm of the adjective *gad-k-ij* 'disgusting'		
gad-k-ij	*gaž-e*	*gad-č-aj-š-ij*
disgusting-AUG-AGR	disgusting-AGR	disgusting-AUG-CMPR-SPRL-AGR
'disgusting'	'more disgusting'	'most disgusting'

are still problematic).

To sum up, we have discussed the three classes of Russian adjectives that pose a problem for the containment hypothesis of Bobaljik 2012. The descriptive contribution of this paper ends here. The next section is devoted to introducing Nanosyntax (but I presuppose the basic knowledge of the main tenets of Distributed Morphology). The section after that presents my Nanosyntactic analysis while introducing recent technical developments of the theory along the way.

3 Nanosyntax: the basics

This section presents the basics of Nanosyntax: its theoretical commitments and the inner workings of Nanosyntactic analyses, using the main building blocks of Russian adjectival morphology as the example (comparative *-ej-/-aj-* and superlative *-š-*). The first subsection presents the basic ideas behind Nanosyntax. The second subsection presents an analysis of regular adjectives in Russian.

Daniar Kasenov

3.1 The basics of Nanosyntax

Nanosyntax (Starke 2009, Baunaz & Lander 2018), like the mainstream Distributed Morphology approach to the syntax-morphology interface (Halle & Marantz 1994), is committed to the Single Engine Hypothesis (Marantz 2001): all complex expressions in languages are built by the same computational system (or module) – syntax. Unlike Distributed Morphology, however, Nanosyntax does not assume that individual syntactic terminals are morphemes / bundles of features (*pace* Embick 2015). Instead, Nanosyntax assumes a version of the One Feature – One Head thesis (Kayne 2005): all features are individual heads (and are, thus, privative). Thus, where DM would have a single "bundle" of, for example, φ-features on AGR nodes on adjectives (as in existing DM work on nominal concord, see Norris 2014 and the tree in Figure 2a), Nanosyntactic work on nominal concord assumes a hierarchy of feature heads (as in Caha 2023 and in the tree in Figure 2b).
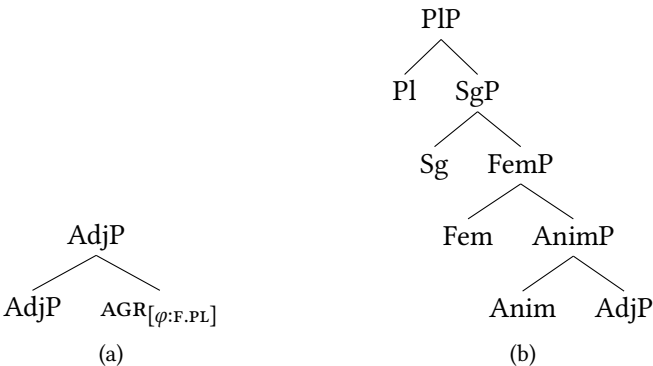
Figure 2: Structure for nominal concord in DM and Nanosyntax

The question is, how are the individual features grouped together to be matched to morphemes? This question requires a two-step answer. The first step is to introduce the notion of phrasal spell-out. While DM assumes that Vocabulary Insertion maps syntactic terminals onto morpho-phonological representations, Nanosyntactic work assumes that Vocabulary Insertion targets constituents. The idea is, then, that the bundles of features form syntactic constituents in order to be lexicalized together. Nanosyntax forms such constituents of features/feature-heads via syntactic movement according to the Spell-Out algorithm provided in (4). The core idea behind this algorithm is that after a new feature-head is merged, the resulting structure must be transformed into a structure that can be spelled-out, or to put it another way, whose subconstituents can be matched to existing lexical entries. The core property of the algorithm is that cumulative exponence

is preferred (the separate exponence of the newly merged F is only possible via the step in 4c).

(4) Spell-Out algorithm

    a. Merge F to XP and spell out

    b. If (a) fails, move Spec,XP to Spec,FP and spell out

    c. If (b) fails, move XP to Spec,FP and spell out

    d. If (c) fails, move to the next option in the previous cycle (backtracking)

Now the question lies in the precise nature of lexical entries in Nanosyntax and matching the feature structures to these entries. In Nanosyntax, lexical entries (or L-trees, to use proprietary terminology) are pairs of morpho-phonological representations and syntactic trees, an example is given in Figure 3.[2]

ZP ⇔ /blick/

Z   YP

Y   XP
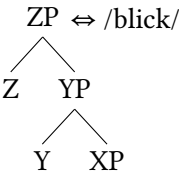
Figure 3: An example of an L-tree

The matching of constituents to L-trees is regulated by the Superset Principle, which states that an L-tree can be matched to any subconstituent of the structure in the L-tree. So, given the structure XP, the two lexical entries in Figure 4 match it (since XP is a subconstituent of both), which requires a principled way of choosing between the two matching L-trees.

ZP ⇔ α YP ⇔ β

Z   YP   Y   XP

Y   XP

Figure 4: Two matching L-trees for XP

---

[2]I do not touch on the topic of the syntax-semantics interface in Nanosyntax (or any Late Insertion theory) due to the complexity of the issue and its lack of direct relevance to the paper.

The choice between L-trees that match to the structure is regulated by the following principle: the L-tree with the least amount of structure not found in the syntactic constituent undergoing spell-out is chosen. So, between the two matching L-trees in Figure 4, the second one is to be preferred since it contains less "excess" structure.

Finally, to end this quick introduction, I want to emphasize that phrasal spell-out and spell-out-driven movement are the only operations available in the Nano-syntax machinery. There is no contextual allomorphy (or readjustment rules, or impoverishment rules, or any other familiar DM operation) in Nanosyntax, only portmanteaux, and, thus, a difference in form implies the presence of additional structure or a phonological analysis.

3.2 A case study: basic adjectival morphology of Russian

To recap the previous subsection, Nanosyntax assumes phrasal spell-out of syntactic constituents consisting of individual feature-heads which are formed via movement. The sequence of features (or f-seq) for degree morphology, according to Nanosyntactic work (see Caha et al. 2019 for the argumentation in favour of the split structure for degree morphology) is provided in (5).

(5) Nanosyntactic f-seq for degree morphology
     AdjP – Q – C1 – C2 – S1 – S2

In this subsection, I will provide a Nanosyntactic analysis of the basic paradigm of regular adjectives in Russian, repeated in Table 11. The main goal is to provide a lexical entry for the comparative affix *-ej-/-aj-* and for the superlative affix *-š-*.

Table 11: The paradigm of a regular adjective

POS	CMPR	SPRL
glup-yj	*glup-ej-e*	*glup-ej-š-ij*
dumb-AGR	dumb-CMPR-AGR	dumb-CMPR-SPRL-AGR
'dumb'	'dumber'	'dumbest'

The split comparative and split superlative structures proposed by Caha et al. (2019) increase the number of analytical choices we are facing. While a simple Bobaljik-style structure would require CMPR being realized as *-ej-/-aj-* and SPRL as *-š-*, the f-seq in (5) allows for various lexicalizations. Since the data of regular adjectives underdetermines the analysis, I will provide the lexical entries, which

allow for the analyses of the three problematic classes that will be presented in the next section. One thing to note is that the constituents in the L-trees for *-ej-* and *-š-* are remnant constituents (constituents, out of which something has moved, as shown by the presence of unary branching at the foot of the tree), see Figure 5. Such remnant constituents are exclusively associated with suffixes in the Nanosyntax literature (see Starke 2018).
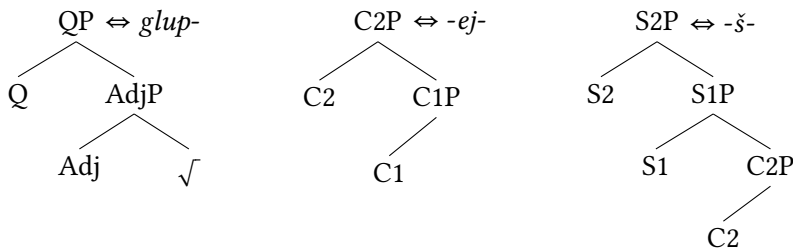
Figure 5: Lexical entries for regular adjective paradigms

Here, we shall go through the derivations step by step to show that the proposed lexical entries result in the observed paradigm. The thing to keep in mind is the standard Nanosyntax Spell-Out algorithm (Baunaz & Lander 2018) repeated in (6).

(6) Standard Nanosyntax Spell-Out algorithm

    a. Merge F to XP and spell out

    b. If (a) fails, move Spec,XP to Spec,FP and spell out

    c. If (b) fails, move XP to Spec,FP and spell out

    d. If (c) fails, move to the next option in the previous cycle (backtracking)

The derivation of the positive form is trivial: AdjP can be realized by the adjectival stem due to the Superset Principle and QP is the exact match of the lexical entry for *glup-*, as shown in Figure 7b. I want to note here that, in presenting the Nanosyntactic derivations, I will match subconstituents to the affixes in their underlying form. While I understand that it hurts the readability of the lexicalizations themselves, the clarity of the paper overall benefits from this decision, in my opinion.

The derivation of the comparative is also rather straightforward, as shown in Figure 7. Since the [C1 [Q AdjP]] structure does not match any lexical entry, as indicated by the double exclamation marks in Figure 7a, the next step is to move

Daniar Kasenov

(a) Start with AdjP

(b) Merge Q to AdjP and spell out
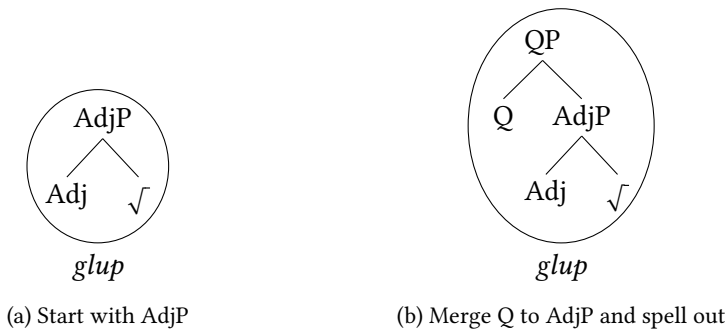
Figure 6: Deriving the positive form *glup-*

the specifier of QP to Spec,C1. However, there is no specifier of QP and, thus, the next step is to move QP to Spec,C1, which results in the proper lexicalization of C1P Figure 7b. Then, the C2 head is merged and the resulting structure does not match any lexical entry Figure 7c, which results in movement of QP to Spec,C2P Figure 7d. This structure results in the observed form *glup-ej-*, given our lexical entries.

The derivation of the superlative form is more complex and requires backtracking, the final step in the spell out algorithm provided in (6). After S1 is merged, there is no licit lexicalization even with movement of QP to Spec,S1P Figure 8b and movement of the whole comparative structure to Spec,S1P Figure 8c. The reason for this is that the lexical entry for the superlative affix *-š-* requires there to be a subconstituent with [C2P [C2]] at its foot, which isn't present at this point in the derivation. Hence, backtracking happens and the procedure goes back to the "next option in the previous cycle" step, namely, movement of C1P to Spec,C2P Figure 8d. After that, merging S1 Figure 8e and moving C1P to Spec,S1P results in a lexicalizable structure Figure 8f. After merging S2 Figure 8g and moving C1P to Spec,S2P Figure 8h, we end up with a structure that is realized as the observed form for the superlative *glup-ej-š-*.

Even though it is not necessary for an analysis of the paradigm of regular adjectives on its own, the core property of the provided analysis is that lexical entries for *-ej-* and *-š-* overlap in their inclusion of the C2 head: superlatives thus require backtracking (informally, splitting of C2 from C1 in lexicalization) and this property of the presented analysis will become relevant in the analysis of zero-comparatives, which is presented in the following section, along with the analyses for augment adjectives and *ABA-violating adjectives.

(a) Merge C1 to QP

(b) Move QP to Spec,C1P

(c) Merge C2 to C1P

(d) Move QP to Spec,C2P
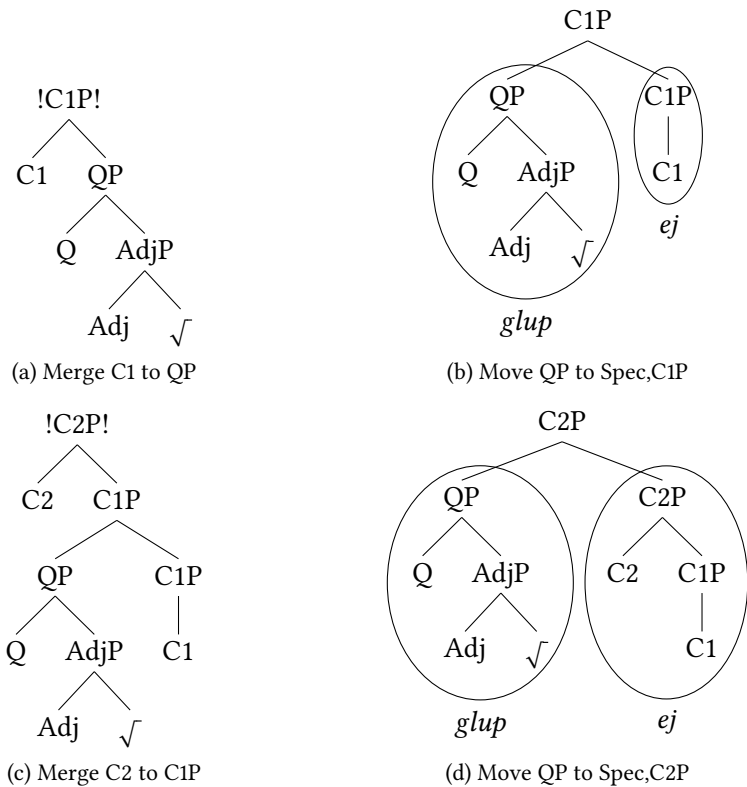
Figure 7: Deriving the comparative form *glup-ej-*

Daniar Kasenov
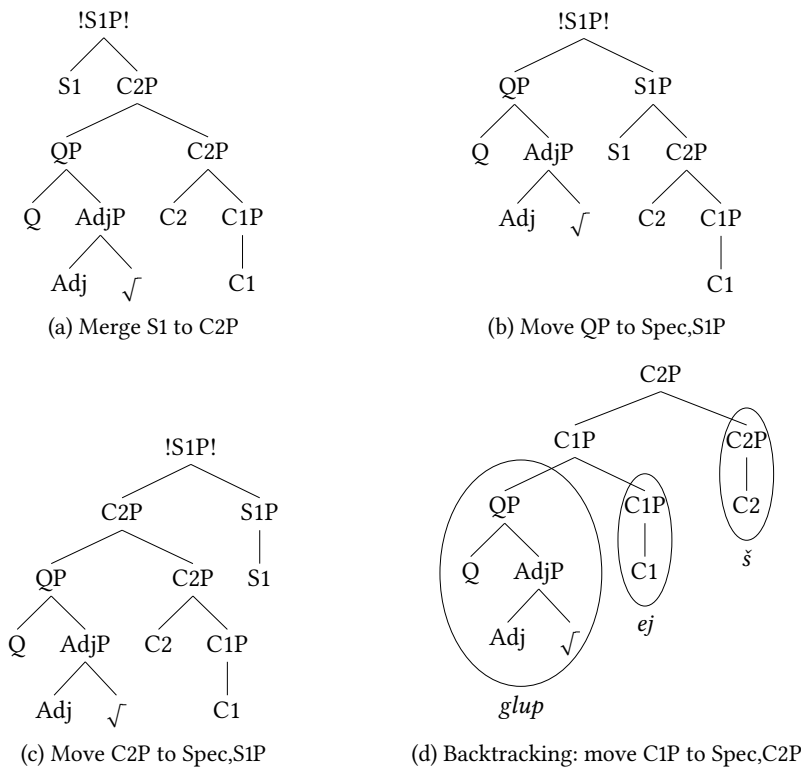
(a) Merge S1 to C2P

(b) Move QP to Spec,S1P

(c) Move C2P to Spec,S1P

(d) Backtracking: move C1P to Spec,C2P

Figure 8: Deriving the superlative form *glup-ej-š-*

(e) Merge S1 to C2P

(f) Move C1P to Spec,S1P

(g) Merge S2 to S1P

(h) Move C1P to S2P
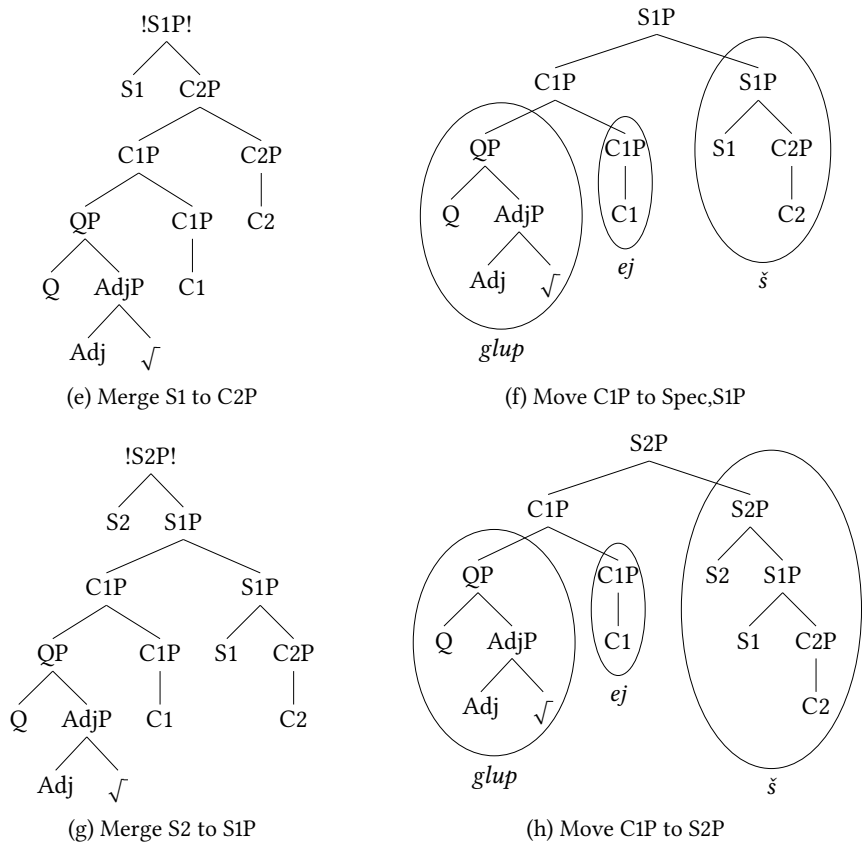
Figure 8: Deriving the superlative form *glup-ej-š-* (continued)

Daniar Kasenov

4 Analysis of three problematic classes

This section presents the Nanosyntactic analysis of the three problematic adjective classes: *strogij*-type adjectives (zero-comparatives), *rezkij*-type adjectives (augment adjectives), and *vysokij*-type adjectives (*ABA-violating adjectives). In §4.1, I present an analysis of zero-comparatives and introduce the notion of Movement-Containing Trees (Blix 2022) along the way. In §4.2, I present an analysis of augment-adjectives and introduce the novel spell-out algorithm of Caha & Taraldsen Medová (2022, 2023). Finally, §4.3 puts the analyses in §4.1 and §4.2 together to derive the ABA pattern established in the introduction. The core idea behind the analyses lies in the backtracking step forced by the L-tree of the superlative affix, as discussed in the previous section: in all three analyses, the backtracking step will trigger re-bundling of the features resulting in the exponence of the comparative and the augment.

4.1 Zero-comparatives: the need for movement-containing trees

Let me repeat the pattern and the problem for a DM-style approach posed by zero-comparative here. The basic pattern is as follows: the comparative affix *-ej-*/*-aj-* is absent from the comparative form itself, but arises in the decomposition of the superlative form, as shown in the paradigm in Table 12 for the adjective *strogij* 'strict'.

Table 12: The degree paradigm of the adjective *strog-ij* 'strict'

POS	CMPR	SPRL
strog-ij	*strož-e*	*strož-aj-š-ij*
strict-AGR	strict-AGR	strict-CMPR-SPRL-AGR
'strict'	'stricter'	'strictest'

The problem for a DM-style analysis was that one appears to need a zero-insertion rule for the CMPR node, which is sensitive to the adjacent adjective. However, this rule needs to be overridden in the superlative form, which results in an accidental homophony for the default VI rule for CMPR and the rule which is sensitive to both adjective and the presence of SPRL.

A basic Nanosyntax model (like the one introduced in Baunaz & Lander 2018) cannot accommodate these findings as well. In Nanosyntax, having a zero-comparative entails that the adjectival root (like *strog-*) has the comparative structure (C1 and C2 heads) in its lexical entry, as shown in Figure 9).

280
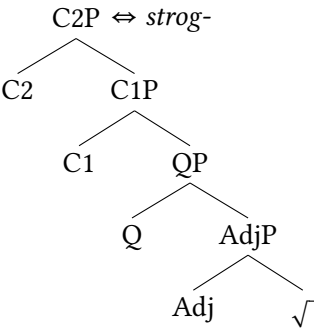
Figure 9: A putative lexical entry (L-tree) for *strog-*

The problem then is that when superlative heads are introduced into the deriva-
tion (S1 and S2), there is no way to trigger the overt comparative affix no matter
the lexical entry for the superlative affix – all comparative structure will be real-
ized either by the adjectival stem or by the superlative affix. Given backtracking,
all structure will be divided into the adjectival stem and the superlative affix, one
way or another (see Figure 10).

The solution for the problem of the comparative affix suddenly being overt in
the superlative form comes from the work of Hagen Blix arguing that phrasal
spell-out entails the possibility of spelling out constituents that "include" move-
ment. To be more substantive, Blix suggests that L-trees like Figure 11 are avail-
able in the lexicon, given the possibility to spell-out whole constituents (see
Blix 2022 for an exploration of this idea based on Kipsigis number morphology).
In accordance with an anonymous reviewer's comments, I emphasize that the
idea of Movement-Containing lexical entries is not a theoretical addition to the
Nanosyntactic project but rather an under-explored representational possibility.

