Advances in formal Slavic linguistics 2017

Edited by

Franc Marušič

Petra Mišmaš

Rok Žaucer

language
science
press

Open Slavic Linguistics

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

Advances in formal Slavic linguistics 2017

Edited by

Franc Marušič

Petra Mišmaš

Rok Žaucer

Freie Universität Berlin

Contents

Contents

Preface

Advances in Formal Slavic Linguistics 2017 is a collection of fifteen articles on formal Slavic linguistics. The articles cover a wide array of topics, such as control verbs, instrumental arguments, and perduratives in Russian, comparatives, negation, n-words, negative polarity items, and complementizer ellipsis in Czech, impersonal *se*-constructions and complementizer doubling in Slovenian, prosody and the morphology of multi-purpose suffixes in Serbo-Croatian, and indefinite numerals and the binding properties of dative arguments in Polish. By exploring these phenomena in individual Slavic languages, the collection of articles in this volume makes a significant contribution to both Slavic linguistics and to linguistics in general.

The articles in this volume were prepared on the basis of talks and posters presented at the conference Formal Description of Slavic Languages 12.5, which was held on December 7-9, 2017, at the University of Nova Gorica. FDSL was initially a biannual conference, hosted in turn by the University of Leipzig and University of Potsdam since 1995, with the University of Göttingen joining the main conference cycle in 2011 and Humboldt-Universität zu Berlin in 2016. In 2006 FDSL was partly turned into an annual event when the University of Nova Gorica hosted FDSL 6.5. Eleven years later, after successful stops in Moscow in 2008 and Brno in 2010 and 2014, the 12.5 conference thus brought FDSL back to Nova Gorica, the place of its first halftime edition. The 18 regular-session talks and 10 poster presentations given at FDSL 12.5 were selected out of 47 submitted abstracts. The conference also featured 5 invited talks. Following the conference, 17 articles were submitted for inclusion in this volume, and 15 successfully completed the two-round review process in which every article was evaluated and commented on twice by two reviewers.

This volume would not have been possible without our extremely helpful reviewers: Nadira Aljović, Svitlana Antonyuk, Boban Arsenijević, Loren A. Billings, Petr Biskup, Joanna Błaszczak, Anna Bondaruk, Olga Borik, Mojmír Dočekal, Jakub Dotlačil, Berit Gehrke, Guillaume Enguehard, Julie Goncharov, Hana Gruet-Skrabalova, Peter Jurgec, Dorota Klimek-Jankowska, Iliyana Krapova, Jonathan E. MacDonald, Christina Manouilidou, Tatjana Marvin, Nataša Milićević, Moreno

Preface

Chapter 1

Object control: Hidden modals

Irina Burukina

Eötvös Loránd University

The paper proposes a novel analysis for object control verbs in Russian. First, I argue that object control verbs are not a homogeneous class, despite the common opinion advocated by Franks & Hornstein (1992); Babby (1998); Boeckx et al. (2010), a.o. In Russian, desiderative object control verbs with a dative argument (*velet'* 'order', *razrešit'* 'allow', etc.) differ significantly in their syntactic and semantic properties from implicative object control verbs with an accusative argument (*zastavit'* 'force', *ubedit'* 'persuade', etc.). However, this distinction does not match existing classifications. Second, I develop a structural description for dative control verbs and argue that (i) the dative argument and the embedded clause form a single constituent that excludes the matrix verb, and (ii) this constituent is headed by a silent modal element. This analysis accounts for many semantic and syntactic properties of dative object control verb including the unavailability of split control with dative control verbs and their distributional similarity with modal predicatives (*možno* 'allowed', *nado* 'necessary').

Keywords: object control, non-finite complement, covert modality, dative, Russian

1 Introduction

The present paper investigates Russian OBJECT CONTROL VERBS (OCVs) that require either a dative or an accusative argument: *velet'* 'order', *prikazat'* 'order', *zastavit'* 'make' etc. During the past decades several major theories of control have been developed (Wurmbrand 2001; Boeckx et al. 2010; Landau 2015), however, Slavic languages have not been sufficiently approached (Franks & Hornstein 1992; Babby 1998). Existing approaches usually draw a line between subject and object control predicates, and the latter are treated as a homogeneous class. The most significant attempt for further sub-categorization has been made by

Landau (2015), and has resulted in the development of the two-tiered theory of control. However, when the Russian data is analyzed, the two-tiered theory of control makes a wrong prediction: the suggested attitude vs. non-attitude distinction does not correspond to the real availability of partial control.

The classification proposed in this paper captures the correlation between syntactic and semantic properties of Russian OCVs, partially inheriting Jackendoff & Culicover (2003) idea to sub-categorize verbs of control according to their meaning. I distinguish between desiderative dative predicates (1a), on the one hand, and implicative accusative predicates (1b), on the other.[1] As will be demonstrated, the two groups differ significantly in their behavior.

(1) a. Petja   razrešil Maše$_i$ PRO$_i$ vzjat'   kuklu.
        Petja.NOM allowed Maša.DAT   take.INF doll
        'Petja allowed Maša to take the doll.'

    b. Petja   zastavil Mašu$_i$ PRO$_i$ vzjat'   kuklu.
        Petja.NOM forced Maša.ACC   take.INF doll
        'Petja forced Maša to take the doll.'

The paper continues with a novel two-part analysis for Russian dative OCVs: (i) the dative argument and the embedded clause form a single constituent that excludes the main predicate head, and (ii) this constituent is headed by a silent modal-like element that takes a non-finite clause as its complement (2).[2] Therefore, the core claim is that the modal item is not merely a part of semantic decomposition, but that it is present in the syntactic structure, separately from the main predicate. I further suggest that, in Russian, this silent modal head belongs to the existing class of modal predicatives (*možno* 'allowed', *nado* 'necessary', etc.).

(2) [$_{vP}$ Petja [$_{VP}$ [$_V$ razrešil][$_{ModP}$ [Maše$_i$] [$_{Mod'}$ *modal* [$_{CP}$ PRO$_i$ vzjat'
        Petja   allow   Maša   take
    kuklu]]]]]
    doll
    'Petja allowed Maša to take the doll.'

The rest of the paper is structured as follows: §2 examines general properties of Russian dative OCVs in comparison with accusative OCVs and addresses the con-

---

[1] I use the terms "desiderative" and "implicative" to refer to these particular verbs following Wurmbrand (2001) and Landau (2013).

[2] I leave the question about the size of the embedded non-finite clause for further research and I mark it as a CP, adopting the traditional Lasnik (1998) approach to infinitives in Slavic languages.

stituency issue; and §3 presents the analysis of sentences with a matrix desiderative control predicate. Finally, §4 discusses the peculiar unavailability of split control in the presence of a desiderative OCVs.

2 Desiderative control predicates and their properties

Desiderative OCVs in Russian include the verbs *razrešit'* 'allow', *pozvolit'* 'allow', *zapretit'* 'prohibit', *prikazat'* 'order', *velet'* 'order', *predpisat'* 'obligate', *poručit'* 'charge', *skazat'* 'tell'. They cannot assign accusative case and require a dative DP. OCVs that select an accusative argument are implicatives *zastavit'* 'force', *vynudit'* 'force', *ugovorit'* 'persuade', *ubedit'* 'persuade' (3).

(3) a. Vrač   zapretil Ivanu$_i$ PRO$_i$ jest'   konfety.
        doctor.NOM forbade Ivan.DAT   eat.INF candies
        'The doctor forbade Ivan to eat candies.'

    b. Vrač   ubedil   Ivana$_i$ PRO$_i$ ne jest'   konfety.
        doctor.NOM persuaded Ivan.ACC   not eat.INF candies
        'The doctor persuaded Ivan not to eat candies.'

As demonstrated in (3), both types of control predicates under consideration can embed a non-finite clause. Aside from this, desiderative OCVs often embed a finite subjunctive clause; importantly, a dative DP is still present and its referent can differ from that of the embedded subject (4).[3]

(4) a. Vrač   zapretil Ivanu$_i$, čtoby on$_i$ jel   konfety.
        doctor.NOM forbade Ivan.DAT so that he.NOM eat.SBJV candies
        'The doctor forbade Ivan to eat candies.'

---

[3]Accusative OCVs only marginally allow embedded finite clauses; in such cases the accusative DP and the embedded subject should be co-referential (i). It is beyond the scope of this paper to examine these constructions in details, and I will only briefly return to this problem in §3.

(i) a. ?? Vrač   zastavil Ivana$_i$, čtoby on$_i$ ne jel   konfety.
         doctor.NOM forced   Ivan.ACC so that he.NOM not eat.SBJV candies
         'The doctor forced Ivan not to eat candies.'

    b. * Vrač   zastavil medsestru, čtoby Ivan   ne jel   konfety.
         doctor.NOM forced   nurse.ACC so that Ivan.NOM not eat.SBJV candies
         Intended: 'The doctor told the nurse that Ivan must not eat candies.'

b. Vrač   zapretil medsestre, čtoby Ivan   jel   konfety.
   doctor.NOM forbade nurse.DAT so that Ivan.NOM eat.SBJV candies
   'The doctor told the nurse that Ivan should not eat candies.'

Constructions with implicative and desiderative OCVs in Russian differ when it comes to structural relations established between a DP argument and an embedded clause. In sentences with an implicative OCV an accusative DP and an embedded clause together do not pass constituency tests; they cannot be separated from the rest of a clause, excluding the main predicate. This is exemplified in (5), where attempted pseudo-cleft and short answer constructions are ungrammatical.

(5) a. * Čto ja zastavil, tak eto Petju   pojti   v   kino.
        what I forced   so that Petja.ACC go.DAT into cinema
        Intended 'What I forced is that Petja would go to the cinema.'

     b. * Čto ty zastavil? Petju   pomyt'   posudu.
        what you forced   Petja.ACC wash.INF dishes
        Intended: 'What did you force him to do? I forced Petja to wash the dishes.'

In contrast, a dative DP and an embedded clause apparently form a single constituent that excludes the main predicate in sentences with a desiderative OCV; see the examples in (6), which are judged as acceptable by native speakers.

(6) a. Čto ja razrešil, tak eto Pete   pojti   v   kino.
        what I allowed so that Petja.DAT go.INF into cinema
        'What I allowed is that Petja would go to the cinema.'

     b. Čto ty razrešil? Pete   pojti   v   kino.
        what you allowed   Petja.DAT go.INF into cinema
        'What did you allow? I allowed Petja to go to the cinema.'

It is possible to suggest that the sentences in (5) are ungrammatical because of the case assignment problems: as a structural case, accusative is licensed by a functional head that must be structurally present. Nevertheless, this does not directly affect the results of the constituency tests for sentences with desiderative predicates, as we would not expected two unrelated constituents to be clefted or questioned.

Furthermore, the boundaries of the immediate constituent that includes the dative DP and the embedded clause and, apparently, does not contain the matrix predicate, become visible in multiple wh-questions and in case of quantifier

stranding. As a result of successive cyclic A-bar movement, an interrogative pronoun or a quantifier from a moved phrase can stay in an intermediate position; the position at the edge of an embedded clause is expected; however, there is another one, to the left of the dative DP. All possible positions where an interrogative pronoun or a quantifier can be realized are shown in (7).

(7) a. Kto (kuda$_i$) razrešil [(kuda$_i$) emu [(kuda$_i$) pojti t$_i$]]?
who where allowed where he.DAT where go.INF
'Who allowed him to go where?'

b. (Vse) knigi$_i$ ja razrešil [(vse) Maše [(vse) pročitat' t$_i$]].
all books.ACC I allowed all Maša.DAT all read.INF
'As for the books, I allowed Maša to read all of them.'

A possible way to account for the examples in (7) is in terms of scrambling across the border of an embedded clause. However, as demonstrated by Bailyn (2003), a.o., the long-distance scrambling in Russian is normally limited to a movement of a constituent from the embedded clause into the right focus position of the matrix clause, and embedded constituents cannot move freely. Furthermore, the ungrammaticality of (8a) suggests that it is also usually unacceptable to put a matrix constituent between the main predicate and the dative DP. Finally, the contrast in speakers' judgments for sentences with a desiderative OCV (7) and an implicative OCV (9) provides additional support for the idea about the presence of a syntactic boundary.[4]

(8) a. *Kto kuda razrešil včera emu pojti?
who where allowed yesterday he.DAT go.INF
Intended: 'Who allowed him yesterday to go where?'

b. Kto kuda včera razrešil emu pojti?
who where yesterday allowed he.DAT go.INF
'Who allowed him yesterday to go where?'

(9) a. ??? Kto zastavil kuda ego pojti?
who forced where he.ACC go.INF
Intended: 'Who forced him to go where?'

b. *Knigi$_i$ ja zastavil [(vse) Mašu [pročitat' t$_i$]].
books.ACC I forced all Maša.ACC read.INF
Intended: 'As for the books, I forced Maša to read all of them.'

---

[4]24 out of 32 native speakers reported (8a) to be ungrammatical; 6 native speakers said that it is 'weird'.