For our purposes, the main consequence of the proposed lexical entry is that
there is no subconstituent of the L-tree in Figure 11 that contains both C1 and
the adjective to the exclusion of C2. Hence, if we force C2 to be spelled-out to-
gether with superlative structure (via the backtracking step, see the previous
subsection), the comparative affix will arise, as shown in Figure 12. Note that the
derivational steps are the same as with the regular adjectives – the only differ-
ence comes from the fact that adjectival stems like *strog-* are able to realize the
whole comparative structure.

To put it informally, Blix's (2022) proposal allows for a formalization of the
intuition that the comparative affix is zero-exponed in comparative only: it is
"inside" a portmanteau form, which is possible in the comparative form only due
to the internal structure of the lexical entry. Once S1 merges, the lexicalization
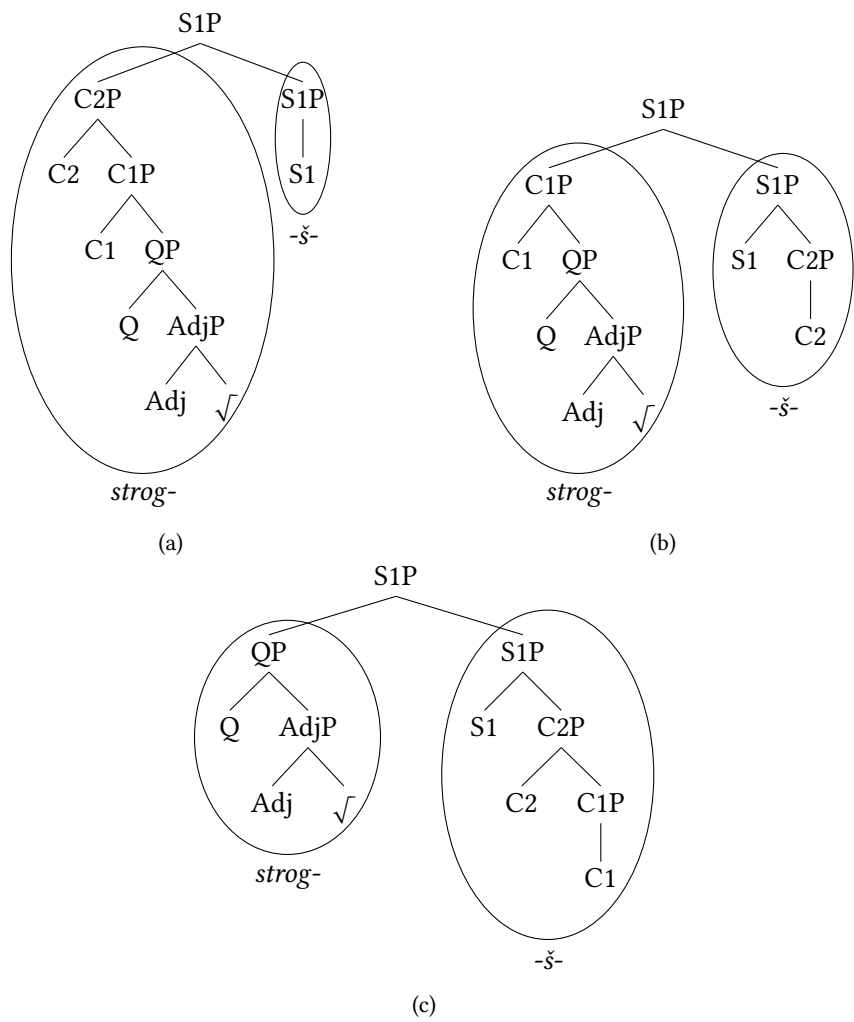
Figure 10: Possible lexicalizations of S1P given the L-tree for *strog-*

9 ABA in Russian adjectives, subextraction, and Nanosyntax
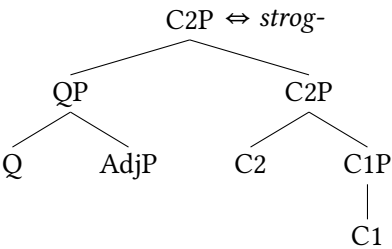
Figure 11: Movement-Containing Tree for the lexical entry of *strog-*
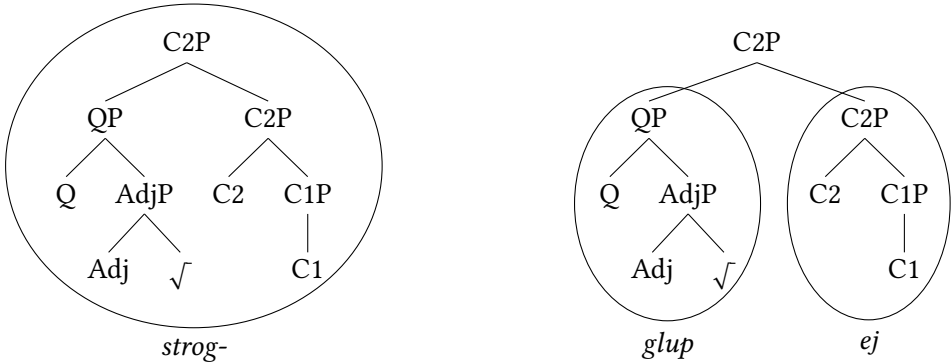
(a) Lexicalization of the comparative *strož-* (cf. *glup-ej-*)
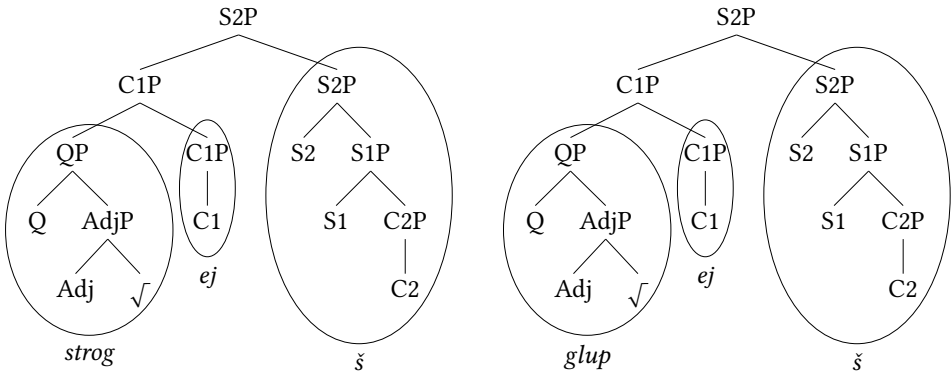
(b) Lexicalization of the superlative *strož-aj-š-* (cf. *glup-ej-š-*)

Figure 12: Lexicalizations of comparative and superlative forms of *strogij* 'strict'

requires bundling C2 together with S1, which results in QP being the only available subconstituent of the L-tree in Figure 11, forcing the exponence of *-ej-* in the superlative form.

4.2 Augment adjectives: the need for subextraction

Although the movement-containing trees (together with backtracking) have allowed us to capture the zero-comparative class of adjectives, the augment adjectives present an additional puzzle: we need the augment affix itself to realize the comparative structure. The desired lexicalization is as follows: there is some right branch that spells out Q, C1 and C2 together in the comparative form, and the superlative form must look like every other superlative form does: Adj, Q, C1 and C2-S1-S2 are lexicalized by distinct affixes.

The question is, how does one come to these lexicalizations given the regular Nanosyntax spell-out algorithm provided in Baunaz & Lander (2018) and Starke 2018. My answer is: it is impossible. Let us see why. Given the regular Nanosyntax algorithm, the fact that the comparatives only have a single affix implies the lexicalization in Figure 13.
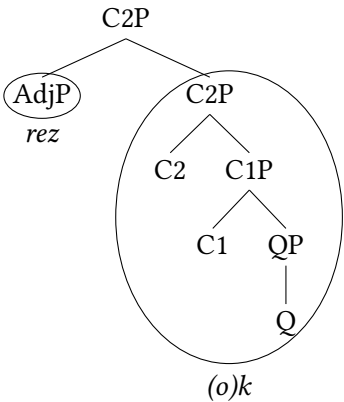
Figure 13: Lexicalization with a single affix according to regular Nanosyntax algorithm

The following problem then arises: any L-tree which matches with the right branch of the tree in Figure 13 will match to the subconstituent without C2 (given the Superset Principle). Hence, the predicted lexicalization for the superlative form does not include the comparative affix, contrary to the data, see Figure 14.

This problem motivates a theoretical addition to the Nanosyntax model. To account for augment adjectives, I employ the novel subextraction Nanosyntax algo-
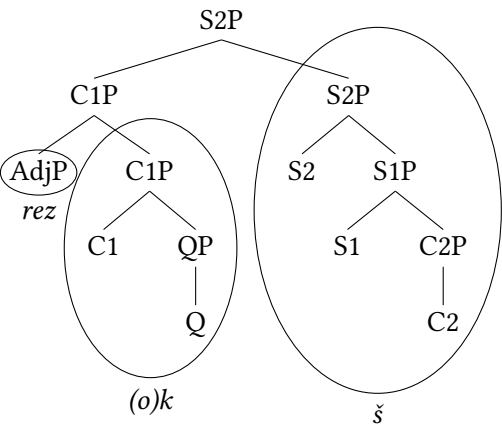
Figure 14: Lexicalization of superlative given the regular Nanosyntax algorithm

rithm, which is given in (7). Here, I take "non-remnant" to mean "not containing a unary branch". See Caha & Taraldsen Medová (2023) for a similar algorithm.

(7)   Subextraction spell-out algorithm (cf. Caha & Taraldsen Medová 2023)

   a. Merge F and spell-out
   b. If (a) fails, move the closest non-remnant constituent to Spec,FP
   c. If (b) fails, move the dominating node to Spec,FP (recursive step)
   d. If (c) fails, try the next option in the previous cycle

Compared to the regular Nanosyntax algorithm, the steps are the same in a single-affix structure like the steps in the standard algorithm: the first step is to move the specifier (the closest non-remnant constituent) and the second step is to move the whole structure (the dominating constituent). The difference comes with multiple affix structures: given the structure in Figure 15a, the first step would be to move the XP to Spec,HP (as in Figure 15d) and not YP (as was the case with the old algorithm, see the step in Figure 15c) – the novel algorithm makes heavy use of subextraction (moving a subconstituent from a specifier).

   The right branch of the structure in Figure 15d is a peculiar one because if there will be backtracking in the derivation, the next step in the derivation is the structure in Figure 15c – hence, if there is an L-tree for the right branch in Figure 15d, backtracking may force the exponence of the affix which realizes [YP [Y]]. The core idea is shown graphically in Figure 16.

Daniar Kasenov

(a) A multi-affix structure

(b) Merge H

(c) Old: move YP to Spec,HP

(d) New: move the closest non-remnant constituent (XP) to Spec,HP
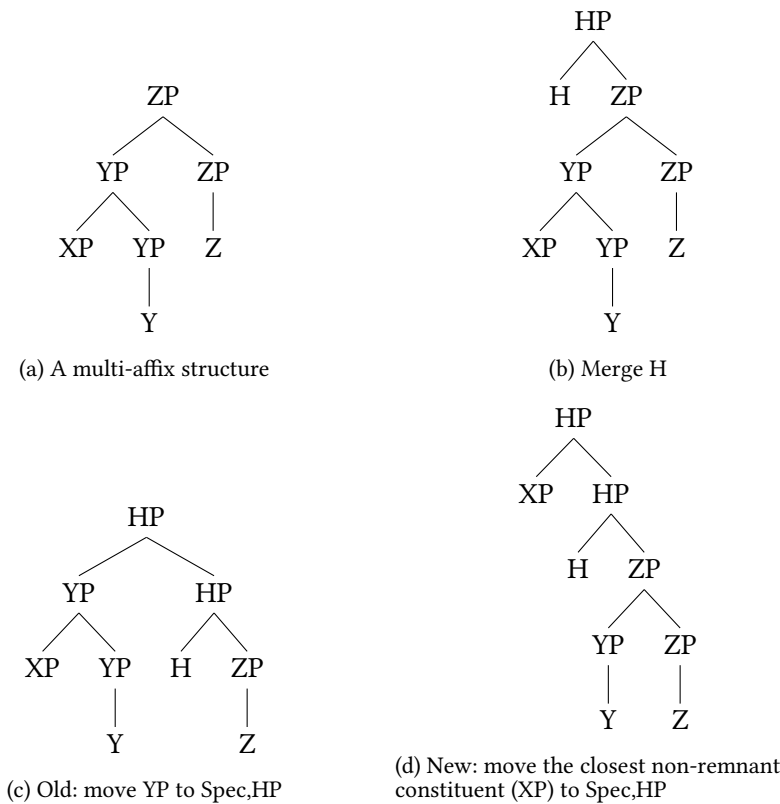
Figure 15: Complex specifier and old/new spell-out algorithm

This property is important given the lexical entry I propose for the augment affix, given in Figure 17. The main idea is that the backtracking step forced by the L-tree for the superlative affix *-š-* will trigger exponence of the affix which realizes [C1P [C1]] (namely, *-ej-*). Before I show the step-by-step derivation for the superlative form, let us go through the necessary steps for such a constituent to arise in the first place. The first step is to provide an L-tree for the adjectival stem, which does not include Q since it it realized by the augment.

Now, let us go through the whole derivation in order to show that the lexical entry in Figure 17, coupled with the subextraction spell-out algorithm, results in the observed morphological pattern of augment adjectives. After Q is merged (Figure 18b), the movement of AdjP to Spec,QP is necessary to lexicalize the structure (Figure 18c).

Then, after C1 is merged (Figure 19a), the movement of AdjP to Spec,C1P does not allow proper lexicalization (Figure 19b) and the next step is done (movement

9 ABA in Russian adjectives, subextraction, and Nanosyntax

(a) Lexicalization of the right branch   (b) The next step (after backtracking)
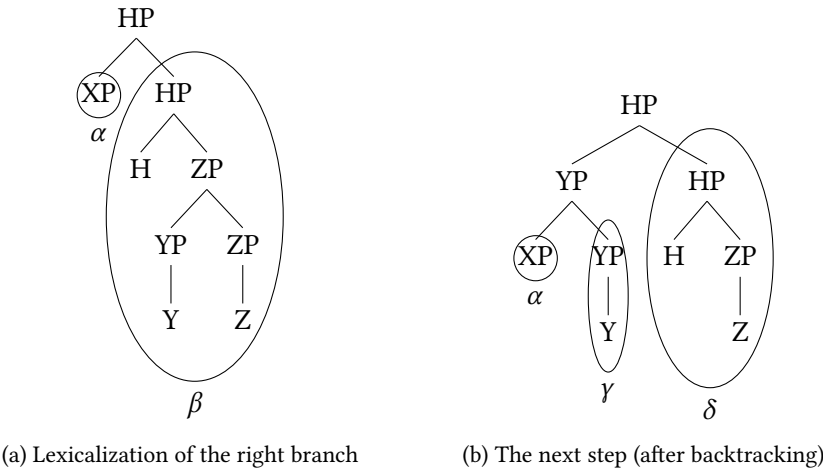
Figure 16: Affix emergence with backtracking and subextraction algorithm
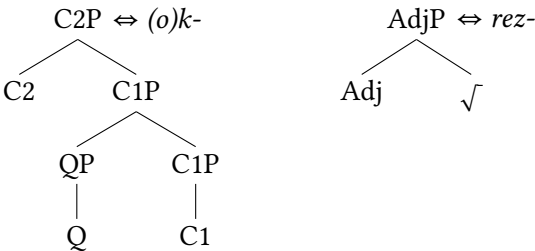
Figure 17: L-trees for the augment and for the stem *rez-*

(a) Start with AdjP   (b) Merge Q   (c) Move AdjP to Spec,QP
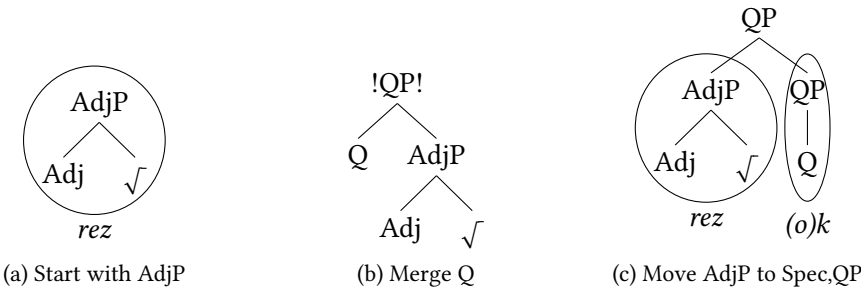
Figure 18: Deriving the positive form *rez-k-*

287

Daniar Kasenov

of QP to Spec,C1P, see Figure 19c, which results in a licit lexicalization. Note that this lexicalization does not correspond to any existing form – that is not an issue since C1P does not occur in absence of C2P.

(a) Merge C1

(b) Move the closest non-remnant constituent (AdjP) to Spec,C1P

(c) Move the dominating constituent (QP) to Spec,C1P
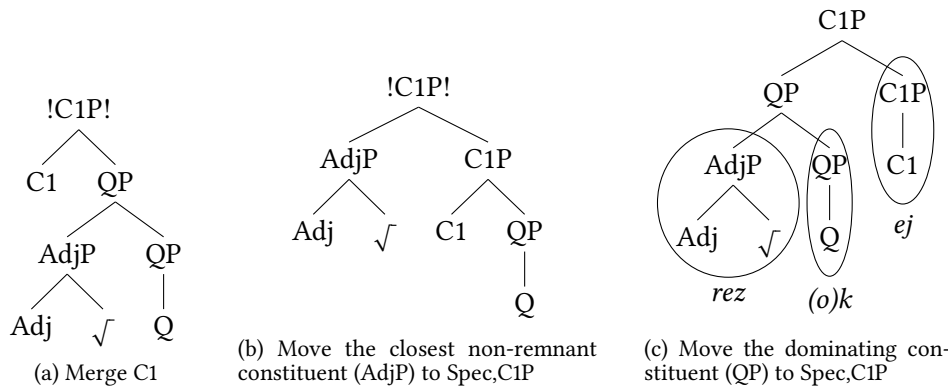
Figure 19: Lexicalizing C1P

After C2 is merged (Figure 20a), the first step of the subextraction algorithm is to move AdjP (and not QP, since it contains a unary branch [QP [Q]]) to Spec,C2P (Figure 20b), which results in a right branch that matches the L-tree for the augment, deriving the fact that the comparative affix is not present in the comparative form of augment adjectives.

(a) Merge C2

(b) Move the closest non-remnant constituent (AdjP) to Spec,C2P
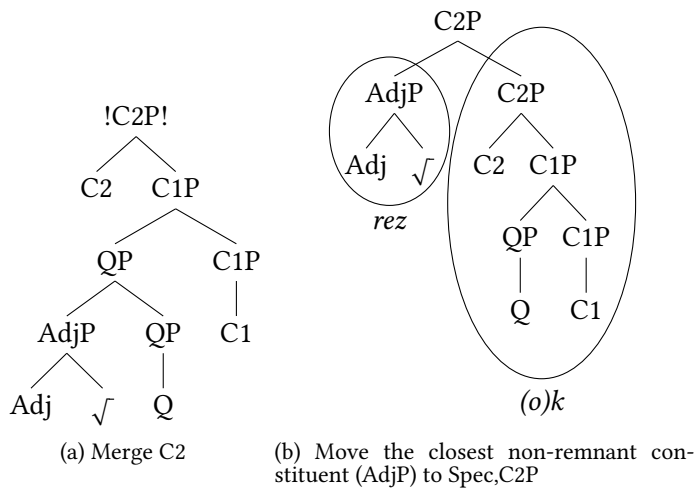
Figure 20: Lexicalizing C2P

Now, let us see what happens when the superlative structure is introduced into the derivation. After S1 is merged (Figure 21a), no operation in the cycle (movement of AdjP in Figure 21b, C2P in Figure 21c) results in a proper lexicalization. Thus, backtracking is necessary.

(a) Merge S1

(b) Move the closest non-remnant constituent (AdjP) to Spec,S1P

(c) Move the dominating constituent (C2P) to Spec,S1P
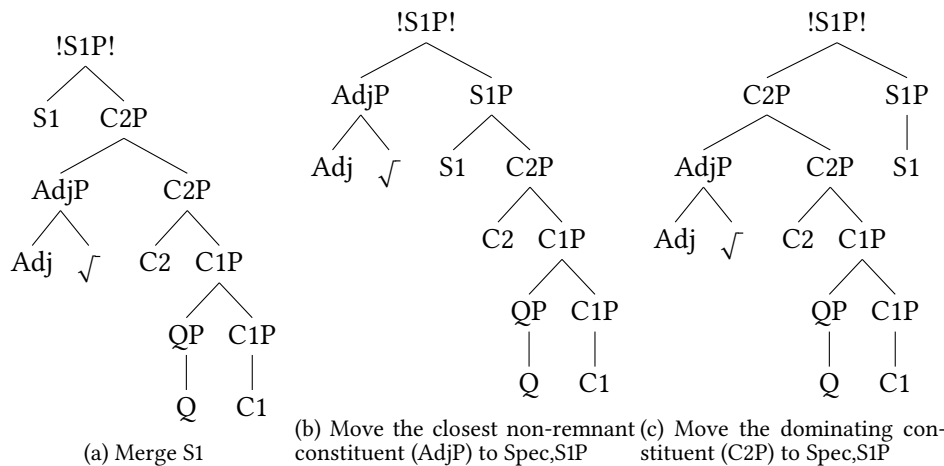
Figure 21: First cycle in lexicalizing S1P

However, the first step of backtracking will be movement of QP to Spec, C2P, resulting in a proper lexicalization more reminiscent of the regular adjectives, as shown in Figure 22.
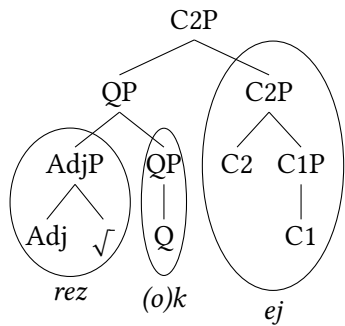
Figure 22: Backtracking: Move the dominating constituent (QP) to Spec,C2P

As was the case with regular adjectives, merging S1 still does not result in a proper lexicalization (as shown in Figure 23), no matter the operation (movement of AdjP, QP, C2P) and, thus, backtracking is necessary once again.