In addition to the syntactic differences described above, desiderative and implicative OCVs can be distinguished with regard to one additional property: availability of partial control for PRO. Desiderative OCVs require exhaustive control, while implicative OCVs allow partial coreference between an embedded PRO and its matrix controller.[5] Evidence for this comes from constructions with a semantically single dative or accusative controller and various embedded elements that require a semantically plural subject. For this purpose, first, collective predicates can be used; in Russian many of those are derived using the pattern *raz + sja*: *razrugat'sja* 'break up, quarrel', *razojtis'* 'disperse', etc. (10).[6]

(10) a. Mal'čiki   razrugalis'.
          boys.NOM.PL broke up
          'The boys broke up.'

      b. Komanda   razrugalas'.
          team.NOM.SG broke up
          'The team broke up.'

      c. *Petja   razrugalsja.
          Petja.NOM broke up
          Intended: 'Petja broke up (with someone).'

Placed in an embedded non-finite clause with a single matrix controller, these verbs are allowed only if the matrix predicate is implicative (11a), and they are prohibited in sentences with a desiderative OCV (11b).

(11) a. Ivan   ubedil'   direktora razojtis' v sem'.
          Ivan.NOM persuaded director.ACC disperse.INF at seven
          'Ivan persuaded the director that they should disperse at seven.'

      b. *Direktor   razrešil Ivanu razojtis' v sem'.
          director.NOM allowed Ivan.DAT disperse.INF at seven
          Intended: 'The director allowed Ivan to disperse at seven.'

Collective modifiers, for example, *vmeste* 'together', behave the same way. When subject-oriented, *vmeste* requires semantically plural PRO; in case of a semantically singular controller, *vmeste* is permitted in constructions with a matrix implicative verb (12a), but not a matrix desiderative predicate (12b).

---

[5]In case of partial control referent of the controller still must be included among referents of PRO.

[6]The same *raz + sja* pattern is used to derive non-collective predicates with different meanings and (often) selection properties. For example, *Petja razrugalsja* can be considered grammatical if the verb is interpreted as the homonymous one meaning 'begin to swear at someone angrily'.

(12)   a.   Direktor   zastavil Ivana   poobedat'   vmeste.
          director.NOM forced   Ivan.ACC have lunch.INF together
          'The director forced Ivan to have lunch together with him.'

    b.   *Direktor   velel   Ivanu   poobedat'   vmeste.
          director.NOM ordered Ivan.DAT have lunch.INF together
          Intended: 'The director ordered Ivan to have lunch together with
          him.'

Taking into account all the differences in the behavior of implicative and desiderative OCVs in Russian, I propose that the two groups require separate structural representations. In this paper I focus on desiderative dative OCVs and I proceed by suggesting a novel way to analyze these predicates. The core idea is that, aside from the main verb, an additional silent deontic modal head should be structurally introduced to connect a dative DP and an embedded clause.[7]

3 Proposed analysis

The two prominent current frameworks that address the problem of detailed sub-categorization of control verbs, namely Wurmbrand's (2001) theory of restructuring configurations and Landau's (2015) two-tiered theory of control, cannot fully account for the Russian data. According to Wurmbrand's classification based on the structural properties of embedded clauses selected by various control predicates, Russian implicative and desiderative OCVs fall into one category of restructuring irrealis predicates; the differences between the two types remain unexplained further.

Landau (2015) distinguishes between attitude and non-attitude predicates, selecting attitude and non-attitude complements. The former refer to the world of the main actor's beliefs; the later receives an interpretation with regard to the real world context. This semantic difference yields different syntactic structures, with additional functional projections above an embedded clause required by attitude predicates. Importantly, the theory predicts that attitude predicates (desideratives, propositionals) must support partial control, while non-attitude

---

[7]At least two options might be suggested for a structural representation of implicative OCVs: (i) an accusative DP and an embedded clause are both internal arguments of the main verb, located in SpecVP and CompVP, respectively (Babby 1998; Bailyn 2012 on Russian), or (ii) an accusative DP and an embedded clause together form a small clause in the complement position of the matrix 'causative-like' predicate (Franks & Hornstein 1992; Landau 2015, a.o.). I am unable to provide a detailed comparison of these two approaches within the limits of this paper and I leave this problem for future investigations.

predicates (implicatives, modals) must require strict coreference between PRO and its controller. As was demonstrated in §2, as for the Russian data, this prediction is not borne out: attitude desiderative dative OCVs allow only exhaustive control, while non-attitude implicative accusative OCVs do not prohibit partial coreference.

Thus, another way should be found to represent the structure of desiderative OCVs. I propose that sentences with these predicates contain a hidden component that is responsible for their peculiar properties – a lexical deontic modal head that, in Russian, belongs to the class of the so-called modal predicatives.

The results of the constituency tests provided in §2 suggest that the dative controller and the embedded clause form a single constituent that excludes the matrix predicate. The question remains about the nature of this constituent; it might be suggested that the two form a small clause or there can be another lexical head that selects a dative DP and a clause as its arguments. I argue that the second option is more plausible and that this lexical head is a silent modal.

In case of desiderative predicates the embedded non-finite clause is a fully saturated proposition; for example, it can be individually substituted by a proform (13a) or an interrogative pronoun (13b).

(13) a. Ja uže   velel Pete   eto.
          I already ordered Petja.DAT this
          'I have already ordered Petja to do this.'

      b. Čto ty velel   Pete?
          what you ordered Petja.DAT
          'What did you order Petja to do?'

It is unlikely that a saturated embedded clause itself functions as a predicate of the dative argument. This is further supported by the availability of an embedded finite clause; as was illustrated in (4), if a finite subjunctive clause is selected, a dative DP argument is still available. Importantly, the latter does not have to be coreferent with the embedded subject (4b), which rules out possible copy-raising analyses. No semantic or syntactic difference can be found between a dative DP present together with an embedded infinitival construction and an argument selected simultaneously with a finite clause. Therefore, I assume that there is no reason to believe that the two are related to different predicates.

I propose that a dative DP and an embedded clause (either finite or non-finite) are selected together by a silent lexical modal head; this modal phrase is later merged as a complement of a desiderative OCV. The structure is schematized in (14). In other words, I argue that deontic modality, intuitively perceived in

desiderative predicates, is represented structurally. At least two properties of desiderative OCV constructions support this idea.

(14) [$_{VP}$ [$_V$ desiderative][$_{ModP}$ [dative DP][$_{Mod'}$ modal [$_{CP}$ infinitival clause]]]]

Firstly, there is the possibility of ambiguous interpretation of examples with sentential negation. Consider the example in (15), for which two readings (15a) and (15b) are available, while interpretation (15c) is prohibited. However, simply assuming that *razrešit'* 'permit' allows Neg-raising, we would expect negation to scope either above the matrix verb (reading (15a)) or above an embedded clause (reading (15c)).

(15) Petja   ne razrešal Maše   ostat'sja.
      Petja.NOM not allowed Maša.DAT stay.INF
      'Petja didn't allow Maša to stay.'

   a. 'Petja said that for Maša it is not possible to stay.'
   b. 'Petja didn't say that for Maša it is possible to stay.'
   c. Not available: 'Petja said that for Maša it is possible not to stay.'

Furthermore, according to von Fintel & Iatridou (2007) and Iatridou & Zeijlstra (2013), a.o., predicates denoting permission typically do not support neg-raising; see, for example, English modal verbs and Russian modal predicatives (16). Although this generalization is originally formulated for deontic modals, deontic modality is a part of constructions with desiderative OCVs like order and permit, and an additional explanation is required for (15) if we assume that this modality is encoded in *razrešit'* itself.

(16) a. Ivan cannot stay.
         = Ivan must leave. ≠ Ivan may leave.
      b. Pete   nel'zja   ostavat'sja.
         Petja.DAT not allowed.SG.N stay.INF
         'Petja is not allowed to stay here.'
         = Petja must leave. ≠ Petja may leave

Introducing a separate deontic head, as shown in (14), splits the structure into two parts: the higher 'communication' component and the lower 'permission' constituent. In (15) negation can scope above either one of them yielding the interpretations (15a) and (15b); however, the may-type modal prohibit neg-raising and the interpretation (15c) becomes impossible.

Second, almost all predicates of communication in Russian have desiderative 'counterparts'. As demonstrated in (17a), Russian verbs of communication, similar to English say, write, whisper, normally embed a finite indicative clause with the *čto* complementizer. However, they can also co-occur with embedded non-finite or finite subjunctive clauses, and such sentences receive a desiderative (modal) interpretation (17b)–(17c). It is true that the contrast between (17a) and (17b)–(17c) could, in principle, be accounted for by postulating two morphologically identical lexical entries for each of the verbs of information transfer. However, there is another possible explanation in line with the hidden modal approach presented in this paper: verbs of communication always select a saturated proposition, that can be either a finite indicative CP or a saturated modal phrase with two arguments (18). Although the behavior of verbs of communication alone does not prove that the proposed analysis is the correct one, taking into account the Neg-raising facts reported above, being able to capture both of these properties of sentences with a desiderative OCV is an important advantage of the hidden modal approach.

(17) a. Petja   {skazal / napisal / šepnul}   Maše,   čto Ivan
        Petja.NOM said   wrote   whispered Maša.DAT that Ivan.NOM
        pomyl   posudu.
        washed dishes
        'Petja {said/wrote/whispered} Maša that Ivan had washed the dishes.'

     b. Petja   {skazal / napisal / šepnul}   Maše   pomyt' posudu.
        Petja.NOM said   wrote   whispered Maša.DAT wash.INF dishes
        'Petja {said/wrote/whispered} that Maša should wash the dishes.'

     c. Petja   {skazal / napisal / šepnul}   Maše,   čtoby ona
        Petja.NOM said   wrote   whispered Maša.DAT so that she.NOM
        pomyla   posudu.
        wash.SBJV dishes
        'Petja {said/wrote/whispered} that Maša should wash the dishes.'

(18) Petja   {skazal / napisal / šepnul}   [Maše   Ø   pomyt'
     Petja.NOM said   wrote   whispered Maša.DAT necessary wash.INF
     posudu].
     dishes
     'Petja {said/wrote/whispered} that Maša should wash the dishes.'
     = 'Petja {said/wrote/whispered} that for Maša it is necessary to wash the dishes.'

In (19) the proposed structure is repeated; at this point I denote the complex modal constituent as ModP and leave the exact size of it for future investigation.

(19)   [$_{VP}$ [$_V$ desiderative][$_{ModP}$ [dative DP][$_{Mod'}$ modal [$_{CP}$ infinitival clause]]]]

Furthermore, I argue that the embedded silent head belongs to the existing class of deontic modals. In Russian, in addition to modal verbs, there is also a group of the so called modal predicatives (*nado* 'necessary', *možno* 'allowed'). Modal predicatives prohibit a nominative subject and require a dative DP argument (20); morphologically, they are invariant and usually end with a neutral singular ending -*o*.

(20)   Ivanu   možno   ne rabotat' segodnja.
       Ivan.DAT allowed.SG.N not work.INF today
       'Ivan is allowed not to work today.'

Similar to desiderative OCVs, modal predicatives embed a non-finite or a finite subjunctive clause (21).

(21)   a. Petja   velel   Maše$_i$ {ostat'sja / čtoby ona$_i$ ostalas'}.
           Petja.NOM ordered Maša.DAT stay.INF   so that she.NOM stay.SBJV
           'Petja ordered Maša to stay.'

       b. Maše$_i$ nužno   {ostat'sja / čtoby ona$_i$ ostalas'}.
           Maša.DAT necessary.SG.N stay.INF   so that she.NOM stay.SBJV
           'Maša should stay.'

However, in contrast with sentences with desiderative OCVs, in constructions with an overt modal predicatives a dative DP and an embedded clause do not form a single constituent: the two cannot be separated together in clefts (22) and short answers (23).

(22)   * Čto možno,   tak eto Pete   posmotret' multiki.
         what allowed.SG.N so that Petja.DAT watch.INF cartoons
         Intended: 'What is allowed is for Petja to watch cartoons.'

(23)   Čto možno?   *Pete   posmotret' multiki.
       what allowed.SG.N Petja.DAT watch.INF cartoons
       Intended: 'What is allowed? It is allowed for Petja to watch cartoons.'

For deontic modal predicatives, I propose the following structural representation (24). Importantly, as was already said about silent modals in desiderative

constructions, I argue that modal predicatives are lexical heads, not functional elements. At first sight, this idea contradicts existing analyses of modals (Cinque 1999; Wurmbrand 1999, a.o.); however the latter usually consider only modal agreeing verbs, whereas the present paper discusses a different class of modal elements. Predicatives in Russian select a wide variety of constructions as a complement; crucially, they select embedded finite subjunctive clauses (25), which is a reflex of the lexical-semantic properties of the modal head.

(24) $[_{\text{ModP}}$ DP $[_{\text{Mod'}}$ modal [embedded proposition]]]

(25) Tebe   neobxodimo, čtoby ty   vypolnil eto zadanie.
      you.DAT necessary.SG.N so that you.NOM do.SBJV   this task
      'It is necessary for you to do this task.'

Going back to sentences with a desiderative OCV, the final structural representation and an example are given in (26).

(26) $[_{v\text{P}}$ Petja $[_{\text{VP}} [_{\text{V}}$ skazal]$[_{\text{ModP}}$ [Maše$_i$]$[_{\text{Mod'}}$ modal $[_{\text{CP}}$ PRO$_i$
      Petja.NOM   said   Mary.DAT
      vzjat'   kuklu   ]]]]]
      take.INF doll.ACC
      'Petja told Maša that for her it is necessary to take the doll.'

4 Split control

The proposed analysis for desiderative OCVs provides a straightforward explanation for the unavailability of split and partial coreference ((27) reproduced from (11) and (12)).

(27) a. * Direktor   razrešil Ivanu   razojtis'   v sem'.
           director.NOM allowed Ivan.DAT disperse.INF at seven
           Intended: 'The director allowed Ivan to disperse at seven.'
      b. * Direktor   velel Ivanu   poobedat'   vmeste.
           director.NOM ordered Ivan.DAT have lunch.INF together
           Intended: 'The director ordered Ivan to have lunch together with him.'

As illustrated in (26), the main predicate (interpreted as desiderative) selects a propositional modal-headed constituent. Within this phrase the control relation is established strictly between the dative argument and the embedded PRO, adherent to the Minimal Distance Principle.

5 Conclusion

In this paper I have used the Russian data to demonstrate that OCVs are not a homogeneous class, and that they can be sub-categorized based on their semantic and syntactic properties. Not rejecting Landau's (2015) attitude vs. non-attitude predicates dichotomy, I propose to distinguish between implicative predicates, which require an accusative argument, and desiderative predicates, which cannot assign accusative case and require a dative controller.

The developed structural representation for desiderative dative OCVs are two-fold: (i) the dative argument and the embedded clause are united into a single constituent that excludes the matrix verb, and (ii) this constituent is headed by a silent deontic modal. I suggest that the central idea of the proposed analysis – syntactic decomposition of desiderative predicates into a verb of communication and a silent modal head – can be further applied to other languages.

Abbreviations

ACC	accusative	OCV	object control verbs
DAT	dative	PL	plural
INF	infinitive	SBJV	subjunctive
N	neuter	SG	singular
NOM	nominative		

Acknowledgements

References

Babby, Leonard H. 1998. Subject control as direct predication: Evidence from Russian. In Željko Bošković, Steven Franks & William Snyder (eds.), *Formal Approaches to Slavic Linguistics 6: The Connecticut Meeting 1997*, 17–37. Ann Arbor, MI: Michigan Slavic Publications.

Bailyn, John F. 2003. A (purely) derivational account of Russian scrambling. In Wayles Browne, Ji-Yung Kim, Barbara Partee & Robert Rothstein (eds.), *Formal Approaches to Slavic Linguistics 11: The Amherst Meeting 2002*, 41–62. Ann Arbor, MI: Michigan Slavic Publications.

Bailyn, John F. 2012. *The syntax of Russian.* Cambridge: Cambridge University Press.

Boeckx, Cedric, Jairo Nunes & Norbert Hornstein. 2010. *Control as movement.* New York: Cambridge University Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective.* Oxford: Oxford University Press.

Franks, Steven & Norbert Hornstein. 1992. Secondary predication in Russian and proper government of pro. In *Control and grammar*, 1–50. Dordrecht: Springer.

Iatridou, Sabine & Hedde Zeijlstra. 2013. Negation, polarity, and deontic modals. *Linguistic Inquiry* 44(4). 529–68. DOI:10.1162/ling_a_00138

Jackendoff, Ray & Peter W. Culicover. 2003. The semantic basis of control in English. *Language* 79(3). 517–56. DOI:10.1353/lan.2003.0166

Landau, Idan. 2013. *Control in generative grammar: A research companion.* New York: Cambridge University Press.

Landau, Idan. 2015. *A two-tiered theory of control.* Cambridge, MA: MIT Press.

Lasnik, Howard. 1998. Exceptional case marking: Perspectives old and new. In Željko Bošković, Steven Franks & William Snyder (eds.), *Formal Approaches to Slavic Linguistics 6: The Connecticut Meeting 1997*, 187–211. Ann Arbor, MI: Michigan Slavic Publications.

von Fintel, Kai & Sabine Iatridou. 2007. Anatomy of a modal construction. *Linguistic Inquiry* 38(3). 445–83. DOI:10.1162/ling.2007.38.3.445

Wurmbrand, Susanne. 1999. Modal verbs must be raising verbs. In Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest (eds.), *WCCFL 18: Proceedings of the 18th West Coast Conference on Formal Linguistics*, 599–612. Somerville, MA: Cascadilla Press.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure.* Berlin: Mouton de Gruyter.

Chapter 2

N-words and NPIs: Between syntax, semantics, and experiments

Mojmír Dočekal

Masaryk University in Brno

In this paper I experimentally approach the following question: do strict negative concord languages like Czech employ two strategies (syntactic and semantic) to encode negative dependency between a verb and its argument(s) or not? And the answer is: beside the default syntactic strategy (n-words), there is a class of negative dependent expressions which are licensed by semantic rules.

Keywords: n-words, negative polarity items, experiments, neg-raising, Czech

1 Introduction

In this article I focus on a problem of dividing negative dependent expressions into two classes: (i) N-words like Czech *nikdo* 'nobody' or Romanian *nimeni* 'nobody' (glossed as N-person) in (1a); (ii) negative polarity items (NPIs) like Czech *sebemenší šance* 'slightest chance' or Romanian *vreun* 'any' in (1b). Despite the long research traditions on both types of expressions (for NPIs see Heim 1984; Ladusaw 1992; Kadmon & Landman 1993; Krifka 1995; Giannakidou 1997; Lahiri 1998; Gajewski 2011; Chierchia 2013; Crnič 2014 among many others; for n-words see Laka 1990; Zeijlstra 2004; 2008 among others) there is still no consensus on the relationship between the two classes of items.[1]

---

[1]Both n-words and NPIs are generally grammatical in sentences with a negated verb. There are of course language-specific differences, e.g. English NPI *any* usually cannot appear in subject position, Romanian *vreun* 'any' in (1b) behaves similarly but many languages allow NPIs to freely occur in subject position – Błaszczak (2001) lists Hindi, Korean, Japanese among many other languages where NPIs are licensed in any position of a sentence with a negated verb. Slavic languages discussed in detail further belong to the set of languages allowing NPIs in subject position, too.

(1)   a.   Nimeni   nu a   venit.
              N-PERSON not has come
              'Nobody came.'

        b. * Vreun student nu a   venit.
             NPI   student not has come
             Intended: 'No student did not come.'

                               (Romanian; Fălăuş & Nicolae 2016: 586, 591)

Nevertheless, it seems to be settled that the division between n-words and NPIs correlates with the division between syntactic licensing and semantic licensing along the following lines:

    1) N-WORDS are syntactically negative dependent expressions;[2]

    2) NEGATIVE POLARITY ITEMS are semantically negative dependent expressions.

Some languages lexicalise the difference between NPIs and n-words, as shown in the example (1) but sometimes the distinction manifests itself only via stress (and usually consequently) focus marking. An example of the second strategy is in (2) from Giannakidou & Zeijlstra (2017) where the non-focused expression *kanenan* 'anybody' is (according to standard criteria) an NPI while the focused expression KANENAN 'N-PERSON' behaves as a n-word.

(2)   a. Dhen idhe kanenan   o   Janis.
           not   saw NPI.person the John
           'John didn't see anybody.'

        b. Dhen idhe KANENAN o   Janis.
           not   saw N-PERSON the John
           'John didn't see anybody at all.'

                        (Greek; Giannakidou & Zeijlstra 2017: 17)

---

[2]This classification is of course very schematic and it can be a bit problematic to apply it to a set of typologically diverse languages. Consider e.g. Romance languages where it seems to be possible to use n-words in questions and in context without overt verbal negation (cases of indirect negative verbs like *doubt* a.o.). Romance languages (and generally all non-strict negative concord languages) allow moreover preverbal n-words in affirmative sentences (as a rule in non-strict negative concord languages, preverbal n-words require positive verb, unless the speaker wants to convey a double negation reading: see Laka (1990) for many examples and further details). But even if the cross-linguistic scenery of n-words is more nuanced than the distinction n-words=syntax, NPIs=semantics suggests, the classification is generally correct and can be applied even to Romance (and generally non-strict negative concord languages), once our syntactic toolbox is supplemented with phonologically null operators which license n-words (see Zeijlstra 2004 a.o. for such a theory) – the licensing of such operators is of course highly constrained (see again Zeijlstra 2004 and Zeijlstra 2008 for details).

Next to the classification of n-words as being basically licensed in syntax (either via agreement or some other standard syntactic process) and NPIs as semantically dependent expressions (occurring only in environments with specific monotonicity properties) there is also an agreed-upon criterion of teasing apart the two classes, one of its recent formalizations can be found in Giannakidou & Zeijlstra (2017) – see (3), their example (16).[3] The criterion is partially meaning based and partially relies on context felicity of n-words. Its working will be exemplified in the following sections.

(3)   X qualifies as an n-word iff:

    a. X can be used with structures with sentential negation or other X with meaning equivalent to one ¬; and

    b. X provides a negative fragment answer.

In this article I discuss mainly experimental evidence from Czech which allows us to answer a research question: do languages like Czech (where the evidence to differentiate between n-words and NPIs is very limited) distinguish between n-words and NPIs (particularly the class of NPIs called strong NPIs)? Why is Czech (and generally strict negative concord languages) a good data source for finding differences between strong NPIs and n-words? Because even if the introduced distinction between syntactically licensed n-words and semantically licensed NPIs is supported by many researchers today (Zwarts 1998; Zeijlstra 2004 and Gajewski 2011 among others), there are very influential theories which subsume n-words under NPIs (Ladusaw 1992) or observe close relationship of the two classes (Laka 1990): in such theories the distinction between syntactic licensing (n-words) and semantic licensing (NPIs) of course disappears. The question of nature (if any … depending on the theory) of the distinction between n-words and NPIs is theoretically still open and empirically is especially vexing in strict negative concord languages because there the environment where a speaker can get positive evidence about the distinction between n-words and NPIs boils down to neg-raising contexts. This is the reason of centrality of neg-raising for NPI debate – see further §2 and §3.1.

To foreshadow the experiments discussed in much bigger detail later, let us consider the following set of Czech sentences (items from one of the experi-

---

[3]Beside n-words and their cross-linguistic variation with respect to the strictness of negative concord, there is also a variation in NPIs: while generally NPIs are bad as negative fragment answers, one particular subtype of them, minimizers provide felicitous fragment answers, see Giannakidou (1998) and Błaszczak (2001) for further details. Thanks to an anonymous reviewer for pressing these points about n-words and NPIs licensing variation.

ments): if asked about grammaticality of such sentences, Czech native speakers would consider (4a) ungrammatical, (4b) perfectly acceptable, (4c) good and (4d) and (4e) bad to some extent. Such graded judgments of sentences containing (what I will argue further to be) strong NPIs, in concreto graded acceptability of strong NPIs depending on the presence of negation and/or the type of embedding verb and some other factors was the original motivation for running the series of experiments resulting in the current article. It is important to notice that there is variation among speakers, variation caused by lexical items used in the tested sentences, etc. This naturally calls for an experimental verification because relying on a researcher's intuitions in such cases can lead to totally conflicting claims: e.g. Bošković & Gajewski (2011) state non-existence of neg-raising in Slavic languages, while Dočekal & Dotlačil (2016a) defend limited existence of neg-raising in Czech. The experimental data and their careful analysis – in the light of current formal semantic theories – allow me to avoid such contradicting claims and eventually isolate the relevant factors behind NPI licensing and an interaction of the licensing with other syntax-semantics phenomena as neg-raising, etc.[4]

(4) a. * Zmizela ani   jedna knížka.
          Lost   not.even one   book
          'A single book is missing.'

     b. Nezmizela ani   jedna knížka.
         NEG.lost   not.even one   book
         'Not a single book is missing.'

     c. Náš nový knihovník nechce,   aby zmizela ani   jedna
         our new librarian NEG.wants COMP lost   not.even one
         knížka.
         book
         'Our new librarian doesn't want even one book to be missing.'

     d. Náš nový knihovník si nepředstavuje, že   zmizela ani   jedna
         our new librarian SE NEG.imagine COMP lost   not.even one
         knížka.
         book
         'Our new librarian doesn't imagine that even one book is missing.'

     e. Náš nový knihovník neslyšel, že   zmizela ani   jedna
         our new librarian NEG.heard COMP lost   not.even one

---

[4]Thanks to an anonymous reviewer for raising importance of this general background question to me.

knížka.
book
'Our new librarian didn't hear that even one book was missing.'

The article is organized as follows: in the first, more theoretically based part (§2) I will illustrate the empirical criteria distinguishing n-words and NPIs, then I will tease apart so called weak NPIs from strong NPIs and lastly I will introduce the basic observations about Czech and negative dependent expression. §3.1, §3.2, and §3.3 will be more of the experimental linguistic character, they are heavily based on the joint work with Jakub Dotlačil (partially reported in Dočekal & Dotlačil 2016a,b; 2017). In concreto, I will report the experimental evidence for distinguishing n-words from NPIs stemming from three classes of phenomena: (i) neg-Raising contexts; (ii) fragment answers; (iii) likelihood manipulated contexts. The nature of this article is more overview-like, the details about statistics, design of the experiments, etc. can be found in Dočekal & Dotlačil (2016a,b; 2017); Dočekal & Šafratová (2018).

2 NPIs vs. n-words: Theory

2.1 N-words

Let us start with introducing some important pieces of linguistic knowledge concerning n-words, the expressions which are generally taken as syntactically dependent on negation and which are different both from negative quantifiers on the one hand and from NPIs on the other hand.