Daniar Kasenov

(a) Merge S1

(b) Move the closest non-remnant constituent (AdjP) to Spec,S1P

(c) Move the dominating constituent (QP) to Spec,S1P

(d) Move the dominating constituent (C2P) to Spec,S1P
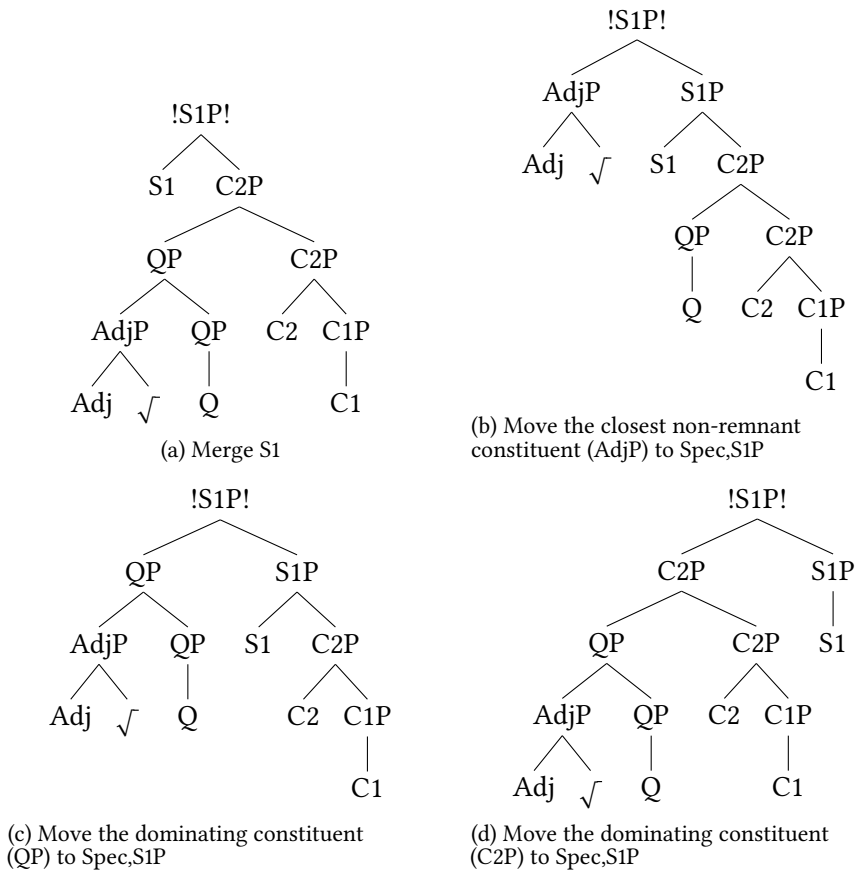
Figure 23: Lexicalizing S1P after first backtracking

The next backtracking step is moving the C1P to Spec,C2P, resulting in a structure that allows for future lexicalization of C2 together with S1 and S2, as shown in Figure 24. Note that this instance of backtracking mirrors the derivational steps necessary to lexicalize the superlative form of the regular adjectives.

After merging S1 (Figure 25a), movement of AdjP (Figure 25b) and QP (Figure 25c) does not result in a proper lexicalization – but movement of C1P does, as shown by Figure 25d.

The same thing happens after merging S2 (Figure 26a): movement of AdjP (Figure 26b) and QP (Figure 26c) does not result in a proper lexicalization, but movement of C1P does (Figure 26d). In the end, we derive the superlative form *rez-č-aj-š-ij* with an overt comparative affix.
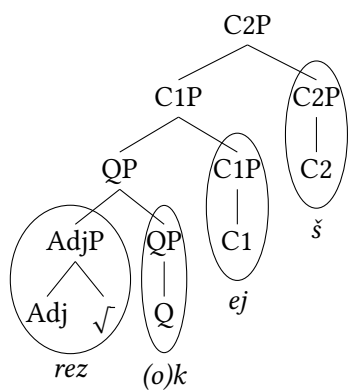
Figure 24: Backtracking: Move the dominating constituent (C1P) to Spec,C2P

This subsection has presented an analysis for the problematic degree morphology pattern of Russian augment adjectives. The core analytical move was the L-tree for the augment: its shape and the novel subextraction spell-out algorithm guarantee that backtracking (which happens due to C2 being in the L-tree for the superlative affix) results in a structure where [QP [Q]] is the only subconstituent matching to the L-tree of the augment, allowing for the independent realization of the [C1P [C1]] subconstituent. In the next subsection, I will synthesize ideas from the proposed analyses of zero-comparatives and augment adjectives in order to account for the *ABA-violating adjectives.

Daniar Kasenov

(a) Merge S1

(b) Move the closest non-remnant constituent (AdjP) to Spec,S1P

(c) Move the dominating constituent (QP) to Spec,S1P

(d) Move the dominating constituent (C1P) to Spec,S1P
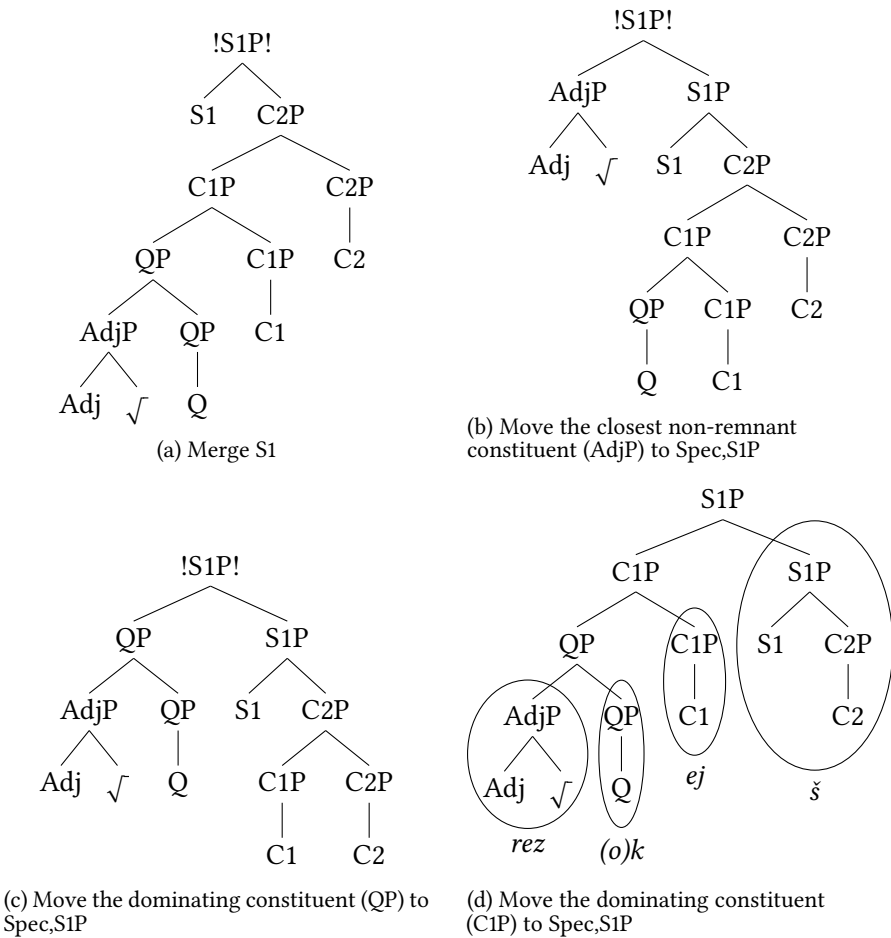
Figure 25: Lexicalizing S1P

(a) Merge S2

(b) Move the closest non-remnant constituent (AdjP) to Spec,S2P

(c) Move the dominating constituent (QP) to Spec,S2P

(d) Move the dominating constituent (C1P) to Spec,S2P
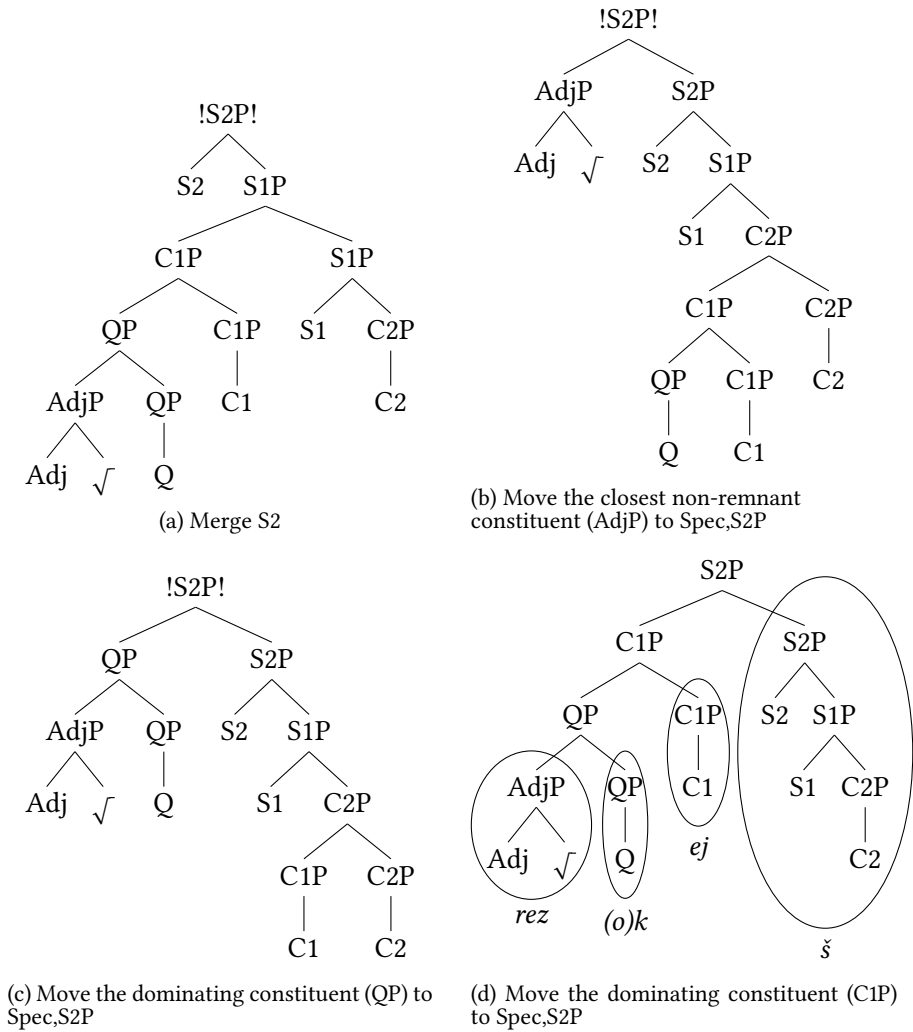
Figure 26: Lexicalizing S2P

Daniar Kasenov

4.3 *ABA-violating adjectives: putting the pieces together

At this moment, the solution to the *ABA-violating class of adjectives should be rather clear. The final lexicalized structure for comparatives of augment adjectives should be the L-tree for the adjectival stems of *ABA-violating adjectives, as given in Figure 27 for *vys-*. The solution to the ABA distribution of zero-exponence of the augment in the *ABA-violating class is thus the same as the solution to the distribution of zero-exponence of the comparative affix in the zero-comparative class: the lexical entry for the adjectival stem is such that the whole structure for the comparative form is a portmanteau.
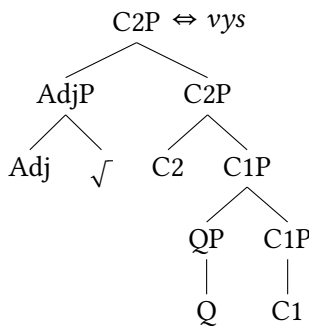
Figure 27: Lexical entry for *vys*

The core property of this proposal is that this lexical entry does not provide any subconstituent with AdjP that isn't AdjP itself or the whole tree, which means that the adjectival root will not be available to spell-out anything but AdjP in the positive form (resulting in *vys-ok-ij*, the lexicalization of which is given in Figure 28a) and the superlative form (resulting in *vys-oč-aj-š-ij*, the lexicalization of which is given in Figure 28b).

Note that, from the derivational point of view, there is no difference in the spell-out steps for augment adjectives and *ABA-violating adjectives – the only difference is the lexical property of *ABA-violating adjectives that they happen to have the correct right branch in their lexical entry, which creates an appearance of an ABA pattern with respect to the overtness of the augment. The core analytical contribution here is that the puzzling ABA pattern results from a combination of two independent phenomena (with theory-laden description) found in the domain of Russian adjectival morphology: the first phenomenon is the pattern of zero-comparatives (the adjectival root triggers zero-exponence of morphosyntactic material in the comparative form only), which is captured by positing a movement-containing L-tree for the adjectival stem. The second phenomenon

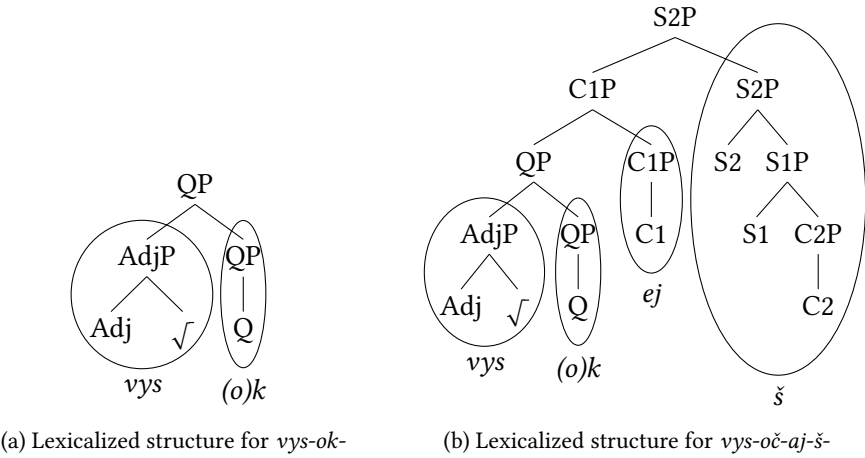(a) Lexicalized structure for *vys-ok-* (b) Lexicalized structure for *vys-oč-aj-š-*

Figure 28

is the pattern of augment adjectives (the augment affix zero-exponence of morphosyntactic material in the comparative form only). When one combines the derivational steps necessary for the analysis of augment adjectives with a movement-containing L-tree for adjectival stems such as *vys-*, an ABA pattern emerges.

Some theorists may take the fact that the morphological theory used in this work generates ABA patterns as a point of worry since many works (this paper included) have taken the impossibility of such patterns as the starting point of the investigation. However, recent research on similar (pseudo-)ABA patterns (Middleton 2021, Davis 2021) have came to the conclusion that the middle cell (the B of ABA) needs to be a portmanteau – this paper can be seen as adding to the body of evidence in favor of this idea.

5 Conclusion

In this paper, I have provided a look into the adjectival morphology of Russian through the lens of the comparative–superlative containment hypothesis put forth by Bobaljik (2012). I have provided evidence for there being a number of adjectives whose morphological behavior in the comparative and the superlative forms is problematic for contemporary proposals that follow Bobaljik's general idea.

Although the reported surface patterns may be taken as counter-evidence to Bobaljik's claims, I have argued that the core ideas of his need not be abandoned and have proposed a Nanosyntactic analysis of the pattern building on the idea

Daniar Kasenov

of Movement-Containing Trees (which are implied by the notion of phrasal spell-out but have been only recently argued for and used by Blix 2022) and the novel spell-out algorithm which allows subextraction from specifiers (Caha & Tarald-sen Medová 2022, 2023).

Given that the morphological patterns discussed in this work are problematic for both Distributed Morphology and the standard version of Nanosyntax found in Baunaz & Lander 2018 and preceding work, it is possible to take the provided analysis as an argument for accepting the generative power of the version of Nanosyntax with the subextraction algorithm presented in this work.

Abbreviations

AUG augment
AGR agreement
CMPR comparative
F feminine
M masculine

PL plural
POS positive
SG singular
SPRL superlative

Acknowledgments

References

Baunaz, Lena & Eric Lander. 2018. Nanosyntax: The basics. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring Nanosyntax*, 3–56. Oxford: Oxford University Press.

Blix, Hagen. 2022. Interface legibility and nominal classification: A nanosyntactic account of Kipsigis singulatives. *Glossa: a journal of general linguistics* 7(1). 1–46. DOI: 10.16995/glossa.5825.

Blumenfeld, Lev. 2003. Russian palatalization and Stratal OT: Morphology and [back]. In Wayles Brown, Ji-yung Kim, Barbara Partee & Robert Rothstein (eds.), *Annual Workshop on Formal Approaches to Slavic Linguistics: The Amherst meeting, 2002* (Formal approaches to Slavic linguistics 11), 141–158. Ann Arbor, MI: Michigan Slavic Publications.

Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words.* Cambridge, MA: MIT Press.

Bobaljik, Jonathan David & Uli Sauerland. 2018. ABA and the combinatorics of morphological features. *Glossa: a journal of general linguistics* 3(1). 15. 1–34. DOI: 10.5334/gjgl.345.

Bonet, Eulàlia & Daniel Harbour. 2012. Contextual allomorphy. In Jochen Trommer (ed.), *The morphology and phonology of exponence* (Oxford Studies in Theoretical Linguistics), 195–235. Oxford: Oxford University Press.

Caha, Pavel. 2009. *The nanosyntax of case.* Tromsø: University of Tromsø. (Doctoral dissertation).

Caha, Pavel. 2023. The layered structure of concord. In Pavel Caha, Pavel Rudnev & Svetlana Toldova (eds.), *The many facets of agreement* (LINCOM Studies in Theoretical Linguistics (LSTL) 68), 3–36. München: LINCOM GmbH.

Caha, Pavel, Karen De Clercq & Guido Vanden Wyngaerd. 2019. The fine structure of the comparative. *Studia Linguistica* 73(3). 470–521. DOI: 10.1111/stul.12107.

Caha, Pavel & Lucie Taraldsen Medová. 2022. On the feminine singular declension of the Russian demonstrative. Conference paper. 15th European Conference on Formal Description of Slavic Languages (FDSL 15). https://is.muni.cz/publication/2222757/fdsl15-dem.pdf.

Caha, Pavel & Lucie Taraldsen Medová. 2023. Syncretism and markedness paradoxes in the Russian demonstrative declension. Manuscript, Masaryk University. https://ling.auf.net/lingbuzz/007511.

Davis, Colin P. 2021. Case-sensitive plural suppletion in Barguzin Buryat: On case containment, suppletion typology, and competition in morphology. *Glossa: a journal of general linguistics* 6(1). 116. 1–26. DOI: 10.16995/glossa.5885.

Embick, David. 2015. *The morpheme: A theoretical introduction.* Berlin, München, Boston: De Gruyter Mouton.

Goncharov, Julie. 2015. *In search of reference: The case of the Russian adjectival intensifier 'samyj'.* Toronto: University of Toronto. (Doctoral dissertation).

Gouskova, Maria & Jonathan David Bobaljik. 2020. Allomorphy and vocabulary insertion. Manuscript, NYU and Harvard. https://ling.auf.net/lingbuzz/005388.

Halle, Morris. 1959. *The sound pattern of Russian: A linguistic and acoustical investigation.* The Hague: Mouton.

Daniar Kasenov

Halle, Morris & Alec Marantz. 1994. Some key features of Distributed Morphology. In Andrew Carnie, Heidi Harley & Toby Bures (eds.), *Papers on phonology and morphology* (MIT Working Papers in Linguistics 21), 275–288. Cambridge, MA: MITWPL.

Kayne, Richard S. 2005. *Movement and silence* (Oxford Studies in Comparative Syntax). Oxford: Oxford University Press.

Kosheleva, Daria. 2016. Rival forms of comparatives in Russian. *Poljarnyj vestnik: Norwegian Journal of Slavic Studies* 19. 42–50. DOI: doi.org/10.7557/6.4059.

Marantz, Alec. 2001. Words. Handout, MIT. https://babel.ucsc.edu/~hank/mrg.readings/Marantz_words.pdf.

Matushansky, Ora. 2002. More of a good thing: Russian synthetic and analytic comparatives. In Jindřich Toman (ed.), *Formal Approaches to Slavic Linguistics 10: The second Ann Arbor Meeting 2001*, 143–161. Ann Arbor, MI: Michigan Slavic Publications.

Middleton, Jane. 2021. Pseudo-ABA patterns in pronominal morphology. *Morphology* 31(4). 329–354. DOI: 10.1007/s11525-021-09377-7.

Norris, Mark. 2014. *A theory of nominal concord.* Santa Cruz, CA: University of California. (Doctoral dissertation).

Starke, Michal. 2009. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1: Special issue on Nanosyntax). 1–6. DOI: 10.7557/12.213.

Starke, Michal. 2018. Complex left branches, spellout, and prefixes. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring Nanosyntax*, 239–249. Oxford: Oxford University Press.

Sudo, Yasutada & Andrew Nevins. 2022. No ABA patterns with fractionals. In Linnaea Stockall, Luisa Martí, David Adger, Isabelle Roy & Ouwayda Sarah (eds.), *For Hagit: A celebration* (QMUL Occasional Papers in Linguistics 47), 1–10. London: Department of Linguistics Queen Mary University of London. https://www.qmul.ac.uk/sllf/media/sllf-new/department-of-linguistics/hagit-borer-celebration/Sudo_and_Nevins.pdf.