N-words crucially differ from Germanic negative quantifiers as the following contrast in (5) shows: English verbal negation and a negative quantifier in (5a) yield only a double negation reading while the word for word translation of (5a) into Czech with the n-word *nikoho* 'anybody' and verbal negation in (5b) is interpretable only with one negation scoping wide over the whole sentence as is clear from the predicate logic formalization. Generally, n-words are syntactically dependent expressions which occur only in languages where some form of negative concord is attested.

(5)   a. John didn't see nobody.   (English)
$\neg \exists x [\text{PERSON}(x) \land \neg \text{SEE}(\text{JOHN}, x)]$

        b. John nikoho neviděl.
John nobody NEG.saw
'John didn't see anybody.'   (Czech)
$\neg \exists x [\text{PERSON}(x) \land \text{SEE}(\text{JOHN}, x)]$

The distinction between n-words and NPIs already mentioned in the criterion in 3 is illustrated in (6): (6b) illustrates the unavailability of NPIs as fragment answers versus the perfect acceptability of n-words in the same context in (6d) – the Czech translation of the (6a) – (6b) mini-dialogue.

(6)   *NPIs ≠ n-words:*
   a.   Whom did you talk to?
   b.   * Anybody. / Nobody.
   c.   S   kým   jsi mluvil?
         with whom AUX spoke?
         'With whom did you speak?'
   d.   S   nikým.
         with nobody
         'With nobody.'

There are at least two influential theories of n-words: the first one treats n-words as non-negative indefinites (predicate of type $\langle e, t \rangle$) which are required to be in the scope of clause-mate negation (so called roofing requirement from Ladusaw (1992), see Giannakidou (1997) for an historical overview). The second type of theory compares n-words to agreement markers which nicely explains their locality requirements, basically their need to be licensed syntactically by clause-mate negation. The second type of approach is developed in Zeijlstra (2004) and Zeijlstra (2008). In this article I will follow the syntactic agreement approach even if nothing hinges too much on the particular framework as far as it constrains the distribution of n-words to clauses with overt verbal negation. This locality constraint is one of the usually mentioned contrasts between n-words and NPIs since unlike NPIs which just need to be in a scope of negative element, n-words need a local negation as the following contrast from Giannakidou & Zeijlstra (2017) shows.

(7)   Dhen prodhosa   mistika [pu eksethesan [kanenan /*KANENAN]].
       not   betrayed.1SG secrets that exposed.3PL anybody   N-BODY
       'I didn't reveal secrets that exposed anybody.'
                                           (Greek; Giannakidou & Zeijlstra 2017: 18)

It should be noted that the locality requirement of n-words varies across languages but for n-words in Slavic languages the locality requirement is very strict as observed already by Progovac (1993). So unlike in Spanish, Italian or Greek

where the licensing of n-words sometimes (especially in case of subjunctive embedding) can span from a negation on the root verb to n-words in embedded clauses, such licensing is ungrammatical in Slavic languages, see the following examples from Czech.

(8)   a.   * Petr neřekl,   že   nikdo   přišel.
          Petr NEG.said that N-BODY came
          Intended: 'Petr didn't say that anybody came.'

      b.   Petr řekl, že   nikdo   nepřišel.
          Petr said that N-BODY NEG.came
          'Peter said that nobody came.'

      c.   * Petr nechce,   aby   tu   nikdo   byl.
          Petr NEG.wants COMP.SBJV here N-BODY were
          Intended: 'Petr doesn't want anybody to be here.'

      d.   Petr chce, aby   tu   nikdo   nebyl.
          Petr wants COMP.SBJV here N-BODY NEG.were
          'Peter wants nobody to be here.'

2.2 NPIs

A prototypical example of an NPI is the English expression *any* – see the seminal work of Kadmon & Landman (1993) (and there for older references). If an NPI occurs in a sentence without negation it results in an ungrammatical sentence – (9). If it occurs in a negated sentence like in (10), the only interpretation is a scope of *any* under negation: (10a) vs. the unavailable interpretation in (10b). In English a quantifier *few students* (which shares with negation the relevant property of downward entailingness – discussed shortly) licenses NPIs in the object: (11).

(9)   * Peter visited anyone.

(10)   Petr didn't visit anyone.
      a.   Available: $\neg\exists x[\text{PERSON}(x) \land \text{VISIT}(\text{PETER}, x)]$
      b.   Unavailable: $\exists x[\text{PERSON}(x) \land \neg\text{VISIT}(\text{PETER}, x)]$

(11)   Few students visited anyone.

Next, negation is not the only expression licensing NPIs which (at least in the case of so called weak NPIs) sets NPIs apart from n-words which are licensed only by negation. Compare the following Czech paradigm in (12) where the

NPI/minimizer *sebemenší šance* 'slightest chance' contrasts with the adjectival n-word *žádnou* (glossed as N-ADJ). The NPI licensing expression in (12a) is the quantifier *málo studentů* 'few students'. The negation and other NPI licensing expressions share the property of reversing the direction of entailment in their argument. Notice how negation reverses entailment in Table 1: logical conjunction entails logical disjunction in a positive case but negated logical disjunction entails negated logical conjunction – notice the tautological status of both formulas in Table 1. Because of the entailment reversion property of NPI licensors their essential quality is called downward entailing (DE) and is generally accepted by scholars as the most probable common denominator of NPI environments (since Ladusaw 1992 at least).[5]

(12) a. Málo studentů mělo sebemenší šanci   složit   tu   zkoušku.
          few   students had slightest chance to.pass the exam
          'Few students had the slightest chance to pass the exam.'

  b. # Málo studentů mělo žádnou šanci   složit   tu   zkoušku.
          few   students had N-ADJ chance to.pass the exam
          Intended: 'Few students had any chance to pass the exam.'

Table 1: Entailment properties of conjunction and disjunction

p	q	$(p \wedge q) \rightarrow (p \vee q)$	$\neg(p \vee q) \rightarrow \neg(p \wedge q)$
1	1	1	1
0	1	1	1
1	0	1	1
0	0	1	1

In natural language the reasoning of monotonicity is frequently applied in relation to sets, subsets and supersets. Notice the predicate logic implications in

---

[5]I will discuss in more detail the distinction between weak and strong NPIs. In the literature there are various attempts to reclassify the landscape of NPIs, one of them Rullmann 1996, following the work of Krifka 1995 and further elaborated in Lahiri 1998 points out that there is a special class of NPIs – in Rullmann's terms *even*-NPIs (*ook maar*-series in Dutch) which seem to be an indefinite incorporated with the semantics of the scalar focus particle *even*. *Even*-NPIs are inherently scalar and interact with focus. Czech *ani*-NPIs have precisely the characteristics of *even*-NPIs as I will discuss later. In this respect they belong to the same class (*even*-NPIs) as stressed English *any* and the *ook maar*-series in Dutch. Thanks to an anonymous reviewer for pointing out the importance of this cross-linguistic comparison.

(13) which corresponds to the patterns from the propositional calculus. If there is some x in the intersection of P and Q denotation then necessarily there is an x in P and Q union (13a). And if there is no x in P and Q union, then there cannot be any x in their intersection (13b). So in a affirmative sentence (in predicate logic non-negated formula) the entailment goes from an subset (intersection) to a superset (union) while in a negated sentence, the entailment is reversed and proceeds from a superset (union) to its subset (intersection). A natural language example is in (14): the denotation of NP *red wine* is a subset of the NP *wine* denotation and in an affirmative sentence (14a) the entailment is from a subset to a superset, not vice versa: (14b). In a negated sentence the entailment reverses: (14c).

(13)   a. $\exists x[P(x) \wedge Q(x)] \rightarrow \exists x[P(x) \vee Q(x)]$
        b. $\neg\exists x[P(x) \vee Q(x)] \rightarrow \neg\exists x[P(x) \wedge Q(x)]$

(14)   red wine $\rightarrow$ wine
        a. John likes red wine. $\rightarrow$ John likes wine.
        b. John likes wine. $\nrightarrow$ John likes red wine.
        c. John does not like wine. $\rightarrow$ John does not like red wine.

The general condition stating that NPIs occur in downward entailing (DE) environments can be stated like (15), from (von Fintel 1999: 100).[6]

(15)   Fauconnier–Ladusaw's Licensing Condition: An NPI is only grammatical if it is in the scope of an α such that $[\![\alpha]\!]$ is DE.

The downward monotonic and upward monotonic reasoning in case of quantifiers works like this: upward monotonic quantifiers allow reasoning from subsets to supersets while downward monotonic quantifiers from supersets to subsets: (16). Natural language examples of upward, downward and non-monotonic quantification are presented in (17).

(16)   a. DET A is upward entailing iff for any B, C $(B \subseteq C)$ DET $A\,B \Rightarrow$ DET $A\,C$

---

[6]The licensing condition has to be understood as necessary, not sufficient: there are cases of intervention in NPIs licensing (see Linebarger 1987 for an early treatment and Homer 2008 for a more recent approach), then cases of NPIs being unacceptable even in simple negative sentences (see Uribe-Echevarria 1994 and Błaszczak 2001). But as none of the experiments reported further addresses such type of data, for the purposes of this paper I stick to (15), as a working definition of NPI licensing.

b. DET A is downward entailing iff for any B, C ($B \subseteq C$)
   DET $A\,C \Rightarrow$ DET $A\,B$

c. if not upward or downward monotonic → non-monotonic

(17) Upward/Downward entailing and non-monotonic determiners:
   a. *some:* Some toys are blue $\Rightarrow$ Some toys are colored
   b. *few:* Few toys are colored $\Rightarrow$ Few toys are blue
   c. *exactly n:* Exactly three toys are blue $\Leftrightarrow$ Exactly three toys are colored

It is important to notice that monotonicity properties belong to a position in a sentence and they are computed compositionally: so a position in a sentence can be upward entailing even if it occurs in the scope of a downward entailing quantifier. In (18) the object position is in the scope of two DE quantifiers and consequently is upward monotonic, as the validity of the entailment pattern shows.[7]

(18) a. [↓ At most three detectives arrested ↓[fewer than four ↑[criminals]]]
   b. →[↓ At most three detectives arrested ↓[fewer than four ↑[humans]]]

2.3 Weak and strong NPIs

There is a class of NPIs, so called WEAK NPIs with prototypical English examples like *any* or *ever*. Weak NPIs occur in all downward entailing environments as illustrated in (19).

(19) a. Bill didn't ever say anything.
   b. No student ever said anything.
   c. Few students ever said anything.
   d. At most 5 students ever said anything.
   e. *Between 5 and 10 students ever said anything.
   f. *{Some/all/most} students ever said anything.

The second class of NPIs instantiated by English expressions like *in weeks*, additive *either*, and punctual *until* are so called STRONG NPIs and as the name suggests,

---

[7]Early discussion of this compositionality which can lead to flip-flop effects in NPIs licensing can be found in Baker (1970), a recent study incorporating some experimental findings is Geurts & van Der Slik (2005).

they occur only in a subset of environments where weak NPIs are grammatical as illustrated in (20).[8]

(20) a.   Bill didn't leave until his birthday.
      b.   No student left until his birthday.
      c.   * Few students left until their birthdays.
      d.   * At most 5 students left until their birthdays.
      e.   * Between 5 and 10 students left until their birthdays.
      f.   * {Some/most/all} students left until their birthdays.

The logical property which licensors of strong NPIs share (negation and *no* in (20)) is a strengthened form of entailment reversal and usually is named anti-additivity.[9] In using anti-additivity as the necessary condition for strong NPI acceptability I follow seminal work of Zwarts (1998). There is a popular alternative explanation of strong NPIs and their behavior in Gajewski (2011) which describes their stricter distribution via downward entailing properties but checked both in at-issue meaning and in the presupposition/implicature part of the meaning. I will stick to the classic theory of anti-additivity here: the definition is in (21). (22) illustrates the anti-aditivity (the quantifier *no* is anti-additive since negation is always anti-additive as is clear from deMorgan's law: $\neg(p \lor q) \leftrightarrow (\neg p \land \neg q)$). But DE quantifiers like *few students* in (23) are not anti-additive – imagine a scenario

---

[8]The interaction of strong NPIs and locality is a vast topic but notice the following pattern from Romoli (2013: 317):

   (i)   a.   John doesn't think that Mary will arrive until tomorrow.
         b.   * John isn't certain that Mary will arrive until tomorrow.

As the pattern shows, licensing of strong NPIs is always possible in case of negated neg-raising predicates like *think* but results in ungrammaticality in cases of negated non-neg-raisers as the predicate like *be certain*.

[9]The full hierarchy of negative strength is the following one: anti-morphicity > anti-additivity > downward entailing (anti-morphicity defined after Krifka 1995: an operator O is anti-morphic iff: $O(\neg X) = \neg O(X)$; negation is anti-morphic unlike English negative quantifier *no* as can be seen from the following equivalence and non-equivalence *John wasn't happy = It's not the case that John was happy* vs. *No student wasn't happy ≠ It's not the case that no student was happy*). Being the strongest negative expression (like verbal negation) entails being classified as anti-additive and downward entailing automatically. Strong NPIs are usually taken to be licensed by operators of at least anti-additive strength – see the grammaticality of (20b). In Slavic languages (strict negative concord) it is not that easy to tease apart anti-additivity and anti-morphicity of negative NPs headed by *no* but it seems that strong NPIs in Slavic require at least anti-additive licensors as well.

with 10 students, three of them drinking and three of them smoking, then ∨ part of (23) is false while ∧ part of (23) is true.[10]

(21)   Anti-additive function: $F(x \lor y) \leftrightarrow F(x) \land F(y)$

(22)   No student smokes or drinks ↔ No student smokes and no student drinks.