Vanden Wyngaerd, Guido, Michal Starke, Karen De Clercq & Pavel Caha. 2020. How to be positive. *Glossa: a journal of general linguistics* 5(1). 23. 1–34. DOI: 10.5334/gjgl.1114.

Chapter 10

Two types of secondary imperfectives: Evidence from Polish and Bulgarian

Dorota Klimek-Jankowska[a] & Vesela Simeonova[b]

[a]University of Wrocław [b]University of Graz

Secondary imperfective (SI) morphology differs in its productivity in Polish (PL) and Bulgarian (BG): in PL, the SI morphology combines with some but not all prefixes. By contrast, almost every BG perfective verb has a SI variant. To our knowledge, there is no research that has attempted to get a closer understanding of the source of this discrepancy. To fill in this niche, we conducted a comparative study of the interaction of SI morphology with different classes of aspectual prefixes in PL and BG and the meaning effects they give rise to. We present novel observations and account for them by proposing that there are two distinct layers at which SI morphemes are generated in BG and only one such layer in PL.

1 Introduction

There is an ongoing debate in the literature on Slavic aspect concerning the status of aspectual morphemes. Little agreement has been reached as to the status of secondary imperfective (SI) /(y)v/ in Slavic. The views vary as to where /(y)v/ is generated in the structure (see Schoorlemmer 1995, Babko-Malaya 1999, Istratkova 2004, Milićević 2004, Svenonius 2004a,b, Romanova 2004, Filip 2005, Di Sciullo & Slabakova 2005, Arsenijević 2006, Romanova 2007, Ramchand 2008a,b, Łazorczyk 2010, Markova 2011, Tatevosov 2011, 2015, Biskup 2012, 2019, Wiland 2012, Žaucer 2012, Rothstein 2020, Klimek-Jankowska & Błaszczak 2022, 2023, Kwapiszewski 2022). In most of these studies, generalizations about the status of secondary imperfective morphology are made based on the data from a single language. However, Slavic languages differ considerably in the productivity

of forming SI verbs. For example, the Bulgarian (BG) secondary imperfective is considerably more productive as compared to Polish (PL): almost every BG perfective verb has a SI variant (see Markova 2011, Rivero & Slavkov 2014; Nicolova 2017: 5.3.25), while in PL, SI morphology combines with some but not all prefixed verbs (see Łazorczyk 2010, Łaziński 2011, 2020, Wiemer et al. 2020, Kwapiszewski 2022, Klimek-Jankowska & Błaszczak 2022, 2023). While it is a well-known observation, there are no works that attempt to explain the locus of variation. Our goal in this paper is to fill that gap. We provide two major novel empirical observations:

1. Within Bulgarian: while all previous works on SI in general and in Bulgarian specifically treat it as a single class, we identify two distinct SI classes within Bulgarian with systematically different sets of formal derivational and semantic properties.

2. Between languages: we identify the locus of cross-linguistic variation in the productivity of SI between Polish and Bulgarian: the two languages share one of the two classes of SI and Polish lacks the second class of SI that Bulgarian has.

We conclude that the difference in the productivity of SI in Polish and Bulgarian is not random, but is systematically determined based on the range of derivational possibilities in the two languages, with respective semantic consequences.

To arrive at these conclusions, we tested the interaction of SI morphology in PL and BG with two classes of prefixes – lexical and purely perfectivizing prefixes – and the meaning effects these different combinations of the tested aspectual prefixes with SI morphemes give rise to. In this paper, the empirical scope of the environments tested is limited to past tense contexts. We discuss possible extensions in §5.

We show that in BG there are two types of SI morphemes that bear different meanings; PL has only one of these SI morphemes. We propose that the two types of SI in BG are realized in two different syntactic layers, following Cinque's (1999) model.

2 Background on the secondary imperfective in Polish and Bulgarian

2.1 Similarities between Polish and Bulgarian

In both PL and BG, aspectual distinctions are encoded on almost all verbs.[1] The least morphologically complex aspectual forms are primary imperfectives (bare, i.e. "unprefixed'' verbs) and they can be perfectivized by means of a prefix, cf. (1). Some prefixes do not change the verb in any way other than its aspectual value; they are called PURELY PERFECTIVIZING or EMPTY prefixes (Bogusławski 1963, Svenonius 2004a,b, Młynarczyk 2004, Willim 2006, Ramchand 2008a).[2]

(1) a. pisać   – napisać   PL
        write.IPFV   write.PFV

    b. piša   – napiša   BG
        write.IPFV   write.PFV

Another class of prefixes are so-called LEXICAL PREFIXES (Babko-Malaya 1999, Svenonius 2004a,b, Romanova 2004, 2007, Ramchand 2008a,b, Biskup 2012, 2019, a.m.o.): they have idiosyncratic meaning in that the prefix changes the lexical interpretation of the verb, but not in a predictable way, for example the prefix *prze-* in (2) and the prefix *pod-* in (3) have very different meaning in various verbs they participate in. Lexical prefixes cause idiosyncratic changes in the meaning of a verbal predicate that is not derivable from either the verb or the prefix, cf. (2) and (3).[3]

---

[1]The encoding of aspectual distinctions can be blocked e.g. for phonological reasons, or in certain loanwords, e.g. *printiram* 'print' only has one form in BG.

[2]Janda & Nesset (2010) emphasize that Russian has at least 16 prefixes forming natural perfectives (those perfectives which are not semantically distinct from the unprefixed base verb), which may suggest that they encode hidden distinctions. They propose that in the case of natural perfectives there is a semantic overlap between the meaning of the prefix and the meaning of the base verb and the diversity of prefixes used in natural perfectives follows from the fact that the base verbs from which they are derived fall into semantically diverse classes. Building on that, Janda & Lyashevskaya (2013) propose that the verbal prefixes act as classifiers in that they select verbs according to broad semantic traits, categorizing them the way numeral classifiers in some languages categorize nouns. We think that irrespective of the terminology used, there is a general consensus that the prefixes in natural perfectives do not modify the meaning of the base verbs but they may only impose selectional restrictions on the base verbs they combine with. Therefore, we will maintain the terminology 'purely perfectivizing'.

[3]Because of these properties, Romanova (2004), Svenonius (2004a,b), Ramchand (2004, 2008a,b), Łazorczyk (2010) argue that lexical prefixes are merged vP-internally.

Dorota Klimek-Jankowska & Vesela Simeonova

(2) a. kupić   – *prze*-kupić   PL
         buy.PFV   bribe.PFV
         'to buy' – 'to bribe'

     b. grać   – *prze*-grać   PL
         play.IPFV lose.PFV
         'play' – 'lose'

     c. łączyć   – *prze*-łączyć   PL
         connect.IPFV switch.PFV
         'connect' – 'switch'

(3) a. seštam *(se) – *pod*-seštam   BG
         recall   REFL remind.PFV
         'recall' – 'remind'

     b. budja   – *pod*-budja   BG
         wake.up.IPFV incite.PFV
         'wake up' – 'incite, instigate'

     c. igraja   – *pod*-igraja   BG
         play.IPFV   mock.PFV
         'play' – 'mock'

Furthermore, lexical prefixes (can) alter the argument structure/selectional restrictions of a verb, cf. (3)–(5).

(4) a. znać   {kogoś   / *uszkodzenia}   PL
         know.IPFV someone   damage
         'to know someone'

     b. do-znać   {*kogoś   / uszkodzenia}   PL
         suffer.PFV someone   damage
         'to suffer damage'

(5) mislja (*se) – za-mislja *(se)   BG
     think   REFL   consider   REFL
     'think' – 'consider'

Lexically prefixed perfective verbs are imperfectivized by means of an *-yw-* or *-a-* suffix in Polish and by a *-va* suffix or vowel alternations in Bulgarian, cf. (6). These imperfective forms derived from perfective verbs are called SECONDARY IMPERFECTIVE (SI). Table 1 shows more verbs from this morphological pattern.

(6) a. podpisać – podpisywać   PL
         sign.PFV   sign.SI

b. podpiša – podpisvam   BG
   sign.PFV   sign.SI

Table 1: Lexical prefixes and SI in Polish and Bulgarian

POLISH		BULGARIAN		ENGLISH
PFV	SI	PFV	SI	
podpisać	podpisywać	podpiša	podpisvam	'sign'
odpowiedzieć	odpowiadać	otgovorja	otgovarjam	'reply'
naprawić	naprawiać	popravja	popravjam	'repair'
wyjaśnić	wyjaśniać	objasnja	objasnjavam	'explain'
sprzedać	sprzedawać	prodam	prodavam	'sell'
opisać	opisywać	opiša	opisvam	'describe'

2.2 A major difference: SI productivity

In the previous section, we showed that SI is possible with lexically prefixed verbs both in PL and in BG. However, there is a major difference between PL and BG in that almost every BG verb can form SI (Dickey 2000: 11; Nicolova 2017: 5.3.25). Most verbs with empty prefixes also have SI forms, as illustrated in (7b). In PL, empty prefixed forms cannot form SI, see the ungrammatical form in (7a). The pattern described in example (7) is systematic in the two languages, as demonstrated in Tables 2 and 3.[4]

(7) a. pisać   – napisać   – *napisywać   PL
         write.IPFV   write.PFV   write.SI

     b. piša   – napiša   – napisvam   BG
         write.IPFV   write.PFV   write.SI

In other words, in BG, there is a morphological triplet for verbs with purely perfectivizing prefixes and a pair for lexically prefixed forms, as in Tables 2 and 4. And in PL, verbs form morphological pairs: either the bare imperfective and a verb with a purely perfectivizing prefix or the lexically prefixed imperfective

---

[4]The fact that one may find some rare instances of these verbs on the internet suggests that someone either used them creatively or mistakenly. Such rare uses may suggest that the two projections that we will argue for in §4 high SI and low SI are universally there in the hierarchy of projections but in some languages such as Polish, for example, the high SI morpheme generally does not merge in this position (it is blocked), but it may exceptionally be unblocked when used creatively or in speech production errors.

Dorota Klimek-Jankowska & Vesela Simeonova

Table 2: Purely perfectivizing prefixes and SI in Bulgarian

BULGARIAN			ENGLISH
IPFV	PFV	SI	
stroja	postroja	postrojavam	'build'
piša	napiša	napisvam	'write'
pŭrža	izpărža	izpăržvam	'fry'
gladja	izgladja	izglaždam	'iron'
broja	prebroja	prebrojavam	'count'
molja	pomolja	pomolvam	'ask'
četa	pročeta	pročitam	'read'
gubja	izgubja	izgubvam	'lose'
merja	izmerja	izmervam	'measure'
zvănja	pozvănja	pozvănjavam	'call'
čupja	sčupja	sčupvam	'break'

Table 3: Purely perfectivizing prefixes and SI in Polish

POLISH			ENGLISH
IPFV	PFV	SI	
budować	zbudować	*zbudowywać	'build'
pisać	napisać	*napisywać	'write'
smażyć	usmażyć	*usmażywać	'fry'
prasować	wyprasować	*wyprasowywać	'iron'
liczyć	policzyć	*policzać	'count'
prosić	poprosić	*popraszać	'ask'
czytać	przeczytać	*przeczytywać	'read'
gubić	zgubić	*zgubiać	'lose'
mierzyć	zmierzyć	*zmierzać	'measure'
dzwonić	zadzwonić	*zadzwaniać	'call'
łamać	złamać	*złamywać	'break'

304

and the derived SI, as in Tables 3 and 5. The crucial difference between the two languages in the two tables is marked with shading.[5]

Table 4: Purely perfectivizing vs. lexical prefixes and SI in Bulgarian

primary imperfective	perfective	secondary imperfective
N/A	podpiša 'sign'	podpisvam
N/A	poleja 'water'	polivam
stroja 'build'	postroja	postrojavam
piša 'write'	napiša	napisvam

Table 5: Purely perfectivizing vs. lexical prefixes and SI in Polish

primary imperfective	perfective	secondary imperfective
N/A	podpisać 'sign'	podpisywać
N/A	podlać 'water'	podlewać
budować 'build'	zbudować	*zbudowywać
pisać 'write'	napisać	*napisywać

While the morphological determinant of the restrictive SI in Polish is well-known – the availability of SI counterparts of perfective verbs depends on the prefix type, as we described above – and it is also well-known that Bulgarian SI is fully productive, these observations raise many questions that remain unanswered to date. The questions we address in this paper are whether there are semantic differences between the SI forms in BG that do not have equivalents in PL and the ones that do have equivalents in Polish and why the BG SI forms are not possible in Polish.

---

[5]We acknowledge that the type of classification of triplets and pairs that we are using to make this claim is not the only one that exists in the literature. In a very recent study on aspectual triplets in Russian, Czech, Polish, Wiemer et al. (2020) identify triplets based on a different set of criteria. They assume that lexically prefixed verbs also form triplets. We explain how we understand triplets and pairs Section, §2.1 that we use in this paper are based on the works cited in §2.

Dorota Klimek-Jankowska & Vesela Simeonova

3 Novel findings: two types of SI

Our first finding is that the pair SI in PL is equivalent to the pair SI in BG and it is ambiguous between the single ongoing and habitual reading. The second finding is that within Bulgarian, the triplet SI is qualitatively different from the pair SI. The triplet SI is habitual only, while the pair SI is ambiguous between single ongoing and habitual reading. This means that the properties of SI are not uniform across languages and even within the same language they are not homogeneous.

3.1 Pair SI in Bulgarian and Polish

Both in Bulgarian and in Polish, SI forms derived from lexically prefixed verbs are ambiguous between an ongoing reading, as in (8), and a habitual reading, as in (9). This is not idiosyncratic of a specific verb, but holds across the morphological paradigm represented in Table 1.

(8)   ONGOING CONTEXT

    a.   Kogato vljazoh   v ofisa na Ivan, toj (točno) {
        when enter.PFV.AOR.1SG in office of Ivan he just
        podpis-va-še   dokumenti / poprav-ja-še   koleleta /
        sign-SI-IMPF.3SG documents   repair-SI-IMPF.3SG bikes
        otgovar-ja-še   na imejli}.   BG
        reply-SI-IMPF.3SG to emails

    b.   Kiedy weszłam   do gabinetu Jana, (właśnie) {
        when entered.PFV.PST.1SG to office   John just
        podpis-yw-ał   dokumenty / naprawi-a-ł   rower /
        sign-SI-PST.3SG documents   repair-SI-PST.3SG bike
        odpowiad-a-ł   na maila}.   PL
        reply-SI-PST.3SG to email
        'When I entered John's office, he was (in the middle of) {signing
        documents / repairing bikes / replying to e-mails}.'

(9)   HABITUAL CONTEXT

    a.   Predi obiknoveno {podpis-va-še   dokumentite /
        before usually   sign-SI-IMPF.3SG documents
        poprav-ja-še   koleletata / otgovar-ja-še   na imejli} po-bărzo
        repair-SI-IMPF.3SG bikes.DEF   reply-SI-IMPF.3SG to emails faster
        ot   men, no veče   ne.   BG
        than me,   but already not

b. Kiedyś zwykle {podpis-yw-ał   dokumenty / na-prawi-a-ł
before usually sign-SI-PST.3SG documents   repair-SI-PST.3SG
rowery / odpowiad-a-ł   na maile} szybciej niż ja, ale teraz
bikes   respond-SI-PST.3SG to emails faster   than I   but now
już   nie.   PL
already not
'In the past, usually he (used to) {sign (the) documents / repair bikes /
respond to e-mails} faster than me but not anymore.'

All the verbs in Table 1 behave in a way analogous to the pattern shown in ex-
ample (9), allowing both habitual and ongoing readings. We were unable to find
any counterexamples.

3.2 Triplet SI in Bulgarian

The triplet SI in Bulgarian cannot be used with ongoing actions, cf. (10) and only
has habitual readings, cf. (11). The examples also show that the bare imperfective
is grammatical in both environments.[6]

(10) ONGOING CONTEXT

a. Kogato telefonăt   zvănna,   točno {păržeh   / *
when   phone.DEF rang.AOR.3SG just   fry.IMPF.1SG.IPFV
izpăržvah}   kjufteta.   BG
fry.IMPF.1SG.SI meatballs
'When the phone rang, I was (right in the middle of) frying meatballs.'

b. Kogato telefonăt zvănna,   točno {gladeh   / *
when   phone   rang.AOR.3SG just   iron.IMPF.1SG.IPFV
izglaždah}   drehi.   BG
iron.IMPF.1SG.SI clothes
'When the phone rang, I was (right in the middle of) ironing clothes.'

c. Kogato telefonăt zvănna,   točno si   {pišeh   / *
when   phone   rang.AOR.3SG just   REFL.GEN write.IMPF.1SG.IPFV
napisvah}   domašnoto.   BG
write.IMPF.1SG.SI homework.DEF
'When the phone rang, I was (right in the middle of) writing my
homework.'

---

[6]While we assume that it is always available in ongoing contexts, we do not claim that it is
always possible in every habitual context.

Dorota Klimek-Jankowska & Vesela Simeonova

(11) Habitual context

    a. Kogato praveh   zakuska, obiknoveno {păržeh   /
       when make.IMPF.IPFV.1SG breakfast usually   fry.IMPF.IPFV.1SG
       izpăržvah}   po   3 kjufteta na čovek.   BG
       fry.IMPF.SI.1SG DISTR 3 meatballs per person
       'When I made breakfast, I used to fry 3 meatballs per person.'

    b. Predi vinagi {gladeh   / izglaždah}   drehite
       before always iron.IMPF.IPFV.1SG iron.IMPF.SI.1SG clothes.DEF
       vednaga   sled prane.   BG
       immediately after washing
       'Before I always ironed the clothes immediately after washing.'

    c. Predi obiknoveno {pišeh   / napisvah}   po
       before usually   wrote.IMPF.IPFV.1SG wrote.IMPF.SI.1SG DISTR
       njakolko knigi na godina, no sega samo po   edna.   BG
       several   books per year, but now only DISTR one
       'In the past, I used to write several books per year, but now only one.'

This semantic pattern observed with triplet SI above is valid across the paradigm of triplet SI, a sample of which was presented in Tables 2 and 3. Since there is no triplet SI in Polish, this part of the data is not directly comparable between the two languages. In both single ongoing and habitual scenarios presented for Bulgarian in (10) and (11), Polish uses primary imperfective verbs only.

    To summarize, while PL uses primary IPFV verbs to render both habitual and ongoing readings in purely perfectivized verbs and the SI forms of such verbs are blocked, BG productively uses the SI forms of those verbs to exclusively "mark" the special kind of habitual reading (consisting of a series of temporally non-overlapping bounded events happening on separate occasions).

    Our novel observation is that the properties of SI morphology are not semantically uniform across the two languages. In the next section, we propose that the two types of SI morphemes merge at different syntactic positions in BG but not in Polish.

4 The syntax of SI

In order to formally capture the observations presented in the previous section, we propose that in BG, the two types of SI morphemes merge at two syntactic layers – one higher and one lower – while in PL, the low SI morpheme merges only in the lower one.

(12)   BG: SI_{high} >> SI_{low}
PL: SI_{low}

This proposal allows us to syntactically distinguish between the properties of lexical prefixes and purely perfectivizing ones (see §2). We argue that SI is not a uniform category within Bulgarian because it merges in two different syntactic positions with different properties each. Moreover, pair SI is equivalent across Bulgarian and Polish because it is merged in the same projection SI_{low} with the same properties. PL does not have aspectual triplets because the high SI morpheme is blocked and it cannot merge in the SI_{high} layer while Bulgarian developed a specialized habitual meaning of the SI_{high} morpheme. Because verbs with SI_{high} have obligatory habitual readings in BG, which are missing in PL, we propose that the SI_{high} morpheme merges in a projection corresponding to Cinque's (1999) Asp_{HAB}.

We assume Baker's (1985) Mirror Principle, according to which syntax reflects morphology and vice versa and the linearization of functional morphemes is syntactically motivated. Additionally, we follow Cinque (1999), who argues that there is a fixed hierarchy of functional projections which regulates the way adverbs and functional morphemes are merged in syntax. Based on a large survey of languages, Cinque shows that among temporal/aspectual affixes, e.g. repetitive, frequentative, terminative, continuative, retrospective, durative, progressive, completive, those that are specifically dedicated to expressing habituality are ordered before and scope the highest of all (Cinque 1999: p. 56; 70). Crucially, this means that the dedicated habitual functional head is syntactically higher than the progressive aspectual head. The complete functional hierarchy is provided below in (13), in which the two functional heads are highlighted with boxes.

(13) $Mood_{speech\ act}$ > $Mood_{evaluative}$ > $Mood_{evidential}$ > $Mood_{epistemic}$ > $T_{(past)}$ > $T_{(Future)}$ > $Mood_{irrealis}$ > $\boxed{Asp_{habitual}}$ > $T_{(Anterior)}$ > $Asp_{perfect}$ > $Asp_{retrospective}$ > $Asp_{durative}$ > $\boxed{Asp_{progressive}}$ > $Asp_{prospective}$/Mod_{root} > $Voice$ > $Asp_{celerative}$ > $Asp_{completive}$ > $Asp_{(semel)repetitive}$ > $Asp_{iterative}$

Cinque (1999: 76)

Our examples in the preceding sections showed that this is the case also in BG and PL. In addition, we showed that the adverbs are optional in the case of SI_{high}, that is, it encodes the habitual reading itself rather than being merely compatible with it. SI_{low}, on the other hand, is compatible with both frequentative and ongoing adverbs just like a null IPFV operator, for example the one proposed

Dorota Klimek-Jankowska & Vesela Simeonova

in Ferreira (2016) selecting for VPs referring to singular or plural events respectively: IPFV [VPsg / VPpl].[7]

In this way the proposal offers a formalization in syntactic terms which captures the differences in the morphological productivity of SI in the two languages, the within-language split (in BG), as well as the fact that the lower SI is equivalent in the two languages.