(23)   Few students smoke or drink ↮ Few students smoke and few students drink

2.4 NPIs vs. n-words

Returning now to the broader question of distinguishing between NPIs (negative dependent expressions licensed in semantics via notions like monotonicity and/or anti-additivity) and n-words (negative dependent expressions licensed in syntax via agreement), it is acknowledged that such a distinction corresponds nicely with a well established modularity architecture of a grammar where usually we distinguish between different forms of well-formedness such as syntactic or semantic, corresponding to well-formedness which is located in different modules of grammar. But the picture is not so clear when we consider recent theories of NPI licensing where the logical properties correlate with syntactic acceptability of NPIs. In concreto: if ungrammaticality of NPIs in upward entailing environments is due to lack of the right monotonicity properties in them, then we are in fact linking the domains of semantics with syntax. And in some theories (Heim 1984; Crnič 2014) of NPIs licensing where the licensing of NPIs is postulated via presupposition the linking goes even further: between the licensing in pragmatics with syntactic acceptability. Recent theories of NPIs (Chierchia 2013) and strong NPIs (Gajewski 2011) seem to point in the same direction.

Before we move to the experimental part of the article, let us have an outlook of Czech data scrutinized in much more detail in the series of experiments I will report. In Czech there are two candidates both at first sight reasonable for the NPI

---

[10]Corresponding to the full scale of negative strength – see the previous footnote – some researchers like Krifka (1995) and Van der Wouden (2002) distinguish weak (licensed in downward entailing contexts), strong (licensed in anti-additive contexts) and super-strong NPIs (licensed in anti-morphic contexts). Due to the strict negative concord properties of Slavic languages (discussed in the previous footnote too) I will stick to the basic dichotomy: weak/strong NPIs where strong NPIs would subsume the strong and the super-strong NPIs from the more nuanced classifications. Thanks to an anonymous reviewer for pointing out the importance of this issue.

or n-word status: *ani (jeden)* 'not even (one)' and *žádný* 'N-ADJ'. As the following example demonstrates, both require clause-mate negation in basic cases, so both can be thought of as either n-words or strong NPIs (the embedded clauses of communicative verbs can be shown to be non-anti-additive: details to follow).

(24) a. Petr neviděl {ani   jednoho / žádného} studenta.
          Petr NEG.saw   even one   N-ADJ   student
          'Petr didn't see any student.'

      b. {*Ani   jeden / *žádný} student přišel.
            even one   N-ADJ student came
          'Not even one/any student came.'

      c. Petr neslyšel, že   {*ani   jeden / *žádný} student přišel.
          Petr NEG.heard that   even one   N-ADJ student came
          'Petr didn't hear that even one/any student came.'

So it is well conceivable that four logical possibilities of classifying *ani jeden* 'not even one' and *žádný* 'N-ADJ' are reasonable. Czech tradition like Havránek et al. (1960) can be interpreted as Table 2 suggests, so basically treating both types of expressions as syntactically dependent on negation.

Table 2: Czech traditional grammar on *ani* vs. *žádný*

item/profile	NPIs	n-words
ani jeden	✗	✓
žádný	✗	✓

And as it is clear from the previous discussion, the division between n-words and strong NPIs is subtle – the only other clause-mate environment (next to negation) which passes the test of anti-additivity are prepositions like English *without* (compare the equivalence of: *John left the pub without paying and saying good bye* ↔ *John left the pub without paying or John left the pub without saying good bye*). So it is reasonable to ask a research question like (25). Neg-concord languages like Czech (and generally all Slavic languages) do employ negative dependency on negation via n-words, so is there a reason for a language to maintain a set of expressions which does nearly the same job but is licensed in semantics? In the rest of the article I will argue for the positive answer to the question: the experimental evidence clearly shows that *ani (jeden)* 'not even (one)' expressions pattern like strong NPIs, not like n-words, while *žádný* 'N-ADJ' are n-words.

(25) Research question: do strict neg-concord languages even allow grammaticalization of strong NPIs?

3 Experimental evidence

In the three following sections I will discuss the experimental evidence which allows us to tease apart n-words from NPIs. First in §3.1 I will report evidence coming from the behavior of NPIs in neg-raising constructions: NEG-RAISING (NR) is a primarily interpretational phenomenon where a negation of verbs like *think, believe* or *want* is most saliently understood as scoping over their embedded verb (*I don't want to leave* ≈ *I want not to leave*, compared with a lack of such interpretation in case of non-NR predicates: *I don't say I will leave* ≉ *I am saying that I will not leave*). In §3.2 the evidence for distinguishing between n-words and NPIs will come from their different acceptability as fragmentary answers to questions. And in §3.3, the two classes will be shown to behave differently with respect to their entailment and likelihood properties.

3.1 Neg-raising

Because NPIs are licensed in the semantic part of the grammar engine, they are (ceteris paribus) expected to be able to be licensed at long distance. N-words as syntactically dependent on negation have to obey strict locality conditions unlike NPIs. And even more importantly, if the licensing of NPIs happens in semantics, their licensing should be sensitive to properties of their embedding verbs, in case of NR-predicates, NPIs should appear in the embedded clauses of NR-predicates but are predicted to be unacceptable in the embedded clauses of non-NR predicates (verbs of communication or causation). If we construe such long distance licensing, the expected pattern should look like the one in Table 3: n-words cannot be licensed across a clausal boundary, while NPIs can be licensed from their embedding clause. Nevertheless in case of NR-predicates, we expect a sharp difference between predicates like *want* or *believe* and non-NR predicates like *hear, say,* or *force*.

  The experimental results which bear on this issue are summarized in more detail in Dočekal & Dotlačil (2016a). Let us call this experiment Experiment 1. Experiment 1 consisted of 5 conditions demonstrated in (26), one of the items of the experiment. The experiment tested acceptability of sentences containing NPIs and focused on neg-raising.

Table 3: n-words vs. NPIs in Neg-raising and non-NR contexts

environment/status	NPIs	n-words
NR embedded	✓	✗
non-NR embedded	✗	✗

(26) a. Ztratila se ani   jedna ovce.
          Lost   SE not.even one   sheep
          'A single sheep is missing.'

      b. Neztratila se ani   jedna ovce.
          NEG.lost   SE not.even one   sheep
          'Not a single sheep is missing.'

      c. Nový bača   v Tatrách nechce,   aby se ztratila ani   jedna
          new shepherd in Tatras NEG.wants COMP SE lost   not.even one
          ovce.
          sheep.
          'The new shepherd in Tatras doesn't want even one sheep to be
          missing.'

      d. Nový bača   v Tatrách si nemyslí, že   se ztratila ani
          new shepherd in Tatras SE NEG.think COMP SE lost   not.even
          jedna ovce.
          one   sheep
          'The new shepherd in Tatras doesn't think that even one sheep is
          missing.'

      e. Nový bača   v Tatrách neříká, že   se ztratila ani   jedna
          new shepherd in Tatras NEG.say COMP SE lost   not.even one
          ovce.
          sheep
          'The new shepherd in Tatras doesn't say that even one sheep is
          missing.'

The sentences represent 5 environments listed below:

(A) an affirmative sentence

(B) a negative sentence

(C) a clause embedded under negated NR predicates of intention and judgement/obligation (e.g. *want, advise*)

(D) a clause embedded under negated NR predicates of opinion (*believe*)

(E) non-NR predicates

Experiment 1 tested only NPIs: *ani jeden* was one two NPIs in it, the second one *až do* 'until' + time expression is not important for this article, n-words were not tested. The descriptive statistics of Experiment 1 is visualized in Figure 1: the x-axis represents the 5 conditions and the y-axis represents the 5-point Likert scale (1 = the least acceptable, 5 = the most acceptable). The boxplots summarize the acceptability in the usual manner. As is evident from the graph, Condition A was the least acceptable, Condition B most accepted, all other conditions somewhere in the interval between the two extremes. The most important difference for this article is the one between the conditions C and D and E where E represents non-NR predicates and was perceived as less acceptable by native speakers. This seems to be result of unlicensed NPI in the embedded clauses of non-NR predicates. I discuss the design of Experiment 1 in more detail here because the following two experiments (viz §3.2 and §3.3) have an analogical design and statistical modeling. When describing the following two experiments, I will be less eloquent.
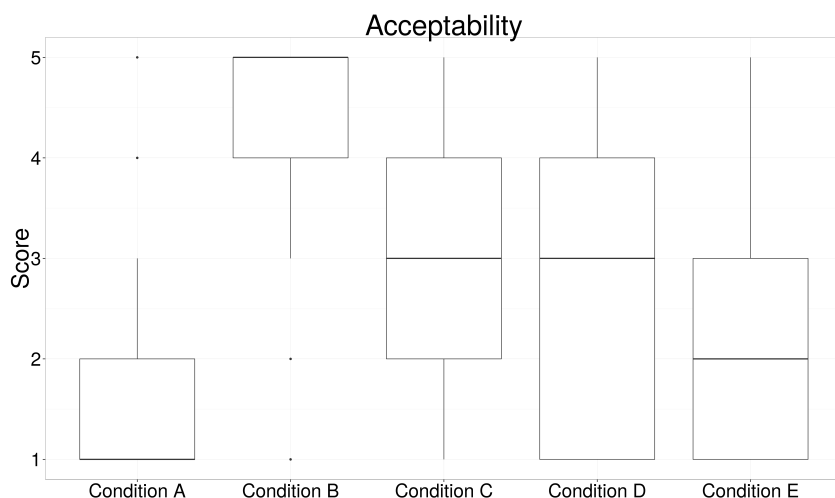

Figure 1: Results of Experiment 1

The results of the experiment can be theoretically explained in the scalar approach to NR (Horn 1973; Romoli 2012; 2013). In the scalar theory of neg-raising NR predicates (beside the assertion – (27a)) contribute the excluded middle (EM) implicature to the semantic composition (27b). And finally the alternatives generated by the implicature are exhaustified by ᴇxʜ – (28).

(27)   a. $[\![\text{NR}]\!] = \lambda p \lambda x . \Box_x[p]$
       b.   ᴀʟᴛ$([\![\text{NR}]\!]) = \{\lambda p \lambda x . \Box_x[p], \lambda p \lambda x . [\Box_x[p] \vee \Box_x[\neg p]]\}$

(28)   ᴇxʜ$(\text{ᴀʟᴛ}(p))(p)(w) = p(w) \wedge \forall q \in \text{ᴇxᴄʟ}(p, \text{ᴀʟᴛ}(p))[\neg q(w)]$

I will illustrate the mechanics of the scalar theory of NR on an example item from Experiment 1: (29). Formula in (30a) shows the alternatives generated by the excluded middle implicature from (27b): it is the negated at-issue meaning ($\neg\text{wᴀɴᴛ}_s[p]$) and the excluded middle part ($\neg(\text{wᴀɴᴛ}_s[p] \vee \text{wᴀɴᴛ}_s[\neg p])$). The excluded middle in this case formalizes the involvement of the subject *s*: he either wants the proposition *p*, or he wants the negation of *p* but he cannot be uninterested with respect to *p*. The excluded middle for other classes of NR-predicates has an analogous meaning: opinionatedness for *know/believe*, clear intentions for *plan*, etc. Compare the lack of such an excluded middle meaning in predicates of communication: a speaker can say *p* or neg *p* but he can be silent about *p* as well. (30b) then shows the exhaustification of the alternatives: the at-issue meaning remains the same but the excluded alternative is negated – the usual strenghtening of the sentence meaning via negating its alternatives. The at-issue meaning and double negated excluded middle alternative then (via deductive reasoning) yield the semantic low scope of negation in the embedded proposition. So, as a consequence of exhaustification of the NR predicate and its excluded middle implicature, the negation is of the NR predicate is interpreted as having low scope (semantically).

(29)   'A new shepherd in Tatra mountains doesn't want even one sheep to be missing.'
       $\neg\text{wᴀɴᴛ}_s[p]$

(30)   a.   ᴀʟᴛ$(\neg\text{wᴀɴᴛ}_s[p]) = \{\neg\text{wᴀɴᴛ}_s[p], \neg(\text{wᴀɴᴛ}_s[p] \vee \text{wᴀɴᴛ}_s[\neg p])\}$
       b.   ᴇxʜ$(\neg\text{wᴀɴᴛ}_s[p]) = \neg\text{wᴀɴᴛ}_s[p] \wedge \neg\neg(\text{wᴀɴᴛ}_s[p] \vee \text{wᴀɴᴛ}_s[\neg p]) \models$
            $\text{wᴀɴᴛ}_s[\neg p]$

Let us recall that strong NPIs are licensed by anti-additive functions: functions which obey deMorgan's laws which naturally is true for negation: a natural

language example is presented in (31a) and (31b) where the entailment is bidi-rectional and in propositional logic in (31c) and (31d) where the same meaning equivalence holds.

(31) a. It didn't rain and it didn't snow.

  b. It didn't rain or snow.

  c. $\neg p \wedge \neg q$

  d. $\neg[p \vee q]$

In the case of NR predicates like *want* in (32) the embedded clause qualifies as an anti-additive environment due to the NR-transfer of negation: (32a) is equivalent to (32b) – both require p and q being false in all possible worlds – see Table 4 with an example of two possible worlds. In such a model both logical formulas in (32c) and (32d) are true.