5 Discussion and conclusion

This study presented a formal description of a systematic difference in the productivity of SI in PL and BG: the Polish perfectives that do not allow subsequent secondary imperfectivization are precisely those cases where Bulgarian SI forms only have a habitual reading and the single ongoing reading is unavailable. These are the perfective forms which in Polish contain purely perfectivizing prefixes. By contrast, Polish perfectives that allow subsequent secondary imperfectivization are those cases where Bulgarian SI forms are ambiguous between a single ongoing and a habitual reading. These are the perfective forms which both in Polish and Bulgarian contain lexical prefixes.

Based on these novel observations we proposed that there are two distinct types of SI in BG: SI_{high} » SI_{low} and only one in Polish, SI_{low}. Crucially, we showed that SI_{low} is uniform in the two languages – it is ambiguous between a single ongoing and habitual reading. SI_{high} is merged in a projection corresponding to Cinque's (1999) Asp_{HAB} and Bulgarian syntax generates this position, while Polish does not.

Previous works on SI cannot capture the novel observations we present here. For example, Rivero & Slavkov (2014) assume that all SIs in BG have both habitual and ongoing readings. This is reflected in their formal account, attributing this duality to context. As we have shown, this is accurate for SI with lexical prefixes in BG and in Polish, but it overgenerates for SI with purely perfectivizing prefixes.

Conversely, Markova (2011) assumes (but provides no evidence) that the SI morphology in BG is in Cinque's (1999) Asp habitual projection. This has the opposite problem: it undergenerates the available ongoing interpretations of SI with lexical prefixes.

One limitation of this study is that the syntactic proposal put forth here still does not explain why it may be that one language is able to generate the higher

---

[7]We assume Tatevosov's (2011, 2015) proposal that aspectual morphology can be lower than the actual aspectual interpretation.

SI layer, while the other language is not. We leave such a comprehensive explanatory account for future work.

Additionally, one may reasonably ask why SI_{high} is not blocked by primary imperfective forms which can also express a habitual reading. In order to address this issue, we show below that in temporal *after*-clauses imposing sequential ordering between two events only SI_{high} is possible and simple imperfective is not, as shown in (14). By contrast in temporal *while*-clauses with two events temporally overlapping only simple imperfective is possible and SI_{high} is not, as shown in (15).

(14) Vseki păt sled kato {*?stroeše / postrojavaše} kăšta, tja se
     every time after when built.IPFV.3SG built.SI.3SG house, she REFL
     srutvaše. BG
     collapsed.SI.3SG
     'Every time 'after' he built (=finished building) a house, it collapsed.'

(15) Vseki păt dokato { stroeše / *postrojavaše} kăšta, imaše
     every time while built.IPFV built.SI.3SG house, have.IPFV.3SG
     incidenti. BG
     incidents
     'Every time while he was building a house, there were incidents.'

One possible answer could be that there are two homophonous SI morphemes in Bulgarian, the one applying higher in syntax being specialized in expressing habituality consisting of a series of temporally non-overlapping bounded events (instead of serving the more general task of "undoing the perfectivizing contribution of the prefix"). Panini's Principle (also referred to as Elsewhere Principle), according to which the application of a specific rule or operation overrides the application of a more general rule, would then link the function of expressing habituality to the specialized SI_{high} form.

What remains to be studied in more detail is the interaction of SI morphology with different classes of superlexical prefixes. It is also necessary to extend the empirical scope to other Slavic languages in order to identify which languages pattern with Bulgarian and which ones pattern parametrically with Polish, and whether there are other possibilities. It also remains to be tested whether the observations we report in this study for past tense contexts can be extended to non-past tense as well.[8]

---

[8]We note that both present tense and imperfective aspect can have a multitude of meanings, e.g. non-actual readings, see Rivero & Slavkov (2014), Nicolova (2017): 364; due to this, a study of the interaction of SI and present tense deserves a longer paper that we leave to future work.

Dorota Klimek-Jankowska & Vesela Simeonova

Finally, Bulgarian and Macedonian have two properties that other Slavic languages lack: they have definite articles and they have preserved the Imperfectum and Aorist tenses. To that end, it would be relevant to test whether the SI morphology interacts with Imperfectum and Aorist and behaves differently in the present and in the past, as well as whether number and referentiality of nominal complements impact the interpretation of either types of SI.

Abbreviations

1	first person	PFV	perfective
3	third person	PP	past participle
AOR	aorist	PRS	present tense
DEF	definite	PST	past
DISTR	distributive	REFL	reflexive
GEN	genitive	SG	singular
IMPF	imperfectum	SI	secondary imperfective
IPFV	imperfective		

Acknowledgments

References

Arsenijević, Boban. 2006. *Inner aspect and telicity: The decompositional and the quantificational nature of eventualities at the syntax-semantics interface* (LOT Dissertation Series 142). Utrecht: LOT.

Babko-Malaya, Olga. 1999. *Zero morphology: A study of aspect, argument structure and case.* Rutgers, NJ: The State University of New Jersey-New Brunswick.

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic inquiry* 16(3). 373–415.

Biskup, Petr. 2012. Slavic prefixes and adjectival participles. In Markéta Ziková & Mojmír Dočekal (eds.), *Slavic languages in formal grammar: Proceedings of FDSL 8.5, Brno 2010*, 271–289. Frankfurt am Main: Peter Lang.

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

Bogusławski, Andrzej. 1963. Czasowniki dokonane i niedokonane i ich stosunki wzajemne. In Witold Doroszewski (ed.), *Z polskich studiów slawistycznych*, 153–159. Warszawa: Polska Akademia Nauk.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective.* Oxford: Oxford University Press. DOI: 10.1093/oso/9780195115260.001.0001.

Di Sciullo, Anna-Maria & Roumyana Slabakova. 2005. Quantification and aspect. In Angeliek van Hout, Henriëtte de Swart & Henk Verkuyl (eds.), *Perspectives on aspect*, 61–80. Dordrecht: Springer. DOI: 10.1007/1-4020-3232-3_4.

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

Ferreira, Marcelo. 2016. The semantic ingredients of imperfectivity in progressives, habituals, and counterfactuals. *Natural Language Semantics* 24. 353–397. DOI: 10.1007/s11050-016-9127-2.

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.

Istratkova, Vyara. 2004. On multiple prefixation in Bulgarian. *Nordlyd, Tromsø Working Papers on Language and Linguistics: Special Issue on Slavic Prefixes* 32(2). 301–321. DOI: 10.7557/12.71.

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. DOI: 10.1353/jsl.2013.0012.

Janda, Laura A. & Tore Nesset. 2010. Taking apart Russian *raz-. The Slavic and East European Journal* 54(3). 476–501. http://www.jstor.org/stable/23345086.

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. DOI: 10.1353/jsl.2022.a923068.

Klimek-Jankowska, Dorota & Joanna Błaszczak. 2023. The interaction of idioms and aspect in Polish. *Natural Language & Linguistic Theory* 42. 153–207. DOI: 10.1007/s11049-023-09577-2.

Kwapiszewski, Arkadiusz. 2022. *Aspect and event structure: The morphosyntax of Polish verbs from a cross-linguistic perspective.* Oxford: University of Oxford. (Doctoral dissertation).

Dorota Klimek-Jankowska & Vesela Simeonova

Łaziński, Marek. 2011. Polish aspectual prefixes, their order and functions: A study based on the National Corpus of Polish. *Word Structure* 4(2). 231–243. DOI: 10.3366/word.2011.0012.

Łaziński, Marek. 2020. *Wykłady o aspekcie polskiego czasownika.* Warszawa: Wydawnictwa Uniwersytetu Warszawskiego. DOI: 10.31338/uw.9788323542346.

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).

Markova, Angelina. 2011. On the nature of Bulgarian prefixes: Ordering and modification in multiple prefixation. *Word Structure* 4(2). 244–271. DOI: 10.3366/word.2011.0013.

Milićević, Nataša. 2004. The lexical and superlexical verbal prefix *iz-* and its role in the stacking of prefixes. *Nordlyd* 32(2). 279–300.

Młynarczyk, Anna Katarzyna. 2004. *Aspectual pairing in Polish* (LOT Dissertation Series 87). Utrecht: LOT.

Nicolova, Ruselina. 2017. *Bulgarian grammar.* Berlin: Frank & Timme.

Ramchand, Gillian. 2004. Time and the event: The semantics of Russian prefixes. *Nordlyd* 32(2). 323–361. DOI: 10.7557/12.72.

Ramchand, Gillian. 2008a. 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. 2008b. *Verb meaning and the lexicon: A first-phase syntax*, vol. 116. Cambridge: University Press Cambridge. DOI: 10.1017/CBO9780511486319.

Rivero, Maria Luisa & Nikolay Slavkov. 2014. Imperfect(ive) variation: The case of Bulgarian. *Lingua* 21. 232–277. DOI: 10.1016/j.lingua.2014.07.019.

Romanova, Eugenia. 2004. Superlexical versus lexical prefixes. *Nordlyd* 32(2). 255–278. DOI: 10.7557/12.69.

Romanova, Eugenia. 2007. *Constructing perfectivity in Russian.* Tromsø: University of Tromsø. (Doctoral dissertation).

Rothstein, Susan. 2020. Perfective is a v operator in slavic, though not in english. *Current development in Slavic linguistics. Twenty years after (based on selected papers from FDSL 11).* 155–172.

Schoorlemmer, Maaike. 1995. *Participial passive and aspect in Russian.* Utrecht: OTS.

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.

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, Leipzig 2013*, 465–494. Frankfurt/M.: Peter Lang.

Wiemer, Björn, Joanna Wrzesień-Kwiatkowska & Marek Łaziński. 2020. Badania aspektu w językach polskim, czeskim i rosyjskim za pomocą korpusów i baz danych (pierwsze podsumowanie tematu). *Forum Lingwistyczne* 7. 45–58. DOI: 10.31261/FL.2020.07.04.

Wiland, Bartosz. 2012. Prefix stacking, syncretism and the syntactic hierarchy. In Markéta Ziková & Mojmír Dočekal (eds.), *Slavic languages in formal grammar: Proceedings of FDSL 8.5, Brno 2010*, 307–324. Frankfurt am Main: Peter Lang.

Willim, Ewa. 2006. *Event, individuation and countability: A study with special reference to English and Polish*. Kraków: Jagiellonian University Press.

Žaucer, Rok. 2012. The syntax of perdurative-prefixed verbs (and the VP-internal/VP-external prefix distinction). In Markéta Ziková & Mojmír Dočekal (eds.), *Slavic languages in formal grammar: Proceedings of FDSL 8.5, Brno 2010*, 339–355. Frankfurt am Main: Peter Lang.

Chapter 11

The (un)expectedly stacked prefixes in Slovenian

Franc Marušič[a], Petra Mišmaš[a] & Rok Žaucer[a]

[a]University of Nova Gorica

When a Slavic verb occurs with multiple prefixes their order is often claimed to follow certain restrictions of a fairly formal character. Firstly, lexical prefixes, which can modify the argument structure of the verb and contribute idiosyncratic interpretations, are always found adjacent to the verbal root, while superlexical prefixes, which do not alter the argument structure and whose interpretative contribution is adverbial, can be stacked over the lexicals. And secondly, when multiple superlexicals stack on a verbal stem, they follow a fixed order. We set out to test these two generalizations with a corpus study. We find that there exist a number of verbs which seem to have more than one lexical prefix, in direct contradiction of the standard assumptions about prefixation.

1 Introduction

In Slovenian and in Slavic languages more generally, simplex verbs consist of a root, a theme vowel [ᴛᴠ] and a tense and agreement ending [ᴛ/ᴀɢʀ], and are typically imperfective (though this is not a rule, cf. e.g. the Slovenian perfective simplex verb *kupiti* 'to buy'). Verbs can also carry one or more prefixes, with the prefixed form generally being interpreted perfectively (unless imperfectivized through, for example, suffixation in the process called secondary imperfectivization [ꜱɪ]). We demonstrate this for the verb *znati* 'to know' and some of its derivatives in Table 1.[1]

Turning to verbal prefixes, these are, in general, all formally related to prepositions (e.g., *ob* 'by/next to', *pri* 'at', etc., cf. Matushansky 2002, Gehrke 2008,

---

[1]Unless indicated otherwise, all examples in this paper are Slovenian.

Marušič, Mišmaš & Žaucer

Table 1: The various parts of the Slavic verb

prefix	prefix	root	SI	TV	T/AGR	Gloss
		zn		a	ti	'to know.IPFV'
	po	zn		a	ti	'to know.PFV'
	po	zn	av	a	ti	'to know.IPFV'
pre	po	zn		a	ti	'to recognize.PFV'
pre	po	zn	av	a	ti	'to recognize.IPFV'

Caha & Ziková 2022, a.o.), but are often assumed to differ among themselves in terms of their position within the verbal domain. Typically, a distinction is made between so-called lexical and superlexical prefixes. The former are often seen as affixal prepositions functioning as VP-internal resultative secondary predicates, similarly to resultative particles in Germanic, the latter as affixal prepositions functioning as VP-external, INFL-level material, e.g., Ramchand (2004), Romanova (2004), Svenonius (2004), and each type is said to behave uniformly with respect to a number of properties. The tree in Figure 1 sketches the relevant positions. A more detailed overview is given in §2.
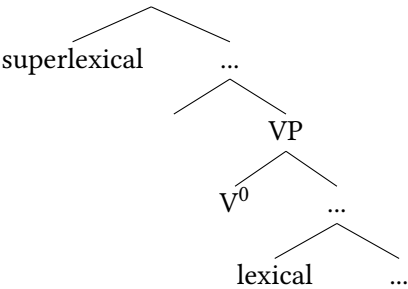

Figure 1: A sketch of the two positions of the two types of prefixes

One important distinction between the two types of prefixes that the literature often seems to convey (even if sometimes unintentionally) is that a verb will – generally – only have one lexical prefix, while superlexical prefixes can stack. The strong tendency that there will only be a single lexical prefix stems from the fact that there is a single position for lexical prefixes, as in Svenonius (2004) or Romanova (2004), or that the semantics of lexical prefixes preclude there being more than one with a given verb, as in Babko-Malaya (2003). When both types of

318

prefixes appear in a verb, the superlexical prefix(es) linearly precede the lexical prefix, and if a verb has multiple superlexical prefixes, these appear in a certain order (e.g., Milićević 2004, Istratkova 2006, Wiland 2012).

The main goal of this paper is to see if we can find a reflection of these generalizations in Slovenian corpus data, if we can use Slovenian corpus data to corroborate these generalizations about the lexical–superlexical division, in particular the view that stacked prefixes will generally not be lexical prefixes and that superlexical prefixes are governed by strict ordering constraints. If we find the generalizations reflected in corpus data, this can be seen as support for the theoretical claims; but note that if we do not, the claims can still be correct, as theoretical possibilities for the existence of specific structures per se do not necessarily imply anything about these structures' frequency in use.

Whereas we find that our corpus data are of limited use for testing fine-grained proposals for orderings of superlexicals, we do also find that they offer corpus support for some aspects of the ordering claims. At the same time, our corpus data also reveal some cases that may appear to be at odds with the expected division. Specifically, while isolated examples of verbs that seem to have two lexical prefixes have been pointed out in the past, e.g., *iz-pod-riniti* 'from-under-drive' and *iz-pod-makniti* 'from-under-move' have been considered in Svenonius (2004: 242), and see also Markova (2011: 260) for Bulgarian and Biskup (2023: 20) for Russian and Czech, our corpus leads us to an expanded set of verbs that display this unexpected combination. Using this set of verbs we then consider how to analyze verbs in which two prefixes both exhibit properties typical of lexical prefixes.

The paper is organized as follows. In §2 we review some widely assumed properties ascribed to the two classes of prefixes. §3 presents a corpus study that focuses on stacked verbal prefixes. §4 discusses the data with potentially unexpectedly stacked prefixes, §5 is the Conclusion.

2 What we know: Lexical and superlexical prefixes in Slavic verbs

A fairly standard division of prefixes that is also characteristic of the more traditional literature (e.g. Vidovič Muha 1993, Toporišič 2000), and is typically assumed to hold for all Slavic languages, establishes two main uses of prefixes. Lexically used prefixes tend to have spatial or idiosyncratic meanings, where "idiosyncratic" is meant to capture situations in which the prefix's addition to the

Marušič, Mišmaš & Žaucer

verb does not lead to a systematically predictable interpretation of the prefix-verb stem complex, as shown in (1). With superlexically used prefixes, on the other hand, the addition of the same prefix predictably adds the same (adverbial) interpretation, and the interpretation of the verb stays transparent and constant across the prefixed verb class, (2).[2]

(1)   ob-delati | ob-soditi | ob-noviti | ob-leteti
      at-work | at-judge | at-new | at-fly
      'to process' | 'to accuse' | 'to renew' | 'to fly around'

(2)   po-sedeti | po-bingljati | po-plesati |
      over-sit | over-dangle | over-dance |
      'to sit for a while' | 'to dangle for a while' | 'to dance for a while' |
      po-igrati se
      over-play REFL
      'to play for a while'

The two classes are said to differ in a number of other properties. Lexical prefixes are said to appear directly on the verb root while superlexicals can be separated from the root by another prefix, and consequently, lexical prefixes can never be stacked, while there should be no such restriction, across the board, for superlexicals. Also, only lexical prefixes are said to be able to affect argument structure. And only lexical prefixes, but not superlexicals, can form secondary imperfectives, cf., e.g., Svenonius (2004: 229) for the diagnostics for superlexical prefixes, though note also that even for Svenonius some subclasses of superlexical prefixes can violate this last constraint (Svenonius 2004: 230). These properties are summarized in Table 2.

     Many aspects of these generalizations, however, have also been questioned. Žaucer (2009), for example, shows that the cumulative prefix *na-* introduces an unselected object – generally considered a hallmark of lexicality – but can, at the same time, also stack over another prefix. A number of authors argued that the split should be in more than two groups: for example, Tatevosov (2008) argues for an independent, third class of *intermediate* prefixes; Babko-Malaya (2003) splits lexical prefixes in two groups; Markova (2011) proposes a four-part division into *outer*, *higher inner*, *lower inner*, and *lexical* prefixes (where the "traditional" lexical prefixes are split into *lower inner* and *lexical* prefixes).

---

[2]For expository reasons, we ignore Slovenian orthography and separate prefixes from the rest of the verb with a hyphen. Prefixes are glossed on the basis of the basic meanings of their prepositional counterparts.

Table 2: Lexical and Superlexical prefixes

LEXICAL PREFIXES	SUPERLEXICAL PREFIXES
adjacent to the root	outside of lexical prefixes
idiosyncratic/PP meanings	adverbial meanings
affect argument structure	don't affect argument structure
form secondary impf.	don't form secondary impf.
generally don't stack	can stack

2.1 Identity of prefixes

What is phonologically one and the same prefix can often be used as either a lexical or a superlexical prefix, as shown in (3)–(4). So if prefixes are defined with their phonological shape one should really only talk of their lexical or superlexical uses, rather than of lexical and superlexical prefixes.

(3)   a. po-liti
          over-pour
          'to spill'

      b. po-sedeti
          over-sit
          'to sit for a while'

(4)   a. do-staviti
          to-put
          'to deliver'

      b. do-od-pirati
          to-off-push
          'to finish opening'

Po- will standardly be analyzed as a lexical prefix resulting in a spatio-idiosyncratic interpretation on the verbal stem in (3a) and as a superlexical prefix with adverbial interpretation in (3b), and *do-* as a lexical prefix added to the verbal stem *staviti* (which never occurs on its own without a prefix in most varieties of Slovenian) and as a superlexical prefix added to an already prefixed stem in (4b).

Moreover, a prefix can have more than one superlexical use, as shown by the Polish example (5), where *po-* serves once as a delimitative and once as a distributive prefix (cf. also Žaucer 2009).

(5) Kucharze po-po-roz-kładali   przez chwilę naczynia i   zajęli
    cooks   po.DELIM-po.DIST-roz-put.SI over all   tables   and began
    się czymś   innym.
    REFL something else
    'The cooks put the dishes on the table for a while and they turned their
    attention to something else.'

                                           (Polish; Klimek-Jankowska & Błaszczak 2022)

2.2 Stacking

As mentioned above, it has been observed that when Slavic verbal prefixes stack
their ordering is not random, but rather reveals certain restrictions of a fairly
formal character. For one, lexical prefixes attach to the verb before superlexi-
cal prefixes, and as a consequence, in any form with multiple prefixes, if the
form includes a lexical prefix, the lexical prefix will appear closest to the verb, as
sketched in (6). The other observation, also sketched in (6), is that superlexical
prefixes (and only superlexical prefixes) can stack even over other superlexical
prefixes so that a single verb can have more than one superlexical prefix but, nor-
mally, just one lexical prefix (cf. Romanova 2004, Svenonius 2004, Gehrke 2008)
(though some authors, e.g. Tatevosov 2008, argue that Russian actually does not
allow stacking of "genuine" superlexical prefixes (i.e., inceptive *za-*, delimitative
po-, cumulative *na-* and distributive *pere-*) but only of "intermediate" prefixes, cf.
above).

(6) superlexical prefix > superlexical prefix > lexical prefix > verb

The restriction to no more than one lexical prefix is taken to reflect the widely
assumed general restriction to one independent resultative secondary predicate
per verb (a.o. Rappaport & Levin 2001, Ramchand 2008), and suggests a further
difference between lexical and superlexical prefixes.[3] Slavic lexical prefixes are
parallel to resultative secondary predicates in languages like English, while su-
perlexicals appear to be something different (cf. also Spencer & Zaretskaya 1998).
   The superlexical prefixes are also said to follow a fixed order when stacked to
the same verbal stem (Istratkova 2006, Wiland 2012, Endo & Wiland 2014, Klimek-
Jankowska & Błaszczak 2022). For example, as claimed by Wiland (2012), who
develops an even more fined-grained, cartography-inspired differentiation of su-
perlexical prefixes, the cumulative prefix *na-* needs to precede the completive

---

[3]This restriction is sometimes also suggested to have a non-structural, conceptual explanation,
e.g. the Single Delimiting Constraint in Tenny (1994) and Filip (2003). In this paper, we focus
on the structural approach.

prefix *do-*, as shown in (7). Istratkova (2006) proposes the order in (8) for Bulgarian, Wiland (2012) proposes the sequence in (9) for Polish, which was later modified by Klimek-Jankowska & Błaszczak (2022) to (10).

(7)   a.   na-do-kładaj   sobie jeszcze
            CUML-COMPL-put self   more
            'get yourself some more (e.g. food)'

      b.   *do-na-kładaj sobie jeszcze   (Polish; Wiland 2012)

(8)   ATT > INCP > COMPL > DIST > CUML > EXC > REP
      po   za   iz   po   na   raz   pre
                                               (Bulgarian; Istratkova 2006)

(9)   DIST > ATT > DELIM > SAT > CUML > EXC > REP > PERD > COMPL > TERM
      po   pod po   na   na   na   prze prze   do   od
                                               (Polish; Wiland 2012)

(10) DELIM > DIST > SAT > CUML > {PERD, EXC, REP, ATT, TERM, PURELY PFV}
      po   po   na   na   prze   prze prze pod od   s/na
                        (Polish; Klimek-Jankowska & Błaszczak 2022)

3 Corpus-study results

In order to get better empirical insight into multiply prefixed verbs in Slovenian, we considered two sets of data. First, we looked at the 3000 most common verbs in Slovenian using the *WeSoSlav* database (see Arsenijević et al. in preparation), to explore the behavior of common verbs with more than one prefix in general (assuming that such a 3000-verb sample is representative of the language). In the second step we created a list of multiply prefixed verbs from the list of all verbs occurring in the *Gigafida 2.0* reference corpus of written standard Slovenian (Čibej et al. 2019).

Starting with *WeSoSlav*, while we were able to confirm that multiple prefixation exists, we found that only 6 out of 3000 verbs had 3 prefixes (no verbs have more), 178 verbs had 2 prefixes, while 2,076 had a single prefix.[4] Table 3 gives the

---

[4]The 6 verbs with three prefixes include two aspectual pairs (i.e. *s-po-raz-umeti* 'to agree/communicate.PFV', *s-po-raz-umevati* 'to agree/communicate.IPFV' and *s-po-pri-jeti* 'to cope/deal with.PFV' *s-po-pri-jemati* 'cope/deal with.IPFV') so that there are really only 4 different verbs with three prefixes. Applying this same aspectual-pair exclusion criterium also to verbs with two and with one prefix, there are only around 125 different verbs with two prefixes and around 1500 different verbs with a single prefix.

relevant results.[5] Note that each verb was counted only once (that is, verbs with three prefixes were not counted also as verbs with one prefix and as verbs with two prefixes).

Table 3: Prefixation in WeSoSlav (Arsenijević et al. in preparation)

NUMBER OF PREFIXES	NUMBER OF VERBS	PERCENT
zero	740	24.67%
(exactly) one	2,076	69.2%
(exactly) two	178	5.93%
(exactly) three	6	0.2%
TOTAL	3,000	

This data leads us to certain conclusions. On the one hand, prefixed verbs are more common than verbs without prefixes (the latter are not necessarily simplex, since some have a suffix, e.g. *kup-ova-ti* 'to buy.ɪᴘꜰᴠ'). But more importantly, while verbs with a single prefix are extremely common, multiple prefixation is not. Given the relatively low number of multiply prefixed verbs, no proper quantificational analysis of the relative order of prefixes can be conducted. In order to create a better empirical base for investigating multiple prefixation, we created a larger list of multiply prefixed verbs.

3.1 Corpus-study results, additional data

The new set of data was created from the list of all 90,000+ verbs found in the *Gigafida 2.0* corpus (Čibej et al. 2019). We only looked at verbs that had more than 5 occurrences in the corpus as the number of typos, misspelled words and incorrectly classified words only increases with less frequent strings of characters. Prefixed verbs were automatically extracted from the list using a simple formula that looked at each individual verb and checked whether it begins with one of the prefixes. The prefix was subtracted from the verb and the verb was checked again if the remaining part of the verb starts with one of the listed prefixes. This procedure was repeated five times. The automatically extracted multiply prefixed verbs were then also checked manually, since in some cases the automatic procedure counted some beginnings of stems/roots as prefixes, as in

---

[5]Verbs that have a non-Slavic prefix like *re-* in *re-organizirati* 'to reorganize' or *dis-* in *dis-kvalificirati* 'to disqualify' were counted as unprefixed. Similarly we also disregarded the negative prefix *ne-*, as in *o-ne-sposobiti* 'to disable'.

the case of verbs like *stati* (incorrectly analyzed as *s-tati*) 'to stand' or *vleči* 'to pull' (incorrectly analyzed as *v-leči*), and some combinations of prefixes could be misparsed as combinations of different prefixes, e.g. *pod-o-...* 'under-about-...', which is string-homophonous with *po-do-...* 'over-to-...', etc.

Table 4: Prefixation in the expanded database

NUMBER OF PREFIXES	NUMBER OF VERBS	PERCENT
zero	4,186	29.45%
one	9,181	64.58%
two	833	5.86%
three	16	0.11%
TOTAL	14,216	

With this procedure we were able to retrieve a list of 849 multiply-prefixed verbs that exhibit at least 5 occurrences in the corpus. As above, the list contains some aspectual pairs, see footnote 4, but we did not exclude aspectual pairs for the figures we made. Verbs with three prefixes are extremely rare in Slovenian (see §3.2), and among the verbs with at least 5 occurrences in the corpus, there were no verbs with more than three prefixes.

In Figure 2 the prefixes are ordered on the basis of their likelihood, increasing from left to right, to appear as the prefix closest to the verb. The first thing to note is that no prefix is restricted to the root-adjacent position: in the presented set of verbs they all appear in the first position of a pair of prefixes at least once.

This last observation is very clearly visible also from Figure 3. Even the prefixes *pod-* 'under-' and *vz-* 'up-', which can be, based on Šekli (2016), taken as essentially exclusively lexical prefixes in Slovenian, appear stacked over another prefix in up to 20% of the cases. Actually, even the prefixes which seem to be most common in the root-adjacent position (*vz-* 'up-', *v-* 'in-', *ob-* 'around-', *pod-* 'under-' according to Figure 2 and Figure 3) also appear stacked over another prefix in at least 10% of the cases. Thus, all prefixes that are possible in the root-adjacent position can also be used as stacked prefixes (cf. Łaziński 2011 for a similar dictionary-based result from Polish) and thus – according to the description so far – as superlexical prefixes. The implication does not go both ways, as *so-* 'co-' is never used as verb-adjacent in multiply prefixed verbs.

The table in Table 5 confirms a tendency for a hierarchy, but does not confirm a true hierarchy. Most pairs of two prefixes only exist in one order of the two prefixes, as is evident from the fact that the lower left half of the table in Figure 5
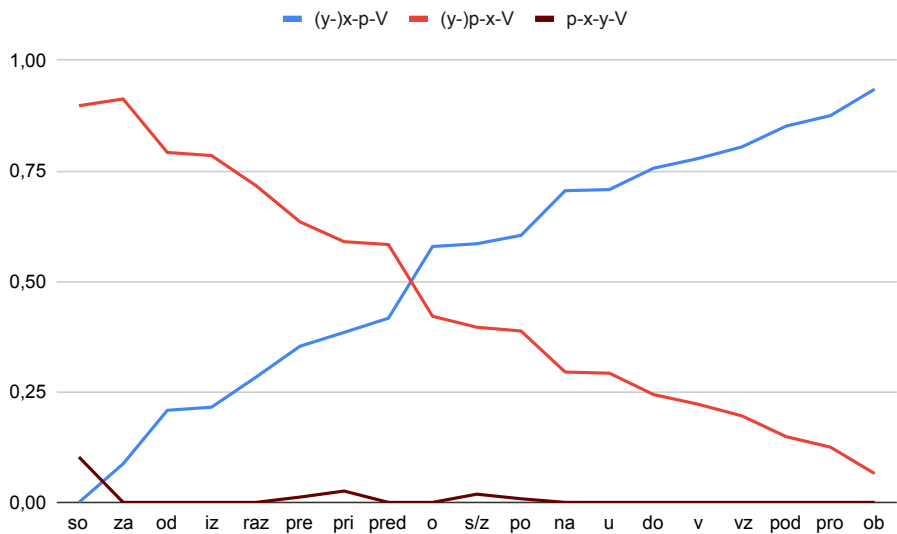
Marušič, Mišmaš & Žaucer

Figure 2: The frequency of prefixes relative to their position in a multiply prefixed verb (counting tokens of combinations)
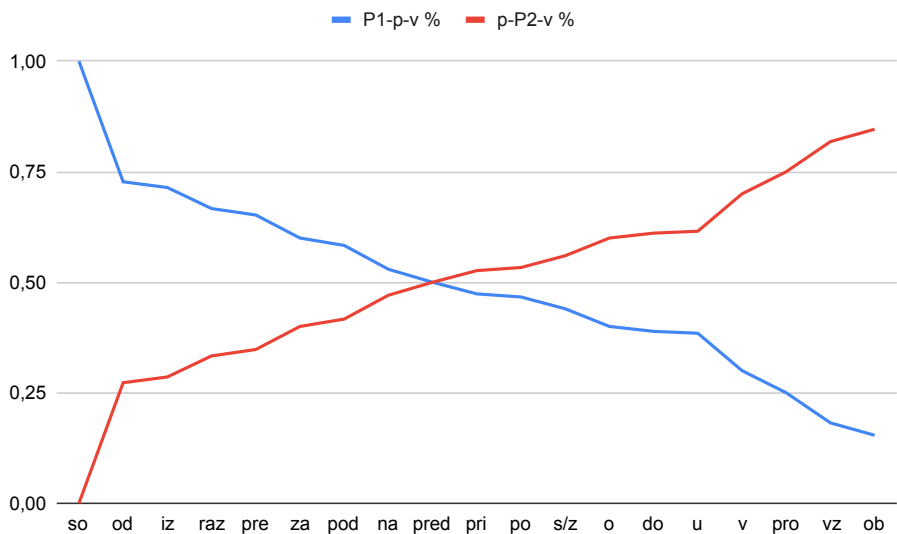

Figure 3: Relative amount of prefixes that a prefix can appear with either when it comes first or second in a pair of prefixes (counting types of combinations).

Table 5: The cross-table of prefix combinations. The first prefix of a pair is listed vertically, the second horizontally. The order of prefixes is a slightly modified sequence from Figure 2.

	so	od	pred	raz	iz	za	pre	o	pri	s/z	po	na	u	v	do	vz	pod	pro	ob
so		3	2			1		3	1	1	4		12	3	4		2	1	2
od						1		1	2	9	1				1	2		2	
pred							1	1			2	3							
raz							2	3		5	15			2			1	7	3
iz							6	4		4	11	4	1		1	2	14		4
za		1					10	9	5	127	20		7	3	8	4		3	12
pre		1			6	5	4		8	5	6	10	27	14	3	1	4		2	12
o						1	4		2	4	11			2					
pri						5		1		14	11	1			5	7		1	3
s/z		1				2	36	5	1		35	2			1	41		8	2
po		1		9	5	5		2	6	18		14	10	8	2	13		2	2
na						1	1			1	5		1	7	4	2			1
u								2	5	9								2	1
v								2	5						1				
do				2	1	1		2	1	2									1
vz											4				5				
pod		1	1			2				1	2	2				2			
pro				2											1	1			
ob				2							1								

has fewer cells filled in than the upper right half of the table. Figure 5 shows that many pairs of prefixes exist with both orders of prefixes, so for example, there are 10 different verbs with the sequence *za-pre-* 'for-over', and 4 different verbs with this sequence reversed, i.e. *pre-za-*. Given that certain prefixes have more than one use, that is, that they can be either used as lexical or superlexical prefixes, one would need to determine case by case whether the second prefix of a sequence of two prefixes is indeed an instance of a superlexical prefix or a lexical prefix (which means coding your data on the basis of previous qualitative data analysis, which we wanted to avoid here as much as possible). Further, some prefixes have even more than one superlexical use (cf. Wiland 2012, Klimek-Jankowska & Błaszczak 2022), so that they can appear in more than one position within the proposed hierarchy of superlexical prefixes. These two facts presumably explain why we find so many different combinations where both orders of the two prefixes are possible, and we can only conclude that automatic extraction of prefixes cannot produce a clear sequence of superlexical prefixes, and therefore none of the proposed orders can be either confirmed or rejected.

If we assume that, generally, only the prefix closest to the root will potentially be a lexical prefix (see §1), we would need to look at verbs with at least three prefixes to be able to get a sequence of superlexical prefixes, but we only have 16 verbs with three prefixes to work with.

Marušič, Mišmaš & Žaucer

3.2 Verbs with three prefixes

Given that prefixes should be able to stack, and that quite some claims have been made on the basis of the possible and impossible ordering patterns in stacking, we expected that we will find substantial numbers of verbs with three or more prefixes. However, this prediction was not confirmed since out of 849 multiply-prefixed verbs no verb included more than three prefixes and only 16 included three prefixes. Specifically, a closer review of the 16 verbs showed that this number is actually even smaller, as "deduplication" of aspectual pairs reduces the number to a mere 10 verbs, listed in (11)–(20).[6]

Moreover, even some of the 10 verbs in (11)–(20) are odd-looking and unknown to us, such as *priopoteči* in (17), but as these verbs' few occurrences in the corpus seem to exhibit similar uses, we did not exclude them manually.[7,8]

(11) pre-raz-po-rediti[PFV] | pre-raz-po-rejati[IPFV]
      over-from-over-order | over-from-over-order
      'to rearrange'

(12) s-po-pri-jeti[PFV] | s-po-pri-jemati[IPFV]
      with-over-at-hold | with-over-at-hold
      'to tackle'

(13) s-pre-ob-rniti[PFV] | s-pre-ob-račati[IPFV]
      with-over-around-turn | with-over-around-turn
      'to convert'

(14) so-u-po-rabiti[PFV] | so-u-po-rabljati[IPFV]
      co-in-over-use | co-in-over-use
      'to co-use'

---

[6]Why are verbs with three or more prefixes so rare in actual language is a question we leave for future work. In discussing the rarity of some predicted orders of superlexical stacking, Markova (2011: 269) suggests that this might have to do with processing constraints.

[7]The verb *prisprehoditi* has 5 occurrences in *Gigafida 2.0* and *priopoteči* has 6 occurrences, and these 5/6 occurrences even include more than one example by the same author, so these are possibly forms that have been used/coined by two or three speakers. *Posprehoditi* has 27 occurences and *porazporediti* 30 occurrences in *Gigafida 2.0*. With the exception of *posprehoditi*, none of these are listed in any of the dictionaries available to us; the translations we provide for these verbs are thus our context- and form-based inferences.

[8]One could perhaps also exclude verbs with the prefix *so-* (similar to the English *co-*), such as (14) and (15). This prefix behaves differently from other verbal prefixes in several respects, can also appear in non-verbal contexts, e.g. *so-avtor* 'co-author', and is consequently often not even included in works on verbal prefixation, e.g. Vidovič Muha (1993).

(15) so-po-vz-ročiti[PFV] | so-po-vz-ročati[IPFV]
      co-over-up-hand | co-over-up-hand
      'to co-cause'

(16) s-po-raz-umeti[PFV] | s-po-raz-umevati[IPFV]
      with-over-from-understand | with-over-from-understand
      'to agree/communicate'

(17) pri-o-po-teči
      at-around-over-run
      'to get somewhere staggering'

(18) po-raz-po-rediti
      over-from-over-order
      'to distribute'

(19) po-s-pre-hoditi
      over-with-over-walk
      'to take a brief walk'

(20) pri-s-pre-hoditi
      at-with-over-walk
      'to get somewhere taking a walk'

Ignoring the interpretation of individual prefixes, we can extract several partial orders of prefixes from the above examples. Partial orders are given in (21). Interestingly *s/z-* 'with-' and *po-* 'over' appear in both orders, which is not surprising if both *po-* and *s/z-* have more than one superlexical use and thus more than one position in the hierarchy of superlexical prefixes.

(21) so > u, po
      s/z > po, pre > raz
      pri > o, s/z
      po > raz, s/z

But what seems to be going on is probably something else. A closer look at the verbs in (11)–(20) reveals that actually none of them seems to have a sequence of two obvious superlexical prefixes, and that for some of them no prefix seems very much like a standard, VP-external-looking superlexical prefix. In verbs like (12) and (13) all three prefixes have some of the properties of lexical prefixes – they affect the argument structure or have spatial PP meanings.

Automatic extraction of prefixes out of a list of verbs has limitations, and even though we were able to show that there is a tendency for a hierarchy we did not arrive at a single order of superlexical prefixes; we were just able to show that there are certain prefixes that prefer to stay closer to the root and others that prefer to be further away, and that this preference is different for different prefixes, but different methods of establishing this preference gave different sequences of prefixes.

We will devote the remainder of this paper to the observed unexpected sequences of prefixes. As mentioned above, even the prefixes that have been suggested as being exclusively lexical appear in up to 20% of cases as the first prefix in a sequence of prefixes. Consider the verb *vz-po-staviti* 'to set up'. The prefix *vz-* generally has the meaning 'up' and is rarely associated with an adverbial meaning (e.g. *vz-ljubiti* 'to start to love') that we generally expect with the outermost prefix of a verb with two prefixes – certainly such a meaning is absent in *vz-po-staviti*. Similarly, the inner prefix of *vz-po-staviti*, as expected, has a meaning that can only be associated with a lexical prefix ('over'). This type of verbs – which we will call *vz-po-staviti*-type verbs – is what we turn to in §4.

4 Examples with two seemingly lexical prefixes

Considering the mainstream view in the literature on prefixation (§1), one expectation is that if a verb has two (or more) prefixes, at most one will tend to be a VP-internal, lexical prefix, while the rest will tend to be superlexical (or intermediate). However, our corpus study presented in §3 turned up a sizeable number of multiply prefixed verbs in which the outermost prefix also contributes a typically lexical meaning (i.e., *vz-po-staviti*-type verbs). Examples (22) to (28) give a sample of such verbs. These examples are presented here in triplets: The first form is the unprefixed version, the last is the relevant example with two prefixes, and the middle example is the form (which is always an attested form) with a single prefix. Example (28) stands out somewhat, as it has three seemingly lexical prefixes, and the version of the verb with just two is not attested in modern Slovenian (though it is attested in older versions of Slovenian). We use a [#]hashtag to mark unprefixed forms that are unattested in modern standard Slovenian and also in many dialects, such as [#]*staviti*, though they are attested in some present-day dialects of Slovenian, in closely related BCMS, or are historically attested. Note also that in *vz-po-staviti*-type verbs, the verb with a single prefix always seems to exist, which makes these different from Žaucer's (2002) examples like (30), discussed in Svenonius (2004), in which the version with a single prefix is

not attested.[9]

(22) klicati | po-klicati | v-po-klicati
      call | over-call | in-over-call
      'to call' | 'to call up' | 'to enlist'

(23) #staviti | po-staviti | vz-po-staviti
       set | on-set | up-on-set
               | 'to set' | 'to set up/establish'

(24) #jeti | pri-jeti | o-pri-jeti
       grab | at-grab | around-at-grab
             | 'to grab' | 'to hold on to'

(25) #peti | vz-peti | po-vz-peti
       pull | up-pull | on-up-pull
            | 'to climb' | 'to climb'

(26) #deti | o-deti | raz-o-deti
       put | around-put | from-around-put
            | 'to wrap' | 'to reveal'

(27) nesti | za-nesti | pri-za-nesti
      carry | behind-carry | at-behind-carry
      'to carry' | 'to carry in' | 'to spare'

(28) #umeti | raz-umeti | #po-raz-umeti | s-po-raz-umeti
       get/understand | apart-get | over-apart-get | with-over-apart-get
                      | 'to understand' | | 'to agree'

(29) nesti | pri-nesti | do-pri-nesti
      carry | at-carry | to-at-carry
      'to carry' | 'to bring' | 'to contribute'

(30) riniti | *pod-riniti | iz-pod-riniti
      push | under-push | from-under-push
      'to push' | | 'to push out'

---

[9]Regarding (24): some varieties do exhibit a verb *jeti*, but only with an aspectual meaning 'to start'. While this is the same root, with the aspectual meaning having developed from the root's basic meaning 'grab'/'hold' or 'take' (Snoj 2009), it is not the root's meaning that the prefixed verb is based on, so we mark *jeti* in (24) with a hashtag.

The meaning contribution of the outermost prefix suggests that these examples contain more than one lexical prefix. In (22) and (23) the addition of *v-* and *vz-*, respectively, leads to an idiosyncratic, or perhaps spatial meaning; in (24) the prefix *o-* adds a spatial meaning; in (25) the contribution of *po-* is not very clear (little discernible meaning change compared to its singly prefixed input); in (27), *pri-* adds an idiosyncratic meaning; etc. This situation is suprising in view of the idea that lexical prefixes generally do not stack.

The question is, then, how these prefixes should be analyzed. Possible answers include: (i) they are, despite their meanings, VP-external superlexicals; (ii) they fall into one of the additional categories of prefixes described in the literature (cf. Babko-Malaya 2003, Tatevosov 2008, Markova 2011, etc.); (iii) they are indeed VP-internal lexicals, but can be stacked because some special conditions are met. The last option then further opens several possibilities that could be explored, such as the possibility that these examples, in a sense, only include one prefix (and the inner prefix is somehow incorporated into the root), or that these are in fact two prefixes which either appear in a double-VP structure with two independent ResultPhrase positions for lexical prefixes, that they are result modifiers, or that they even require a completely different approach, perhaps one in which all prefixation is introduced above the VP (cf. Biskup 2023). In what follows, we explore these options.