(32) a. Susan does not want to sleep and she does not want to dance.

  b. Susan does not want to sleep or dance.

  c. $\square \neg p \wedge \square \neg q \leftrightarrow$

  d. $\square \neg (p \vee q)$

Table 4: A fragment of possible worlds for (32)

world/proposition	p	q
w_1	0	0
w_2	0	0

But consider an example of non-NR predicates like *say* in (33a) and (33b). (33b) does not follow from (33a) since non-NR predicates if negated allow only the high scope of negation interpretation: (34a) – and such an interpretation is the following: it requires there to be at least some possible worlds where the propositions p and q are false. But (34a) is stronger: it requires both propositions p and q to be false in all possible worlds. (34a) would be true in a valuation of propositions across possible worlds in Table 5 but (34b) would be false in such a model. In other words: non-NR predicates do not create anti-additive environment in their embedded clauses. And since strong NPIs need anti-additivity, they are unlicensed in the embedded clauses of non-NR predicates.

(33) a. Susan didn't say that she will sleep and she didn't say that she will dance.
      b. Susan didn't say that she will sleep or dance.

(34) a. $\neg\Box p \wedge \neg\Box q$ (true in the table)
      b. $\neg\Box[p \vee q]$ (false in the table)

Table 5: A fragment of possible worlds for (33a)/(33b)

world/proposition	p	q
w_1	0	1
w_2	1	0

Returning now to the initial predictions: Experiment 1 confirmed the NPI status of *ani (jeden)* – if *ani (jeden)* were an n-word, the contrast between NR predicates (*ani (jeden)* licensed) and non-NR predicates (*ani jeden* not acceptable) would be unexplained since syntactic licensing should not be sensitive to semantic distinctions between anti-additive and non-anti-additive environments. So we can conclude this section with a first clear experimental confirmation of classifying *ani (jeden)* as a strong NPI. Moreover it was established that anti-additivity is a necessary condition for licensing the strong NPI *ani (jeden)*. Experiment 1 itself did not establish contrast between strong NPIs (*ani (jeden)*) and n-words but its results would be unexpected if *ani (jeden)* were not a strong NPI. Experiment 1 did not test intuitions for *žádný*, the reason for that is the following one: *žádný* is perceived by native Czech native speakers to be grammatical only if it appears in a sentence with local negation (26b) type of sentences. So unlike in case of *ani (jeden)* where the judgments are much more graded, there is no need to experimentally establish the acceptability of *žádný*.

Table 6: N-words vs. NPIs in Neg-raising environments

environment/status	NPIs	n-words
NR embedded	✓	✗
non-NR embedded	✗	✗

3.2 Fragment answers

Another distinction mentioned already in criterion 3 is the distinction between n-words and NPIs with respect to their ability to be fragmentary answers to questions. Roughly, n-words are good fragmentary answers, while NPIs are generally not acceptable as fragmentary answers. Similarly to the situation in NR contexts reported in the last section, the acceptability of *ani (jeden)* as a fragmentary answer seems to be more varied than in case of n-words which are always good as fragment answers. I pre-experimentally noticed that especially in cases where the question supplies more context, the NPIs seem more acceptable, following the pattern in (35). The fragment answers were tested in two expriments; first in Experiment 2 the fragment answers were tested against minimal context questions.

(35) a. Kdo byl dneska večer   na náměstí?
          who was today   evening on square?
          'Who was today in the evening on the square?'

      b. ? Ani jeden člověk.
          NPI one   human
          'Not even one man.'

      c. Kdo tu   dneska byl?
          who here today   was
          'Who was here today?'

      d. ??? Ani jeden člověk.
          NPI one   human
          'Not even one man.'

In Experiment 2 (details can be found in Dočekal & Dotlačil 2017), there was a negative interaction of *ani* and ellipsis in non-negative questions like (36). In other words, as expected n-words were judged by speakers as better fragmentary answers than NPIs. The statistical outcome is visualized in Figure 2 – the relevant condition is ELLIPSIS and blue bar for n-words, red for NPIs.

(36) Kdo odešel z   hospody?
      who left   from pub?
      'Who left the pub?'

      a. Žádný student.
          N-ADJ student
          'No student.'

b. ?? Ani jeden student.
    NPI one   student
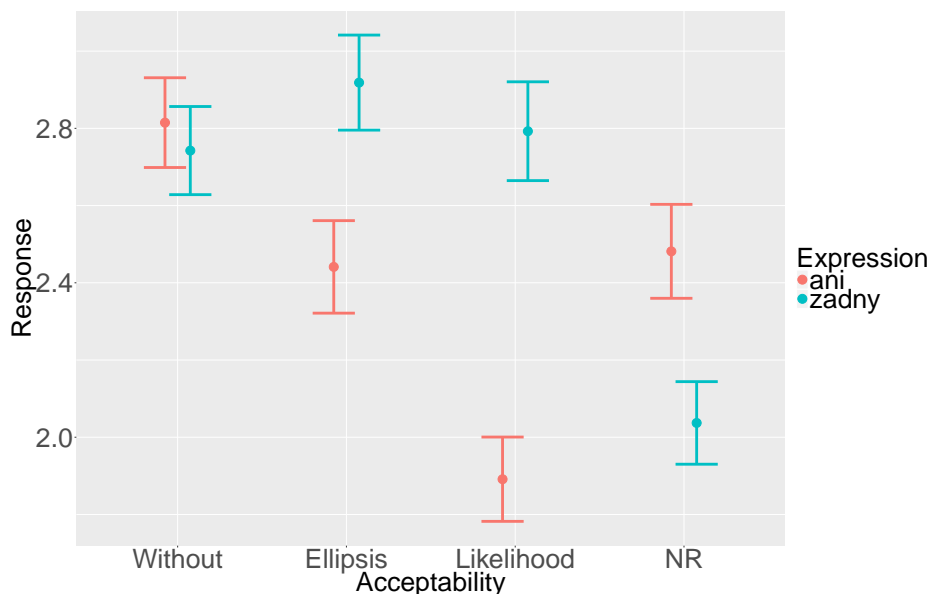   'Not even one student.'

Figure 2: Results of Experiment 2

The theoretical explanation of this known difference is usually provided via a possible reconstruction of n-words and unavailability of reconstruction for NPIs. Because NPIs are usually not able to reconstruct under a possible licensor in their scope (De Swart 1998) like in the following example where NPI *any* student in the cleft cannot reconstruct to its base object position under the quantifier *no professor* which would license it.[11]

(37)   *It is any student that no professor likes.

We further elaborated the fragment answer distinction in Experiment 3 (details can be found in Dočekal & Dotlačil 2017) where we provided more contextual

---

[11] Again the ban on NPI reconstruction can be side-stepped with a carefully constructed example as the following sentence from Uribe-Echevarria (1994: p.17) shows: *A doctor who knew anything about acupuncture was not available.* It seems though that in such cases it is the whole subject NP (containing the NPI) reconstruction which saves grammaticality of NPI and this type of construction seems to be highly restricted. Nevertheless thanks to an anonymous reviewer for pointing this out.

informations like in the example item (38). In this experiment the correlation disappeared: see Figure 3 – conditions FRAGNPI vs. FRAGNWORD with no difference in acceptability.

(38) Koho   vyhodil profesor Palný včera   ze   zkoušky?
      whom fired   prof   Palný yesterday from exam?
      'Who was fired by prof Palný during yesterday's exam?'

   a. Žádného studenta.
       N-ADJ   student
       'No student.'

   b. Ani jednoho studenta.
       NPI one   student
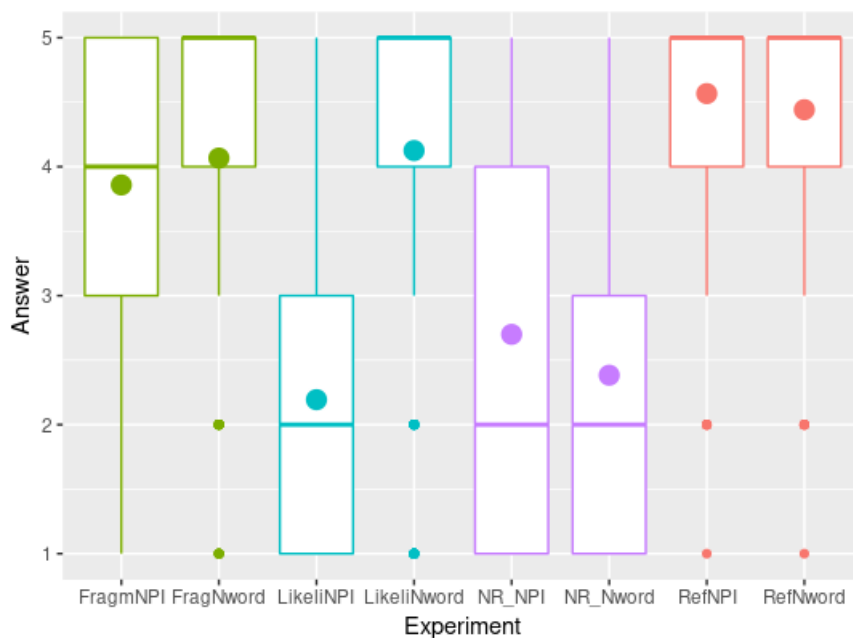   'Not even one student.'

Figure 3: Results of Experiment 3

The ability of n-words to appear as fragmentary answers is usually taken as the standard distinction of n-words against NPIs. But in a recent paper Fălăuş & Nicolae (2016) observe a strikingly related phenomen: the authors claim (based

on data from many strict neg-concord languages) that in strict neg-concord lan-
guages n-word answers to negative questions can have (surprisingly) a Double
Negation (DN) reading. This observation goes against the n-words vs. NPIs cri-
terion as it falsifies the meaning part of it: n-words and (reconstructed) negation
yield only one semantic negation. I checked Fălăuș & Nicolae's (2016) claims with
10 native speakers of Czech and they seem to be valid – see example (39): there
seems to be even a preference (8/10) for the DN reading – (39a) but the negative
concord reading (39b) is considered to be possible (for 2 out of 10 speakers).

(39) Kdo nepřečetl žádný článek?
      who neg.read N-ADJ article
      'Who didn't read any article?'

    a. Nikdo.
       Nobody. (None of us read a single one.) NC (2/10):
       $\neg\exists x, y[\text{PERSON}(x) \land \text{ARTICLE}(y) \land \text{READ}(x, y)]$
       $\equiv \forall x[\text{PERSON}(x) \rightarrow \neg\exists y[\text{ARTICLE}(y) \land \text{READ}(x, y)]]$

    b. Nikdo.
       Nobody. (Each of us read an article) DN (8/10):
       $\neg\exists x[\text{PERSON}(x) \land \text{ARTICLE}(y) \land \neg\text{READ}(x, y)]$
       $\equiv \forall x[\text{PERSON}(x) \rightarrow \exists y[\text{ARTICLE}(y) \land \text{READ}(x, y)]]$

Fălăuș & Nicolae (2016) solve the availability of DN reading of n-words via pos-
tulating another (focus-related) position for covert negation (CN): in the left pe-
riphery of a clause as in the tree in Figure 4. The position is according to Fălăuș
& Nicolae (2016) licensed via n-word movement to the left peripheral position
above TP. A negation in the left periphery is a second negation in a sentence,
next to the reconstructed negation from the question (surface negation, SN). So
the first negation in (39b) is the interpretation of covert negation, the second one
of the verbal negation. If we follow Fălăuș & Nicolae (2016), we can explain the
puzzling disappearance of contrast between n-words and NPIs (Experiment 3)
as a consequence of the covert negation – if such a negation appears in a clause,
the NPIs are licensed because they do not need to reconstruct under the scope
of verbal negation and then the contrast between n-words and NPIs disappears.
There are many questions raised by postulating such covert negation, especially
with respect to possible over-generation – at the end n-words in strict negative
concord languages cannot appear in sentences without negation but postulating
covert negation leaves this robust observation unexplained. Fălăuș & Nicolae
(2016) try to resolve such problems by restricting the covert negation only to
controllable set of cases, all somehow related to focus movement of n-words to
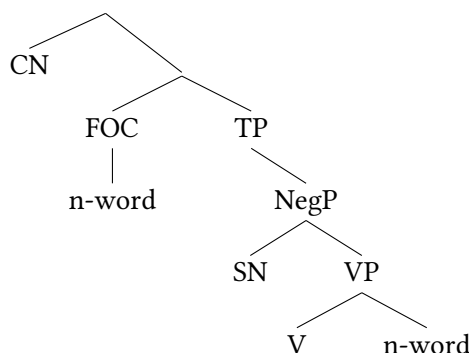
Figure 4: Covert negation, syntax

the left periphery. I tried to verify their claims and conducted a small survey
again with the same 10 speakers of Czech and it seems that Fălăuş & Nicolae's
general idea is confirmed with an interesting twist. Let us start with a basic case
– (40) is interpreted only with NC reading as is visible from the ranking in (40a)
and (40b) – a double negation reading is simply non-existent.