4.1 Option 1: They are superlexical

If the outer prefixes of the *vz-po-staviti*-type verbs were instances of *v*P-external, superlexical prefixes, then one would expect them to exhibit properties typical of superlexical prefixes. One such property is their placement and the ability to stack – since they appear on top of a prefix they could, in principle, be taken as superlexical.

However, there are arguments against this claim. Firstly, they do not carry typical superlexical, adverbial meanings. If we consider the verb *pri-za-nesti* in (27), adding the prefix *pri-* results in an idiosyncratic meaning shift from 'to carry in' to 'to spare', which cannot be the result of one of the two possible adverbial readings that *pri-* has, according to Šekli (2016), namely, a delimitative or an inchoative reading, as in *pri-preti* 'open a little' and *pri-žgati* 'to light up', respectively.

Also, superlexical prefixes are typically said not to allow secondary imperfectivization (see §2). Except for *vpoklicati*[PFV] 'to conscript' in (22), all other verbs given in (22)–(30) have well-attested secondary imperfectives: *vzpostavljati*[IPFV]

'to establish', *oprijemati*[IPFV] 'to hold on to', *povzpenjati*[IPFV] 'to climb', *razode-vati*[IPFV] 'to reveal', *prizanašati*[IPFV] 'to spare', *sporazumevati*[IPFV] 'to communi-cate', *spozabljati se*[IPFV] 'to forget oneself'.[10] It should be emphasized that these do not seem to be cases of a prefix combining with an imperfective base – if this were the case, the resulting verb should be, contrary to fact, perfective. Rather, the imperfectivized verbs match the meaning of the perfective form (except in aspect), suggesting that these are in fact imperfectivizations of the doubly prefixed verbs:

(31) a. Veter je {za-nesel[PFV] / za-našal[IPFV]} listje na dvorišče.
wind AUX behind-carry behind-carry leaves..ACC on yard
'The wind carried leaves to the yard.'

b. * Veter je {pri-za-nesel[PFV] / pri-za-našal[IPFV]} listje na
wind AUX at-behind-carry at-behind-carry leaves..ACC on
dvorišče.
yard
Literally: 'The wind spared leaves to the yard.'

(32) a. * Sodišče ni { za-neslo[PFV] / za-našalo[IPFV]} osumljencem.
court NEG.AUX behind-carry behind-carry suspects.DAT
Literally: 'The court didn't carry to the suspects.'

b. Sodišče ni {pri-za-neslo[PFV] / pri-za-našalo[IPFV]} kriminalcem.
court NEG.AUX at-behind-carry at-behind-carry criminals
'The court didn't spare the criminals.'

And finally, according to Svenonius (2004) superlexical prefixes normally do not appear in nominalizations, in particular root/zero nominalizations (cf. also Caha & Ziková 2016). While it should be noted that not all verbs in Slovenian derive root nominalizations, several of these *vz-po-staviti*-type verbs do:

(33) iz-po-staviti | iz-po-stav-a / iz-po-stav-e
out-over-stand | out-over-stand-F.SG.NOM out-over-stand-F.SG.GEN
'to single out' | 'branch'

---

[10]The fact that at least for many speakers, *vpoklicati*[PFV] 'to conscript' does not have a natural imperfective counterpart is not problematic, given that it is also not the case that every per-fective verb with a single prefix has a secondary imperfective counterpart, e.g., *za-bresti*[PFV] 'to get stuck' does not. In fact, the input of *vpoklicati*[PFV], i.e., *poklicati*[PFV] 'to call up', also does not have a secondary imperfective counterpart.

Marušič, Mišmaš & Žaucer

(34) do-pri-nesti | do-pri-nos-∅   / do-pri-nos-a
      to-at-carry | to-at-carry-M.SG.NOM to-at-carry-M.SG.GEN
      'to contribute' | 'contribution'

(35) za-pri-seči | za-pri-seg-a   / za-pri-seg-e
      behind-at-reach | behind-at-reach-F.SG.NOM behind-at-reach-F.SG.GEN
      'to pledge' | 'pledge'

(36) v-po-klicati | v-po-klic-∅   / v-po-klic-a
      in-over-call | in-over-call-M.SG.NOM in-over-call-M.SG.GEN
      'to call in, enlist' | 'conscription'

Root nominalizations are usually assumed not to contain structure above the VP, and following Svenonius (2004), the existence of root nominalizations can be taken as an argument that these prefixes are structurally similar to lexical prefixes, merged inside the verb phrase.

The only reason to consider the outermost prefix in the verbs under discussion to be superlexical, then, would be their placement, whereas their other properties speak against their being superlexical. In what follows, we will therefore further explore the option that they are not superlexical.

4.2 Option 2: They are neither lexical nor superlexical

Whereas a binary split into VP-internal lexical prefixes and a possibly internally diverse group of superlexical prefixes is the most common stance taken in the literature (present also in several cartography-like accounts such as Wiland 2012), some authors have proposed systems with more than two circumscribed groups of prefixes. In this section, we consider whether the stacked prefix in our *vz-po-staviti*-type verbs could belong to one of these additional classes, and conclude that it could not. Note that we will always leave the highest-merging prefix type of these systems out of the discussion: that the stacked prefix in our *vz-po-staviti*-type verbs cannot be any of these highest merging types follows from the discussion in §4.1.

Tatevosov (2008) analyzes lexical prefixes as merging in a result phrase inside the VP and superlexical prefixes as merging outside the vP. He suggests that between the lexical and the superlexical prefixes there is a third group – intermediate prefixes, such as the Russian completive *do-* – which merges somewhere above the VP and below the superlexicals.

While *vz-po-staviti*-type verbs share certain properties with verbs with intermediate prefixes (e.g. being able to be imperfectivized), they also have characteristics that set them apart. According to Tatevosov (2008), intermediate prefixes

(among other characteristics) yield compositional meanings and never influence argument structure. As we already saw in §4.1, the outermost prefix in *vz-po-staviti*-type verbs can lead to non-compositional meanings, such as *vz-* in (23) ('to set' > 'to establish') or *pri-* in (27) ('to carry in' > 'to spare'), which come with concomitant argument structure effects (shown with more detail in §4.3). As was also already mentioned in §4.1, *vz-po-staviti*-type verbs often serve as the basis for root nominalizations, as in (33), which following Svenonius (2004) also suggests that their prefixes do not originate above the VP. We therefore conclude that our *vz-po-staviti*-type verbs are not simply intermediate prefixes.[11]

In a similar vein, Markova (2011) presents an account in which lexical prefixes, which she merges inside the VP as head adjuncts to V^0, are joined by three groups: outer prefixes, which are above vP; higher inner prefixes, which originate between VP and *v*P; and lower inner prefixes, which originate in a PathP complement to V^0.

Given that Markova's (2011) higher inner prefixes are positionally the same as Tatevosov's (2008) intermediate prefixes, the same arguments that we just presented against viewing the stacked prefix in *vz-po-staviti*-type verbs as Tatevosov's intermediate prefixes will also apply to the possibility that these prefixes would be Markova's higher inner prefixes. At the same time, the stacked prefixes in *vz-po-staviti*-type verbs will also not be Markova's lower inner prefixes, since she reserves this position for spatial and causative prefixes, whereas the stacked prefixes in a number of our *vz-po-staviti*-type verbs are neither spatial nor causative: see again, for example, (27). Also, Markova's lower inner prefixes cannot contribute idiosyncratic meanings, which she reserves for lexical prefixes, but the stacked prefixes in our *vz-po-staviti*-type verbs can contribute idiosyncratic meanings.

Note, however, that somewhat in passing, Markova (2011: 260) also mentions the possibility that a verb hosts two lexical prefixes, in a V^0 combining two prefixes and a verb, that is, in a V^0 to which two prefix heads have been adjoined. From what we can tell, this structure, which assumes the possibility for idiosyncratic meanings for both prefixes, can actually successfully derive our *vz-po-staviti*-type verbs. Though Markova does not mention this, her account probably also predicts the possibility that a verb hosts a lexical prefix as well as a stacked lower inner prefix, a structure that presumably can derive some of our *vz-po-staviti*-type verbs. We return to this in §4.3.1.

---

[11]In the spirit of Žaucer (2013), an argument could also be made on the basis of relative scope with respect to VP adverbials, the restitutive 'again' and adverbs of completion, all of which scope over the outer prefix. For a demonstration of some of this, see §4.3.3 below.

Marušič, Mišmaš & Žaucer

Another account that proposes more than two groups of prefixes was put forth in Babko-Malaya (2003). As a version of the superlexical category, Babko-Malaya has Aktionsart-prefixed verbs, in which the prefix merges outside the VP (for which see §4.1). In addition, she has lexically prefixed verbs, in which the prefix is adjoined to V^0, and resultatively prefixed verbs, in which the prefix (itself part of a complex head) is adjoined to V^0. As explained by Babko-Malaya (2003: 27) herself, the semantics derived from those structures is such that double prefixation is only possible when a stacked prefix is an Aktionsart prefix (i.e., a superlexical prefix in the terminology from §4.1), while it actually prevents double prefixation with either two lexical prefixes, two resultative prefixes, or a combination of the two. So the stacked prefixes in our *vz-po-staviti*-type verbs will clearly not be either the lexical or the resultative prefixes of Babko-Malaya (2003).

Note, however, that as pointed out by a reviewer, the account from Babko-Malaya (2003) is presumably not incompatible with the existence of stacked prefixes of the type of *vz-po-staviti* if such stacked prefixes are analyzed as result modifiers in the sense of Žaucer (2013) (even though Babko-Malaya herself does not discuss this type of data). This would be a version of the view that these stacked prefixes are VP-internal, lexical prefixes, which is the option we discuss next, having determined now that our *vz-po-staviti*-type verbs can be neither superlexical nor intermediate, or something of the sort.

4.3 Option 3: They are lexical

If prefixes in *vz-po-staviti*-type verbs are VP-internal lexical prefixes, we expect them to exhibit properties typically ascribed to lexical prefixes. Again, an argument against such an analysis is that the prefixes under discussion stack, while for lexical prefixes it is assumed that they generally do not stack, see §1 and §2. The explanation for this restriction is structural. Because lexical prefixes are generally assumed to be resultative and originate in a VP-internal Result Phrase [RP], as shown in Figure 4 (based on Svenonius 2004: (80)), and because verbal structure is assumed to be able to host only one result/one RP (Rappaport & Levin 2001, Ramchand 2008), it should normally not be possible to have more than one lexical prefix per verb.[12]

However, as already indicated in §4.1, these prefixes display several other properties that can be taken as arguments for a VP-internal, lexical analysis. In addition to non-superlexical interpretations, the availability of secondary imperfec-

---

[12]Though see den Dikken (1995) for a different understanding of the structure used for particles and prefixes and the restrictions it imposes.
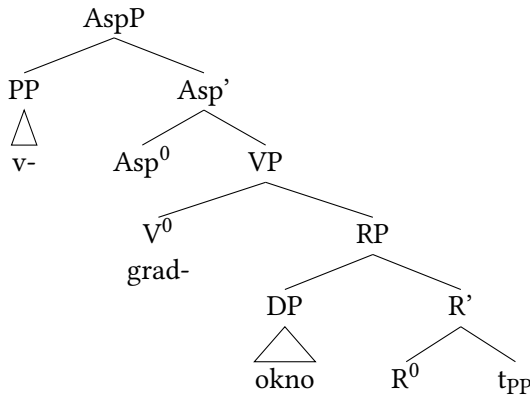
Figure 4: Structure dictates the one-lexical-prefix restriction.

tivization and root nominalizations, the outer prefixes in *vz-po-staviti*-type verbs also exhibit some argument-structure effects.

For example, the "singly" prefixed verb *pri-jeti* 'to grab' can select for a single accusative object, the reflexive clitic sᴇ, and an optional genitive object, or an optional reflexive clitic sᴇ and a prepositional phrase, as shown in (37a). The "doubly" prefixed *o-pri-jeti* 'to hold on to', on the other hand, is unacceptable (in most modern varieties) with a single accusative object, requires the genitive object with a reflexive clitic sᴇ, and simply does not tolerate prepositional objects, as shown in (37b). Differences in the argument structure of the singly- and doubly-prefixed counterparts are observed also in other cases, as shown in (38)–(39).

(37)   a. pri-jeti {ročaj   / se   (ročaja)   / (se)   za ročaj}
           at-grab handle.ᴀᴄᴄ ʀᴇꜰʟ handle.ɢᴇɴ   ʀᴇꜰʟ for handle.ᴀᴄᴄ
           'to grab the handle/ to grab on (to the handle)'

       b. o-pri-jeti   {*ročaj   / se   *(ročaja)   / *(*se) za
           around-at-grab handle.ᴀᴄᴄ ʀᴇꜰʟ handle.ɢᴇɴ   ʀᴇꜰʟ for
           ročaj}
           handle.ᴀᴄᴄ
           'to grab on (to the handle)'

(38)   a. pri-seči (*pričo)
           at-reach   witness.ᴀᴄᴄ
           'to swear, take an oath'

    b. za-pri-seči   (pričo)
       behind-at-reach witness.ACC
       'to take an oath; to swear in a witness'

(39) a. za-nesti   {skrbi   Vidu   / *Vidu   (s   skrbmi)}
       behind-carry worries.ACC Vid.DAT   Vid.DAT with worries
       'to carry worries to Vid'

    b. pri-za-nesti   {(*skrbi)   Vidu   / Vidu   (s   skrbmi)}
       at-behind-carry   worries.ACC Vid.DAT   Vid.DAT with worries
       'to spare Vid the worries'

Given that we seem to be led to the conclusion that the outer prefix in *vz-po-staviti*-type verbs is a lexical prefix, it should be noted that different authors have previously observed that VP-internal prefixes are not a homogeneous group. A natural question to ask, then, is whether the outer prefixes in *vz-po-staviti*-type verbs share any of the properties of those proposed subgroups.

4.3.1 Option 3.1: They are lexical – but these verbs contain only one prefix

This option presents itself as a possibility especially in view of the fact that some of these apparently doubly-prefixed verbs are no longer used without a prefix. For example, while (40) exists in some Slovenian dialects (and in BCMS), it does not exist in standard Slovenian, nor in many other dialects that normally use *vz-po-staviti*. Similarly, (41) does not exist in modern Slovenian (though it does exist in BCMS), and neither does (42).

(40)   #*staviti* 'set' (exists in some Western Slovenian dialects)

(41)   #*peti* 'pull' (but exists in BCMS)

(42)   **jeti* 'grab'/'hold'

Given that these simplex forms are not attested (or are at best very limited) synchronically, it could be the case that the innermost prefix, even if historically a prefix, is just a part of the root (cf. Fowler 1996), or in other terms, as suggested in Markova (2011: 260) for all prefixes resulting in idiosyncratic meaning shifts, is adjoined to V^0, forming a complex verbal head. According to this analysis, a verb can have more than one lexical/X^0-adjoined prefix, and since prefixes are adjoined to v^0, they are freely ordered.

    On the one hand, it seems to us that Markova's proposal could be seen as consistent with *vz-po-staviti*-type verbs, especially for those built on verbs like

po-staviti 'to set' or *pri-jeti* 'to grab', whose unprefixed bases are not attested syncronically, as well as for those whose outer prefix seems somehow related to a spatial use, such as in *v-po-klicati* 'to enlist'. On the other hand, for a number of *vz-po-staviti*-type verbs aspect presents an issue. Several of these verbs, such as *vz-peti* 'to climb.PFV', are based on stems that were historically imperfective, and just like most lexically prefixed verbs (and unlike most native unprefixed verbs), these verbs generally form secondary imperfectives, e.g. *po-stavljati* 'to stand.IPFV', *vz-penjati* 'to climb.IPFV', *pri-jemati* 'to hold.IPFV'. This suggests that these inner prefixes trigger perfectivity. It is unclear to us how such adjunction could account for the change of aspect. In Svenonius's (2004) account, for example, the perfectivizing effect arises when a prefix moves from the RP into a VP-external aspect projection; if the prefix is part of a complex V^0, such movement does not seem to be possible. For those *vz-po-staviti*-type verbs which exhibit singly-prefixed counterparts even in modern Slovenian, such as *v-po-klicati* 'to enlist' or *za-pri-seči* 'to take an oath, to swear somebody in', this aspectual concern regarding treating their inner prefix as V^0-adjoined is even more obvious.

In addition, whereas some of these *vz-po-staviti*-type verbs synchronically do not exhibit unprefixed versions, they do occur in a modern Slovenian with several different prefixes, (43)–(45), resulting in forms with either clearly related or with idiosyncratic meanings. We can take this as an argument against an analysis on which the innermost prefixes are simply part of the root: While we agree with Romanova (2004), who considers similar examples of "cranberry roots" in Russian, that these roots are light (according to Romanova they can have no semantics at all), a comparison of the same root with different prefixes implies some common meaning (for (43), this could be paraphrased as 'to place') while the prefixes add a predictable spatial meaning.

(43) **na**-*staviti* | **po**-*staviti* | **v**-*staviti* | **pre**-*staviti* | **do**-*staviti* | **od**-*staviti* ...
on-set | over-set | in-set | over-set | to-set | from-set
'set' | 'set' | 'insert' | 'move' | 'deliver' | 'remove'

(44) *na-peti* | *vz-peti* | *v-peti* | *raz-peti* | *pri-peti* | *od-peti* ...
on-pull | up-pull | in-pull | apart-pull | at-pull | from-pull
'stretch'/'string' | 'climb' | 'fasten' | 'spread' | 'attach' | 'detach'

(45) *na-jeti* | *pri-jeti* | *za-jeti* | *ob-jeti* | *vz-eti* ...
on-grab | at-grab | behind-grab | around-grab | up-grab
'hire' | 'grab' | 'scoop' | 'hug' | 'take'

And finally, assuming that the forms in (43)–(45) are unprefixed poses a problem for the varieties in which the simplex forms of the verbs in (43)–(45) do exist, and

it also does not account for those *vz-po-staviti*-type verbs that are perfectly normally attested both in standard Slovenian and across Slovenian dialects without the prefix (e.g., *klicati* 'to call', the root of the doubly prefixed verb *v-po-klicati* 'to enlist'). We thus conclude that despite some merits, Markova's account falls short of fully explaining our *vz-po-staviti*-type verbs.

4.3.2 Option 3.2: They are lexical – but these verbs have two VPs (=double resultative structure)

As mentioned in §4.3, the restriction to a single lexical prefix per verb has been derived as a consequence of the structural position of lexical prefixes; because the clausal structure can only have one RP, there can normally only be one lexical prefix per verb phrase (and consequently per verb). However, Žaucer (2009) discusses a class of verbs in Slovenian that seem to have two resultative prefixes, and ultimately analyzes these as having a double-VP structure (cf. also Tatevosov 2022). In the discussion of the cumulative (/accumulative/saturative) prefix *na-*, a crucial piece of support for the double-VP structure is argued to be the two sets of unselected objects, (46) and (47).

(46)   *(pre)-igrati^{PFV} Maradono
       over-play   Maradona.ACC
       'fake out Maradona'

(47)   *(na)-*(pre)-igravati$^{PFV/IPFV}$ se   Maradone
       on-over-play   REFL Maradona.GEN
       'get / getting one's fill of faking out Maradona'

As is evident from our examples in §4.3, the *vz-po-staviti*-type verbs do not behave like this. They do not appear to introduce two unselected objects.

   Furthermore, the outermost prefix in (47) and this type of examples require an imperfective input, which is not the case in *vz-po-staviti*-type verbs. Also, (47) and this type of examples are normally read perfectively, with the outermost prefix there triggering perfectivity; in other words, an example such as (47) does not necessarily get an imperfective reading despite the presence of the imperfective suffix *-ava*. At the same time, though, the imperfective affix *can* be interpreted as scoping over the outermost prefix – in this case the interpretation of (47) is 'getting one's fill of faking out Maradona'. Unlike (47), and as shown in (48), the outermost prefix of *vz-po-staviti*-type verbs never perfectivizes its input and the imperfective affix always scopes over the outermost prefix, which further means that the whole verb is interpreted as imperfective.

(48) a. pri-jeti[PFV] | pri-jemati[IPFV] || o-pri-jeti[PFV] | o-pri-jemati[IPFV]
          at-grab | at-grab.sɪ || around-at-grab | around-at-grab.sɪ
          'to grab' | 'to grab' || 'to grab on to' | 'to grab on to'

      b. pri-nesti[PFV] | pri-našati[IPFV] || do-pri-nesti[PFV] | do-pri-našati[IPFV]
          at-carry | at-carry.sɪ || to-at-carry | to-at-carry.sɪ
          'to carry to' | 'to carry to' || 'to contribute' | 'to contribute'

      c. po-staviti[PFV] | po-stavljati[IPFV] || iz-po-staviti[PFV] | iz-po-stavljati[IPFV]
          over-stand | over-stand.sɪ || out-over-stand | out-over-stand.sɪ
          'to set' | 'to set' || 'to single out' | 'to single out'

While Žaucer (2009) discusses other properties of examples that can be analysed as including two VPs, we take these differences as evidence enough to conclude that prefixes in *vz-po-staviti*-type verbs are not similar to the cumulative *na-*.

4.3.3 Option 3.3: They are lexical – result modifiers, not main result predicates

The literature has identified one further group of prefixes that does not fully respect the standard division into lexical and superlexical. As discussed by Žaucer (2013), prefixes such as excessive (*pre-*), repetitive (*pre-*), attenuative (*pri-*, *po-*), and distributive (*po-*) have adverbial, superlexical-like meanings, can stack, and do not affect argument structure at least when stacked, which makes them look like ordinary superlexical prefixes. An example of this type of prefix is given in (49).