(40) Nikdo   ničemu nevěří.
      n-person n-thing neg.believes
      'Nobody believes anything.'
      a.   NC: $\forall x[\textsc{person}(x) \rightarrow \neg\exists y[\textsc{entity}(y) \wedge \textsc{believe}(x, y)]]$
      b. * DN: $\forall x[\textsc{person}(x) \rightarrow \exists y[\textsc{entity}(y) \wedge \textsc{believe}(x, y)]]$

But in case of information structure manipulation like in (41), which is even an
affirmative sentence, the double negation reading surprisingly emerges. A simi-
lar pattern is observed in (42). The sentences moreover seem to have the double
negation reading only. This confirms Fălăuş & Nicolae's hypothesis about focus
position of the CN: example (40), where there is no object movement to the left
periphery (unlike in (41) and (42)), has only the expected NC reading. In this ar-
ticle it is not possible to explore more details of this interesting appearance of
double negation reading in a negative concord language like Czech but more im-
portantly: it seems to be reasonable to postulate another position for negation
in the left periphery of a clause, such a position (because it is somehow licensed
via focus) can then blur the picture of the fragmentary answer criterion and the
fluctuation of acceptability of NPIs as fragmentary answers observed in Experi-
ment 3 is no longer a mystery, context manipulation can lead to a focus related
CN licensing of even strong NPIs as fragment answers.

(41) V nic   nikdo   nevěří.
   in N-THING N-PERSON believes
   'Nobody believes in anything.'

   a. * NC (0/10): $\forall x[\text{PERSON}(x) \rightarrow \neg \exists y[\text{ENTITY}(y) \land \text{BELIEVE}(x, y)]]$
   b.   DN (10/10): $\forall x[\text{PERSON}(x) \rightarrow \exists y[\text{ENTITY}(y) \land \text{BELIEVE}(x, y)]]$

(42) Nic   při té zkoušce nikdo   nenapsal.
   N-THING at the exam   N-PERSON NEG.wrote
   'Nobody wrote anything during the exam.'

   a. * NC (0/10): $\forall x[\text{PERSON}(x) \rightarrow \neg \exists y[\text{ENTITY}(y) \land \text{WRITE}(x, y)]]$
   b.   DN (10/10): $\forall x[\text{PERSON}(x) \rightarrow \exists y[\text{ENTITY}(y) \land \text{WRITE}(x, y)]]$

Summary of this section: there seems to be some evidence for classifying *ani* as an NPI and *žádný* as an n-word which stems from the fragment answer experiments. When the results diverge from the expected dichotomy, there seems to be a reasonable explanation via postulation of a second covert negation in the sentence.

3.3 Likelihood scenarios

The last environment discussed in this article concerns the semantic properties of sentences where n-words vs. NPIs occur. The straightforward predictions are the following:

   1) n-words (licensed in syntax) should not be sensitive to logical properties of their environment (they require just sentential/verbal negation)

   2) NPIs are licensed in semantics and by definition are dependent on semantic properties like DE, anti-additivity, etc.

I will pursue the line of distinguishing NPIs from n-words via the NPI sensitivity to monotonicity and likelihood. And I will base my reasoning on a very influential theory of NPI licensing, the so called simple *even* hypothesis of NPI licensing (Heim 1984; Krifka 1995; Crnič 2014 – I will call the theory Heim/Crnič theory further). The theory describes NPIs using the following three ingredients:

   • NPIs associate with covert *even* – the formalization can be via a formal [*even*] feature carried by the NPIs, etc.

   • NPIs (like focused element) generate sets of possible alternatives;

- covert *even* associates with the alternatives and generates presupposition of its prejacent being the least probable member of the set of alternatives (entailing all the alternatives) – in case of association with *even* (some authors suggest different covert licensors of NPIs too);

The immediate predictions of the Heim/Crnič theory is that NPIs should be sensitive to probability and entailing properties. The first and the second one are logically related: a proposition p cannot be more likely than a proposition q, if p entails q: intuitive illustration – p being *Rambo killed 100 enemies*, q being *Rambo killed 99 enemies*, p entails q and p is less likely than q; q does not entail p and is more likely than q – see Crnič (2011) for details of relating entailing and likelihood. The theoretical intricacies away, the prediction that NPIs should be sensitive to logical properties like entailing or probability while n-word not is uncontroversial, see Table 7 for a visualization of these predictions.

Table 7: N-words vs. NPIs in probability manipulated environments

property/item	entailment/probability
n-words	✗
NPIs	✓

And exactly this prediction was tested in Experiment 2 and Experiment 3. In both we found a strong correlation of *ani* and probability. As a side note: a corpus survey (the biggest national Czech corpus, Křen et al. (2015)) confirms the likelihood sensitivity of *ani* – a prototypical example in (43) shows that *ani* usually associates with weak scalar items (*ani jeden* is the second most frequent collocation, the first one another minimizer *ani slovo* 'not a single word'). which via scalar reasoning entails all other scalar alternatives ($\neg \exists X[\text{CUSTOMER}(X) \wedge \#X = 1 \wedge \text{ENTER}(X)] \rightarrow \neg \exists X[\text{CUSTOMER}(X) \wedge \#X > 1 \wedge \text{ENTER}(X)]$). And due to this entailment the sentence with *ani* and a weak element associated with *ani* is the least probable (entailing all other alternatives).

(43) tento nyní úspěšný   podnikatel   […] v prvním měsíci neměl   [ani
      this now succesfull businessman   in first   month NEG.had NPI
      jednoho zákazníka]
      one   customer
      'This currently succesfull businessman did not have even one customer in
      the first month.'

In Experiment 2 the acceptability of *ani* with strong scalar items was tested – example item in (44) where the scale of catholic hierarchy is most probably ⟨*priest, bishop, cardinal*⟩ – *cardinal* being high scalar item in any case. The scale entails contextual (not proper formal logical) entailment due to the facts of world we know the following implicational hierarchy: $\exists x[\textsc{become cardinal}(x) \rightarrow \textsc{become bishop}(x) \rightarrow \textsc{become priest}(x)]$ and its reversal as invalid: $\exists x [\textsc{become priest}(x) \nrightarrow \textsc{become bishop}(x) \nrightarrow \textsc{become cardinal}(x)]$. To acquire the grade of cardinal entails acquiring (ceteris paribus) acquiring all lower ranks of catholic hierarchy but not the other way round. The scalar item *cardinal* is the strongest (in the ad hoc scale), it entails all other items in the scale and is consequently least likely (which fits the natural intuitions). If *ani* prefers weak scalar items, it should be degraded with strong items, while n-words (as they are not picky about semantic environments) should be more acceptable.

(44) […] nestal   se {ani / žádným} kardinálem
      NEG.became SE NPI   N-ADJ   cardinal
    'He didn't become even a cardinal.'

And we found out that people overall preferred *žádný* (n-word) with strong scalar items. The reason is that n-words do not have semantic requirements unlike NPIs: *ani* prefers weak scalar items. The statistical results of Experiment 2 are in Figure 2, the pertinent condition LIKELIHOOD: *ani* (red) had mean acceptability very much below the n-word's mean acceptability (blue) (around 2.8 for n-words).

Experiment 3 was partially an elaboration of Experiment 2 – while Experiment 2 used an acceptability task, in Experiment 3 the truth value judgment task was used in case of testing likelihood properties of *ani*. An example item is in (45). Again it was tested how much worse is the acceptability of strong scalar items with *ani*. In this scenario the scale is ⟨PhD, MA, BA⟩: here the scale is contextually based on the likelihood of passing the exam (if the scale were based simply on academic hierarchy, as in the acceptability testing in (44), it would be ⟨PhD, MA, BA⟩ but in (45) the scale is reversed as passing the exam is prototypically negatively correlated with the academic rank). The scale is (due to the context) again based on contextual entailment: $\forall x[\textsc{BA}(x) \rightarrow \textsc{pass}(x)] \rightarrow [\forall x[\textsc{MA}(x) \rightarrow \textsc{pass}(x)] \rightarrow \forall x[\textsc{PhD}(x) \rightarrow \textsc{pass}(x)]]$. Therefore *ani* associates again with the strongest scalar item (in the positive version of a tested sentence entailing all its scalar alternatives). And as the statistical summary in Figure 3 shows (the relevant condition LIKELI_NPI s. LIKELI_Nword – blue color), speakers again preferred n-words to *ani* NPIs. This again follows from *ani*'s semantic requirements (it associates with weak items which in negative contexts become

least likely among alternative scalar items) vs. n-words which do not have any semantic sensitivity and are therefore more acceptable than *ani*.

(45) Scenario: prof. Novák yesterday examined an easy course which BA, MA and PhD students attend. PhD students pass the exam always, MA in most cases but BA only rarely. Včerejší   zkoušku u prof. Nováka
                                yesterday exam   at prof. Novák
nesložili   {ani / žádní} bakaláři.
NEG.passed NPI   N-ADJ BA-students
'No bachelors passed the yesterday's exam by prof. Novák.'

Empirically both experiments strongly support the classification of *ani* as an NPI which associates with weak scalar items and *žádný* as an n-word licensed in the syntax (and consequently without any particular semantic sensitivity).

Table 8: *Ani* vs. *žádný* in probability manipulated environments

property/item	probability/entailment
žádný	✗
ani	✓

The theoretical explanation of *ani* being an NPI which obligatorily selects weak scalar items can be the following. The first thing to note is that the facts observed in the experiments are only a piece of a bigger pattern where *ani* competes in some environments with another scalar particle *i* 'even'. In a recent experiment (Dočekal & Šafratová 2018) it was confirmed that *i* obligatorily selects strong scalar items, while *ani* weak items. Illustrated on a data pattern close to the catholic hierarchy from Experiment 2 Czech native speakers are prone to the following judgments (where * should be understood as total unacceptability in experiments, ?? as in-between-acceptability and ✓ as nearly total acceptability – statistic noise away – but of course only in case the judgments are related to the set up scale, catholic hierarchy in (46).

(46) a. *Upward entailing contexts:*
          i. Petr se nakonec stal   {✓i   kardinálem / ??i   knězem}.
              Petr SE at-end   became   even cardinal   even priest
              'Petr in the end became { even a cardinal / even a priest }.'

      ii. Petr se nakonec stal   {*ani   kardinálem / *ani
          Petr SE at-the-end became   not.even cardinal   not.even
          knězem}.
          priest.
          'Petr in the end didn't become even a cardinal.'

    b. *Downward entailing, non anti-additive contexts:*

      i. Jestli se Petr stal   {✓i   kardinálem / ??i   knězem}, tak…
         If   SE Petr became   even cardinal   even priest   then
         'If Peter became even a cardinal, then …'

      ii. Jestli se Petr stal   *ani   kardinálem, tak …
         If   SE Petr became   not.even cardinal,   then
         'If Peter didn't become even a cardinal, then …'

    c. *Downward entailing, anti-additive contexts:*

      i. Petr se nakonec nestal   {*i   kardinálem / *i
         Petr SE at-the-end NEG.become   even cardinal   even
         knězem}.
         priest
         'Petr didn't become even a cardinal at the end.'

      ii. Petr se nakonec nestal   {??ani   kardinálem /
         Petr SE at-the-end NEG.become   not.even cardinal
         ✓ani   knězem}.
           not.even priest
         'Petr didn't become even a cardinal at the end.'

The pattern we observe is the following: *i* in upward entailing contexts and downward entailing contexts prefers strong elements on a scale but it is unacceptable with weak or strong scalar items in anti-additive contexts; *ani* prefers weak scalar items in anti-additive contexts but it is unacceptable in upward entailing contexts with both weak and strong scalar items (and in simple DE contexts). Such a pattern is explainable (following the logic of argumentation in Crnič 2011) as *i* and *ani* spelling out the following features:

(47)   a. *i* … [EVEN]
        b. *ani* … [EVEN,AA]

The feature [EVEN] requires the association with covert *even* defined below in (48) following Crnič (2014) among many others. The feature [AA] requires the item to occur in an anti-additive environment. The items form a scale in (49) and

compete for insertion via the usual Maximize presupposition principle which requires the speaker to make her contribution presupposing as much as possible (for the original formulation see Heim 1991).

(48) $[\![\text{even}]\!]^w(C)(p)$ is defined only if $\forall q \in C[q \neq p \rightarrow q >_{\text{LIKELY}} p]$

(49) $\langle i, ani \rangle$

The observed distribution of *i*/*ani* and their strong/weak association is explainable as follows:

1. Upward entailing environments: *i* is licit but only with strong scalar items as then the *even* presupposition is satisfied, *ani* cannot be inserted as UE environments clash with *ani* [AA] feature.

2. Downward entailing environments: *i* is licit with *even* scoping below the DE operator: [→ [[even C] antecedent … *i* …] consequent], *ani* cannot be used due to the [AA] feature requirement.

3. Anti-additive environments: *i* cannot be inserted because Maximize Presupposition dictates the insertion of the most specific item (*ani* in this case), *ani* associates with weak scalar items: the scope [even C] [¬ …*ani* …].

4. The association of *i*/*ani* with 'wrong' scalar items is perceived as bad (??) but not totally ungrammatical – weak scalar item for *i* in upward entailing contexts and strong scalar items for *ani* in anti-additive environments.

The last point seems to point to the existence of possible reversed scoping: [even C][→ [antecedent …*i* …] consequent] for *i* and [¬ [even C] …*ani* …] for *ani* which would explain their allowed (even if not preferred) 'crossed' association. But as it was confirmed by Experiment 2 and Experiment 3 *ani* associates with weak items, while *i* with strong scalar items (see Dočekal & Šafratová 2018 for details) by default. This default scope exchange of *i*/*ani* which happens exactly in anti-additive contexts (*i* prefers strong elements, *ani* weak elements but only in the scope of negation – negation being the anti-additive licensor in 99%) reveals their unified semantics where the flip-flop is a consequence of entailment/likelihood reversal caused by the negation. The only difference between *i* and *ani* is the formal feature [AA] which formalizes the morphological incorporation of negation into *ani*. It would be possible to encode the scope differences via different features ([SOLO] of Crnič 2011 for the weak elements) but such a move would miss the nice competition pattern which emerged from the data: namely *i* is in principle

expected in anti-additive environments but cannot be inserted as a consequence of *ani* being more specific ([EVEN,AA]).