(49) pre-na-polniti
      over-on-fill
      'overfill'

However, Žaucer (2013) argues, contrary to what would be expected given the properties listed above, that these prefixes nevertheless merge VP-internally, supporting this claim, for example, with the fact that they scope below VP-adverbials, as shown in (50). The proposed analysis is that these prefixes are result modifiers, thus a sort of adverbial prefixes, but ones that modify the result phrase directly, before it is merged together with the verb.[13]

---

[13]As already mentioned, this is a possibility not considered by Babko-Malaya (2003), whose analysis explicitly rules out stacked lexicals and resultatives, but it is, as pointed out to us by a reviewer, a possibility that is in fact perfectly compatible with that system.

(50) U-steklenčil sem tole vino sicer na roke, pre-u-steklenčil ga bom pa
      in-bottled   AUX this wine PTCL on hand over-in-bottled   it will PTCL
      z   mašinco.
      with machine
      'Though I bottled this wine manually, I'll re-bottle it with a machine.'
                                                          (Žaucer 2013: 292)

What (50) says is that the first time the wine was bottled it was bottled manually, while the second time it was bottled this was done with the use of a machine, which indicates that the repetitive *pre-* is inside the scope of the 'with'-adverbial, which, in turn, means that *pre-* does not originate above the VP.

   Interestingly, the same scopal facts can be observed with *vz-po-staviti*-type verbs. As shown in (51) the entire verb *oprijeti* 'to hold on to' is in the scope of the 'with'-adverbial, suggesting that all parts of the verb originate VP-internally.

(51) Vejo   sem sicer pri-jel z   roko, o-pri-jel   se   je bom pa
      branch AUX PTCL at-hold with hand around-at-hold REFL it AUX PTCL
      z   rokavico.
      with glove
      'I grabbed the branch with my hand, but I'll hold on to it with a glove.'

The two sets of prefixes also behave the same with respect to the restitutive reading of *spet* 'again'. That is, both the excessive/measure prefix in (52) and the outer prefix in *vz-po-staviti* 'establish' in (53) take narrow scope with respect to the restitutive reading of *spet* 'again'.

(52) Juš je   hladilnik spet   pre-na-polnil.
      Juš AUX fridge   again over-on-filled
      'Juš restored the fridge to an overfilled state.'
      Not: Juš was overly involved in filling up the fridge.   (Žaucer 2013: 293)

(53) Miha je   stike   z   očetom spet vz-po-stavil.
      Miha AUX contacts with father   again up-over-set
      'Miha restored contacts with his father.'
      (No other interpretation.)

While Žaucer's (2013) result-modifying prefixes have a predictable adverbial interpretation and the outer-most prefixes in *vz-po-staviti*-type verbs do not seem to, both of these types of prefixes behave comparably with respect to scopal tests, suggesting that they share the same structural position.[14]

---

[14]Žaucer (2013) does not discuss nominalization possibilities, but root nominalizations from

4.3.4 Option 3.4: They are lexical and parallel to particles

It is well known that there exist parallels between Germanic particles and Slavic prefixes, e.g. Spencer & Zaretskaya (1998), Svenonius (2004). In fact, similarly to doubly-prefixed verbs of the *vz-po-staviti*-type verbs in Slovenian, we can also observe particle recursion in Germanic, see for example den Dikken (1995: 80). den Dikken (1995) claims that particle recursion is structurally possible but, for unclear reasons, rare. He analyzes recursive particles using his basic structural template from Figure 5 by simply having the second particle as the head of XP, as in Figure 6.

(54)   I'll send the letter on over to Grandma's house.
       den Dikken (1995: (116b)), quoting Di Sciullo & Klipple (1994)
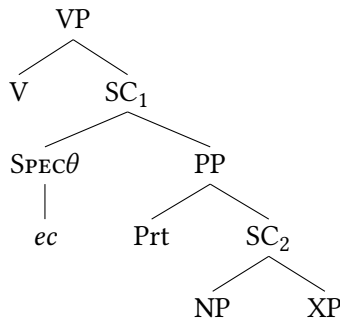
Figure 5: The basic structural template of den Dikken (1995)

4.3.5 Option 3.5: They are some of the lowest projections above VP

There is yet another set of accounts that we have not discussed, namely, accounts that merge all prefixes, including lexical ones, outside the VP. One part of these accounts is represented by systems which at least implicitly still subscribe to two groups, lexical prefixes and a group of higher prefixes, with a single slot for lexical prefixes (e.g. Slabakova 2005, Istratkova 2006, Wiland 2012); like the accounts discussed above, with lexical prefixes originating VP-internally, these accounts thus generally also end up with a restriction to a single lexical prefix. In addition, it is also not clear to us that such systems can really explain argument structure

---

verbs with those result-modifying prefixes are not difficult to find, e.g, *pri-vz-dig* 'a partial lift', *pre-u-stroj* 'remodeling', *pre-u-redba* 'reorganization'. The same holds also of our *vz-po-staviti*-type verbs, cf. (33)–(36) above.
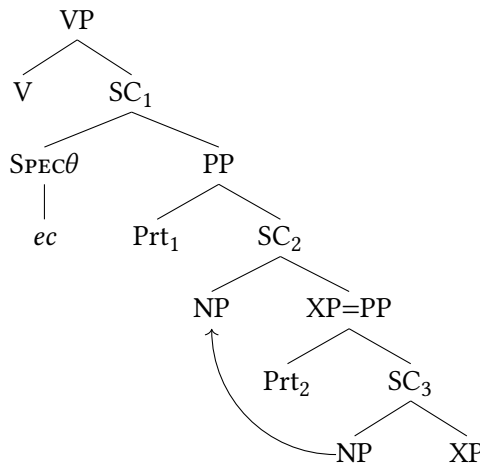
Figure 6: Using den Dikken's (1995) basic structural template to explain unexpected multiple prefixation

effects of lexical prefixes well, cf. Žaucer (2009: 16–18). Most recently, Biskup (2023) also develops a system with all prefixes merged outside the VP, but his version presumably allows more flexibility than the previous all-prefixes-outside-the-VP accounts as it does not really seem to subscribe to two groups, and it does not limit the number of lexical prefixes structurally but rather by appealing to conceptual reasons; for a similar case as our *vz-po-staviti*-type verbs, it explicitly allows two lexical prefixes hosted in two separate internal-prefix phrases above the VP. The approach looks promising to us for approaching our *vz-po-staviti*-type verbs, however, in addition to the concern regarding argument-structure effects already stated above, it is also not clear to us – assuming a universal clausal spine – what the nature of the lexical-prefix projections introducing the multiple lexical prefixes could be, and why they could be freely remergeable.

4.4 Instead of a conclusion—a partial proposal

We have shown that the outer prefixes in *vz-po-staviti*-type verbs, even though they are stacked on top of another prefix, do not behave like other superlexical prefixes but rather much more like VP-internal, lexical prefixes. Table 6 presents a comparison of our *vz-po-staviti*-type verbs, or rather, their outer prefixes, lexical prefixes, superlexical prefixes and result-modifying prefixes on the basis of the six most typically considered properties. Some of these properties are clearly related to one another, so for example, a prefix's VP-internal position is related

to its ability to form a secondary imperfective, which is merged outside the VP and thus scopes over it. Similarly, as already explained in §4.3, placing lexical prefixes in a dedicated VP-internal Result Phrase means that a verb should not host a stack of such prefixes. Additionally, idiosyncratic meaning and argument-structure effects of lexical prefixes also seem to be related to their position inside the VP.

Table 6: Lexical, Superlexical, and other types of prefixes

	Lexical	*vz-po-staviti*	result mod.	Superlex.
VP-positioning	internal	internal	internal	external
meaning	idiosyn./spati.	idiosyn./spati.	adverbial	adverbial
affect arg. struct.	Yes	Yes	No	No
form sec. imperf.	Yes	Yes	Yes	No
form root nomin.	Yes	Yes	Yes	No
stacking	No	Yes	Yes	Yes

So far we mentioned 12 different *vz-po-staviti*-type verbs that used 10 different prefixes as the outer prefix. Most likely, then, the outer prefixes of *vz-po-staviti*-type verbs do not form a homogeneous class of prefixes, so we actually need not expect to find a single explanation for all of them.

The type of verbs that had been discussed by Žaucer (2002) and Svenonius (2004), *iz-pod-riniti* 'to push out' and *s-pod-makniti* 'to jerk away', are probably just instances of a complex prefix which realizes both Path and Place parts of the preposition phrase inside a single result phrase, as suggested by Svenonius (2004).[15]

Some prefixes have a relatively clear spatial meaning, such as *o-* in *o-pri-jeti* 'hold on to', which is comparable in meaning to verbs where *o-* is more clearly lexical like *o-kleniti* 'grab on to', *o-graditi* 'to put a fence around', or *o-črtati* 'to draw a line around' (in some cases the (core) spatial meaning got obscured by a more metaphorical interpretation) and *v-* in *v-po-klicati* 'to enlist', which can even be doubled by a preposition phrase with the same prefix, as in (55).

---

[15]The two combinations *iz-pod-* and *s-pod-* are synonymous. One can find both versions of these two verbs in written Slovenian – *iz-pod-riniti* and *s-pod-riniti* both with the same meaning 'to push out' and likewise *s-pod-makniti* and *iz-pod-makniti* both meaning 'to jerk away'. Spoken Slovenian hardly makes a distinction between the two pronunciations of these two forms, so we are treating them as just two realizations of the same lexical unit.

Marušič, Mišmaš & Žaucer

(55) Trener ga je   v-po-klical v reprezentanco.
      coach   him AUX in-over-call in national-team
      'The coach called him up into the national team.'

In cases like these, the outer prefix may seem to be a proper lexical prefix that would require a result phrase of its own, which would mean that we need two RPs inside the VP, which seems like a problem – but cf. Markova (2011) and Biskup (2023). Note that even though these verbs have a different argument structure from their unprefixed counterpart, the contribution of the prefix to the argument-structure change is not very clear, suggesting that potentially one of the two prefixes can receive an alternative interpretation.

In many respects, our *vz-po-staviti*-type verbs seem to behave similarly to doubly-prefixed verbs in which the prefixes are "result modifiers", the main difference being the interpretation of prefixes/prefixed verbs – while the "result modifiers" in Žaucer (2013) have a clear adverbial reading, prefixes in *vz-po-staviti*-type verbs lead to anything between a slight modification in the interpretation of the input to an full-scale idiosyncratic meaning shift compared to the input. Despite this, we propose that the prefixes in *vz-po-staviti*-type verbs should be subsumable under a result-modifier analysis.[16]

Based on Žaucer (2013), we thus propose that the structure in Figure 7 captures the two positions for the prefixes in *vz-po-staviti*-type verbs. Note that the result-modifying prefix (on its own) here cannot introduce an unselected object (perhaps unlike the structure in Figure 6).
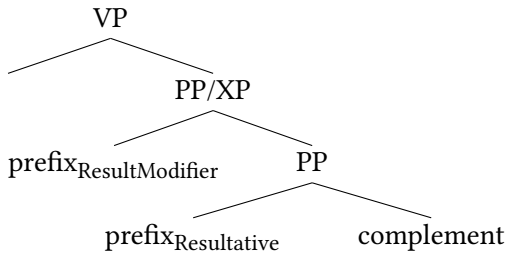

Figure 7: The structure with the two positions of the two prefixes of *vz-po-staviti*-type verbs

---

[16]One could say that just like standard lexical prefixes, which sometimes contribute a compositional spatial interpretation and sometimes a non-compositional idiosyncratic interpretation, result-modifying prefixes also have these two options: contributing either a compositional adverbial interpretation or a non-compositional idiosyncratic interpretation, which we observed with many *vz-po-staviti* type verbs.

5 Conclusions

Our corpus data show that even prefixes which have been claimed to serve (almost) exclusively as lexical prefixes appear stacked over another prefix in up to 20% of their occurrences, which ultimately means that no prefix is used exclusively as a lexical prefix, or that lexical prefixes can sometimes also stack. Our corpus data also confirms a tendency for a hierarchy, but as multiple prefixes have more than one use and since all of them can be used either as lexical or as superlexical prefixes and can appear in more than one position, a true hierarchy of superlexical prefixes could only be determined, perhaps, if prefix occurrences were coded for specific prefix uses – a task that unfortunately seems quite unrealistic, but also one that would inevitably end up drawing in individual researcher's subjective decisions. Our corpus study also showed that whereas prefixed verbs are very common in Slovenian, verbs with stacked prefixes are very rare, all in all making the use of corpora rather poorly suited for investigating prefix stacking options in Slovenian.

On the other hand, our corpus investigation also turned up a sizeable set of verbs with two prefixes in which the outer prefix does not seem to have any of the typical superlexical characteristics, other than the fact that it occurs stacked over another prefix. Zooming in on these verbs, which we called *vz-po-staviti*-type verbs, we compared their outer prefixes to superlexical prefixes, to intermediate (and other types of in-between) prefixes, and to some types of stacked prefixes that had previously been proposed to instantiate lexical prefixes despite being stacked. We argued that both the inner and the outer prefix in *vz-po-staviti*-type verbs are lexical and cannot be explained away easily. We found that the outer prefixes in these verbs do not seem to form a homogeneous class, and so it is quite likely that it need not be just one explanation that will solve all of these examples. Some of the discussed cases can be explained relatively easily, and at least for a large part of them they seem best treated as (a version of) result-modifying prefixes, though some cases may need alternative approaches, which we leave for future research.

Marušič, Mišmaš & Žaucer

Abbreviations

ACC	accusative	IPFV	imperfective
ATT	attenuative	M	masculine
AUX	auxiliary	NOM	nominative
BCMS	Bosnian/Croatian/	NEG	negation
	Montenegrin/Serbian	PERD	perdurative
COMPL	completive	PFV	perfective
CUML	cumulative	PTCL	particle
DAT	dative	REFL	reflexive
DELIM	delimitative	REP	repetitive
DIST	distributive	SAT	saturative
EXC	excessive	SG	singular
F	feminine	SI	secondary imperfective
GEN	genitive	TERM	terminative
INCP	inceptive	TV	thematic vowel

Acknowledgments

References

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

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

Biskup, Petr. 2023. Deriving morphological aspect in Russian: A matryoshka way. In Christina Clasmeier & Julia Golbekand (eds.), *Linguistische Beiträge zur Slavistik: XXIX. JungslavistInnen-Treffen vom 09. bis 10. September 2021 an der Ruhr-Universität Bochum*, 9–30. Berlin, Frankfurt am Main: Peter Lang.

Caha, Pavel & Markéta Ziková. 2022. Prefixes in Czech zero-derived nominalizations and verbs. In Petr Karlík & Lucie Taraldsen Medová (eds.), *Nominalizations and Participles in Czech and beyond*, 3–61. München: Lincom.

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.

Čibej, Jaka, Špela Arhar Holdt, Kaja Dobrovoljc & Simon Krek. 2019. *Frequency lists of words from the Gigafida 2.0 corpus*. Slovenian language resource repository CLARIN.SI. http://hdl.handle.net/11356/1273.

den Dikken, Marcel. 1995. *Particles: On the syntax of verb-particle, triadic, and causative constructions*. Oxford: Oxford University Press.

Di Sciullo, Anna Maria & Elizabeth Klipple. 1994. Modifying affixes. In Sharon Hargus, Gerald R. McMenamin & Vida Samiian (eds.), *Proceedings of the Western conference on linguistics XXIII (WECOL 93)*, 68–80. Fresno, CA: Department of Linguistics, California State University. https://wecol.weebly.com/proceedings.html.

Endo, Yoshio & Bartosz Wiland. 2014. The symmetric syntax of Japanese complex verbs and Slavic prefixes. In Anna Cardinaletti, Guglielmo Cinque & Yoshio Endo (eds.), *On peripheries: Exploring clause initial and clause final positions*, 141–159. Tokyo: Hituzi Syobo Publishing.

Filip, Hana. 2003. Prefixes and the delimitation of events. *Journal of Slavic linguistics* 11(1). 55–101. https://www.jstor.org/stable/24599705.

Fowler, George. 1996. An articulated theory of aspect and prefixation in Slavic. In Jindřich Toman (ed.), *Annual workshop on Formal Approaches to Slavic Linguistics: The College Park meeting 1994*, vol. 20, 97–12. Ann Arbor, MI: Michigan Slavic Publications.

Gehrke, Berit. 2008. *Ps in motion: On the semantics and syntax of P elements and motion events* (LOT Dissertation Series 184). Utrecht: LOT.

Istratkova, Vyara. 2006. When prefixes become adverbs. In Tanya Ivanova-Sullivan, Andrea D. Sims & Miriam Whiting (eds.), *Proceedings of the second graduate colloquium on Slavic linguistics*, vol. 6 (Ohio Slavic Papers), 1–31. Columbus, OH: Department of Slavic, East European Languages & Literatures, The Ohio State University.

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. DOI: 10.1353/jsl.2022.a923068. http://ojs.ung.si/index.php/JSL/article/view/99.

Łaziński, Marek. 2011. Polish aspectual prefixes, their order and functions: A study based on the National Corpus of Polish. *Word Structure* 4. 231–243. DOI: 10.3366/word.2011.0012.

Markova, Angelina. 2011. On the nature of Bulgarian prefixes: Ordering and modification in multiple prefixation. *Word Structure* 4(2). 244–271. DOI: 10.3366/word.2011.0013.

Matushansky, Ora. 2002. On formal identity of Russian prefixes and prepositions. *MIT Working Papers in Linguistics* 42. 217–253.

Milićević, Nataša. 2004. The lexical and superlexical verbal prefix *iz-* and its role in the stacking of prefixes. *Nordlyd* 32(2). 279–300. DOI: 10.7557/12.70.

Ramchand, Gillian. 2004. Time and the event: The semantics of Russian prefixes. *Nordlyd* 32(2). 324–360. DOI: 10.7557/12.72.

Ramchand, Gillian. 2008. *Verb meaning and the lexicon: A first-phase syntax*, vol. 116. Cambridge: Cambridge University Press.

Rappaport, Malka & Hovav Beth Levin. 2001. An event structure account of English resultatives. *Language* 77(4). 766–797. https://www.jstor.org/stable/3086847.

Romanova, Eugenia. 2004. Superlexical versus lexical prefixes. *Nordlyd* 32(2). 255–278. DOI: 10.7557/12.69.

Šekli, Matej. 2016. Pomeni glagolskih predpon v slovenščini. *Filološke studije* 14(1). 273–287. https://hrcak.srce.hr/203716.

Slabakova, Roumyana. 2005. Perfective prefixes: What they are, what flavors they come in, and how they are acquired. In Steven Franks, Frank Y. Gladney & Mila Tasseva-Kurktchieva (eds.), *Formal Approaches to Slavic Linguistics 13: the South Carolina Meeting*, vol. 13, 324–341. Ann Arbor, MI: Michigan Slavic Publications.

Snoj, Marko. 2009. *Slovenski etimološki slovar*. Ljubljana: Modrijan.

Spencer, Andrew & Marina Zaretskaya. 1998. Verb prefixation in Russian as lexical subordination. *Linguistics* 36(1). 1–39. DOI: 10.1515/ling.1998.36.1.1.

Svenonius, Peter. 2004. Slavic prefixes inside and outside VP. *Nordlyd* 32(2). 206–253. DOI: 10.7557/12.68.

Tatevosov, Sergei. 2008. Intermediate prefixes in Russian. In Andrei Antonenko, Christina Bethin & John Baylin (eds.), *Proceedings of the annual workshop on formal approaches to Slavic linguistics 16. the stony brook meeting*, vol. 16, 423–445. Ann Arbor, MI: Michigan Slavic Publications.

Tatevosov, Sergei. 2022. *Multiple prefixation, reflexivization, and event identification*. Paper presented at FASL 31; abstract available at https://fasl.humanities.mcmaster.ca/program/.

Tenny, Carol. 1994. *Aspectual roles and the syntax-semantics interface* (Studies in Linguistics and Philosophy 52). Dordrecht: Springer.

Toporišič, Jože. 2000. *Slovenska slovnica*. Fourth edition. Maribor: Založba Obzorja.

Vidovič Muha, Ada. 1993. Glagolske sestavljenke: Njihova skladenjska podstava in vezljivostne lastnosti. *Slavistična revija* 41. 161–192. https://srl.si/ojs/srl/article/view/COBISS_ID-35424256.

Wiland, Bartosz. 2012. Prefix stacking, syncretism and the syntactic hierarchy. In Markéta Ziková & Mojmír Dočekal (eds.), *Slavic languages in formal grammar: Proceedings of FDSL 8.5, Brno 2010*, 307–324. Berlin, Frankfurt am Main: Peter Lang.

Žaucer, Rok. 2002. *The role of verbal prefixes in Slavic: Evidence from Slovenian locative denominal verbs*. Ottawa, ON: University of Ottawa. (MA thesis).

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

Žaucer, Rok. 2013. A case for result-modifying prefixes. In Alexander Podobryaev (ed.), *Formal Approaches to Slavic Linguistics 20: the Second MIT Meeting*, 290–304. Ann Arbor, MI: Michigan Slavic Publications.

Advances in Formal Slavic Linguistics 2022

Advances in Formal Slavic Linguistics 2022 brings together a collection of 22 articles originating as talks presented at the 15th Formal Description of Slavic Languages conference (FDSL 15) held in Berlin on 5–7 October, 2022. The contributions cover a broad spectrum of topics, including clitics, nominalizations, l-participles, the dual, verbal prefixes, assibilation, verbal and adjectival morphology, lexical stress, vowel reduction, focus particles, aspect, multiple wh-fronting, definiteness, polar questions, negation words, and argument structure in such languages as BCMS, Bulgarian, Czech, Macedonian, Polish, Russian, Slovenian, Ukrainian, and Upper Sorbian. The wide range of topics explored in this volume underscores the the diversity and complexity of Slavic languages. The contributions not only advance our understanding of languages belonging to the Slavic group but also offer fresh perspectives for the linguistics more broadly.