Summary of this section: *ani (jeden)* 'not even (one)' behaves like a strong NPI – this behavior was confirmed by Experiment 2 and Experiment 3 where association with strong scalar items was sanctioned (against relatively acceptable n-words modifying strong scalar items). Furthermore, *ani* competes with *i* – the former prefers strong scalar items which was experimentally confirmed too. The association with weak scalar items and competition with *i* would be unexpected if *ani* were n-word.

3.4 Summary

Let us end this article by answering the question asked at the beginning: do n-words and strong NPIs co-exist in natural language? And if yes (in some languages like English they do co-exist for sure), does this distinction hold even in strict neg-concord languages where the boundary between strong NPIs and n-words is even more subtle? The experiments, their results and their theoretical interpretation described in this article bring very strong support of the existence of both classes of negatively dependent expressions even in a strict neg-concord language like Czech. This result allows us to maintain the standard assumptions concerning n-words (they are licensed syntactically) and NPIs (they are licensed in semantics/pragmatics). More importantly, the data patterns of Czech NPIs seem to strongly favor the NPI theories which base their licensing on concepts like anti-additivity and likelihood (Zwarts 1998 in the first case, Heim 1984 and Crnič 2014 in the second). Another issue touched in this article is unreliability of our intuitions: it seems that distinguishing between n-words and strong NPIs has to be based on such subtle data which can only be obtained by experimental methods. The subtlety of judgments can explain differing stances on this distinction in the previous literature where such opposing views as: n-words are a subclass of NPIs (Ladusaw 1992, Fălăuş & Nicolae 2016 a.o.) versus n-words are a separate class (Zeijlstra 2008 and Giannakidou & Zeijlstra 2017 a.o.) were maintained. There is another pertinent question raised by the data: do all speakers agree with respect to the distinction between n-words and strong NPIs? And if no, is there a real dialectal variation or at least some correlation? The results of the experiments in fact bear direct evidence on this fascinating question but the space of this article is alas filled completely.

Abbreviations

N-PERSON	n-word for persons	SG	singular number
N-THING	n-word for things	COMP	complementizer
NEG	negation	AUX	auxiliary verb
NPI	negative polarity item	N-ADJ	n-word for properties
SBJV	subjunctive	AA	anti-additive
DE	downward entailing	SE	reflexive clitic
PL	plural number	NR	Neg-raising

Acknowledgements

References

Baker, C Lee. 1970. Double negatives. *Linguistic Inquiry* 1(2). 169–186. DOI:10.1080/08351816909389104

Błaszczak, Joanna. 2001. *Investigation into the interaction between the indefinites and negation.* Berlin: Akademie Verlag.

Bošković, Željko & Jon R. Gajewski. 2011. Semantic correlates of the NP/DP parameter. In Suzi Lima, Kevin Mullin & Brian Smith (eds.), *NELS 39: Proceedings of the 39th Annual Meeting of the North East Linguistic Society*, 121–134. Amherst, MA: GLSA Publications.

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention.* Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199697977.001.0001

Crnič, Luka. 2011. *Getting even.* Cambridge, MA: MIT. (Doctoral dissertation).

Crnič, Luka. 2014. Against a dogma on NPI licensing. In *The art and craft of semantics: A festschrift for Irene Heim*, vol. 1, 117–145. Cambridge, MA: MITWPL.

De Swart, Henriëtte. 1998. Licensing of negative polarity items under inverse scope. *Lingua* 105(3-4). 175–200. DOI:10.1016/S0024-3841(98)00021-7

Dočekal, Mojmír & Jakub Dotlačil. 2016a. Experimental evidence for neg-raising in Slavic. *Linguistica* 56(1). 93–109.   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 21*, 387–399. Edinburgh: University of Edinburgh. https : / / sites . google . com / site / sinnundbedeutung21/.

Dočekal, Mojmír & Jakub Dotlačil. 2017. *Strong NPIs vs. n-words: Acceptability experiment in Czech.* Paper presented at Sinn und Bedeutung 22, Berlin. https://sinnundbedeutung22.wordpress.com/.

Dočekal, Mojmír & Iveta Šafratová. 2018. *Even* hypothesis of PPIs licensing: An experimental study. In Joseph Emonds, Markéta Janebová & Ludmila Veselovská (eds.), *Language use and linguistic structure: Proceedings of the Olomouc Linguistics Colloquium 2018*, 253–273. Olomouc: Palacký University.

Fălăuş, Anamaria & Andreea Nicolae. 2016. Fragment answers and double negation in strict negative concord languages. In Mary Moroney, Carol Rose Little, Jacob Collard & Dan Burgdorf (eds.), *SALT 26: Proceedings from the 26th Conference on Semantics and Linguistic Theory*, 584–600. Washington, DC: Linguistic Society of America.   DOI:10.3765/salt.v26i0.3813

Gajewski, Jon R. 2011. Licensing strong NPIs. *Natural Language Semantics* 19(2). 109–148.   DOI:10.1007/s11050-010-9067-1

Geurts, Bart & Frans van Der Slik. 2005. Monotonicity and processing load. *Journal of Semantics* 22(1). 97–117.   DOI:10.1093/jos/ffh018

Giannakidou, Anastasia. 1997. *The landscape of polarity items.* Groningen: University of Groningen. (Doctoral dissertation).

Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency.* Amsterdam: John Benjamins.

Giannakidou, Anastasia & Hedde Zeijlstra. 2017. The landscape of negative dependencies: Negative concord and n-words. In *The Wiley Blackwell companion to syntax, Second edition.* New York: Wiley–Blackwell. DOI:10.1002/9781118358733.wbsyncom102

Havránek, Bohuslav et al. 1960. *Slovník spisovného jazyka českého.* Praha: Nakl. Československé akademie věd.

Heim, Irene. 1984. A note on negative polarity and downward entailingness. In Charles Jones & Peter Sells (eds.), *NELS 14: Proceedings of the 14th Annual Meeting of the North East Linguistic Society*, 98–107. Amherst, MA: GLSA Publications.

Heim, Irene. 1991. Articles and definiteness. In Arnim von Stechow & Dieter Wunderlich (eds.), 487–535. Berlin: de Gruyter.   DOI:10.1515/9783110126969.7.487

Homer, Vincent. 2008. Disruption of NPI licensing: The case of presuppositions. In Tova Friedman & Satoshi Ito (eds.), *SALT 18: Proceedings of the 18th Semantics and Linguistic Theory*, 429–446. Ithaca, NY: CLC Publications. DOI:10.3765/salt.v18i0.2483

Horn, Laurence. 1973. *On the semantic properties of logical operators in English.* Los Angeles, CA: University of California. (Doctoral dissertation).

Kadmon, Nirit & Fred Landman. 1993. *Any. Linguistics and Philosophy* 16(4). 353–422.   DOI:10.1007/BF00985272

Křen, Michal, Václav Cvrček, Tomáš Čapka, Anna Čermáková, Milena Hnátková, Lucie Chlumská, Tomáš Jelínek, Dominika Kováříková, Vladimír Petkevič, Pavel Procházka, Hana Skoumalová, Michal Škrabal, Petr Truneček, Pavel Vondřička & Adrian Zasina. 2015. *SYN2015: Reprezentativní korpus psané češtiny.* http://www.korpus.cz. Praha: Ústav Českého národního korpusu, FF UK.

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25(3–4). 209–257.   DOI:10.3765/salt.v0i0.2462

Ladusaw, William A. 1992. Expressing negation. In Chris Barker & David Dowty (eds.), *SALT 2: Proceedings from the 2nd Conference on Semantics and Linguistic Theory*, 237–260. Columbus, OH: Ohio State University Press. 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

Laka, Itziar. 1990. *Negation in syntax: On the nature of functional categories and projections.* Cambridge, MA: MIT. (Doctoral dissertation).

Linebarger, Marcia C. 1987. Negative polarity and grammatical representation. *Linguistics and Philosophy* 10(3). 325–387.   DOI:10.1007/BF00584131

Progovac, Ljiljana. 1993. Negative polarity: Entailment and binding. *Linguistics and Philosophy* 16(2). 149–180.   DOI:10.1007/BF00985178

Romoli, Jacopo. 2012. *Soft but strong. neg-raising, soft triggers, and exhaustification.* Cambridge, MA: Harvard University. (Doctoral dissertation).

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. 1996. Two types of negative polarity items. In Kiyomi Kusumoto (ed.), *NELS 26: Proceedings of the 26th Annual Meeting of the North East Linguistic Society*, 335–351. Amherst, MA: GLSA Publications.

Uribe-Echevarria, Maria. 1994. *Interface licensing conditions on negative polarity items: A theory of polarity and tense interactions*. Storrs, CT: University of Connecticut. (Doctoral dissertation).

Van der Wouden, Ton. 2002. *Negative contexts: Collocation, polarity and multiple negation*. London: Routledge.

von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148.

Zeijlstra, Hedde. 2004. *Sentential negation and negative concord*. Amsterdam: University of Amsterdam. (Doctoral dissertation).

Zeijlstra, Hedde. 2008. *Negative concord is syntactic agreement*. Manuscript, University of Amsterdam. https://ling.auf.net/lingbuzz/000645.

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 3

Whom to oblige?

Julie Goncharov

Hebrew University of Jerusalem / University of Göttingen

In this paper, I argue that deontic modals can be relativized on a par with epistemic modals, contrary to what is generally believed (see, for instance, Rett 2016). The evidence comes from the behaviour of Russian deontic modals under negation. In Russian, these constructions have an aspectual restriction: they are well-formed with the imperfective, but not with the perfective aspect. This restriction, however, can be circumvented when the modal is relativized to the addressee rather than the subject. This obviation effect shows that the relativization of the deontic modal base (i.e. whose obligations are relevant) is not a function of the context, but is rather encoded in the grammar. The analysis is implemented within the grammaticalized speech act system (Speas & Tenny 2003, Wiltschko 2017, a.o.). The account proposed here is extended to imperatives providing support for the presence of a deontic component in imperatives (Han 1999; Ninan 2005; Kaufmann 2012 a.o.). I also discuss cross-Slavic variation when it comes to the aspectual restriction with deontic modals and imperatives.

Keywords: necessity modals, deontic modals, negation, aspect, Russian

1 Two core observations

We begin straightaway with examining the two observations that are the focus of this paper. The first observation is that negated strong deontic modals in Russian cannot be used with the perfective aspect (aspectual restriction). The second observation is that the aspectual restriction can be circumvented when the modal is relativized to the addressee rather than the subject. I use these observations to argue that: (i) deontic modals can be relativized on a par with epistemic modals (contra Rett 2016) and (ii) the relativization takes place in the grammar rather than being a function of the context.

1.1 The aspectual restriction

Russian strong deontic modals, like *dolžen* 'must', can be used with a verb in either IMPERFECTIVE (IPFV) or PERFECTIVE (PFV) in positive sentences, (1a). There are minimal interpretative differences between IPFV and PFV in (1a) due to aspect, which will not concern us here. The important observation is that under negation PFV is unavailable, (1b).[1] This aspectual restriction on Russian strong deontic modals under negation has been widely discussed in the literature, e.g. Forsyth (1970); Rappaport (1985); de Haan (1997); Zalizniak (2006); Paducheva (2013).[2]

(1) a. Ivan dolžen   {uxodit'   / ujti}.
       Ivan must.PTCP leave.IPFV   leave.PFV
       'Ivan must leave.'

    b. Ivan ne dolžen   {uxodit'   / *ujti}.
       Ivan not must.PTCP leave.IPFV   leave.PFV
       'Ivan doesn't have to leave.'

The observation in (1) is also true for other strong deontic modals and modal expressions in Russian, (2)–(3). However, for reasons of space, I limit the presentation of data and discussion to *dolžen*.

(2) a. Ivanu   nužno   {uxodit'   / ujti}.
       Ivan.DAT need.PTCP leave.IPFV leave.PFV
       'Ivan needs to leave.'

    b. Ivanu   ne nužno   {uxodit'   / *ujti}.
       Ivan.DAT not need.PTCP leave.IPFV leave.PFV
       'Ivan doesn't need to leave.'

(3) a. Ivan objazan   {?uxodit'   / ujti.}
       Ivan obliged.PTCP leave.IPFV leave.PFV
       'Ivan is obliged to leave.'

    b. Ivan ne objazan   {uxodit'   / *ujti}.
       Ivan not obliged.PTCP leave.IPFV   leave.PFV
       'Ivan is not obliged to leave.'

---

[1]Note that (1b) with PFV can have an epistemic reading irrelevant here. Unless marked otherwise, all modal bases in this paper are interpreted as deontic. I gloss *dolžen* as 'must' because like English *must* it can have both deontic and epistemic interpretations. However, in negative sentences, I translate *dolžen* as 'have to' because this translation better represents the fact that *dolžen* scopes below negation, see the discussion around example (4).

[2]The aspectual restriction applies only to strong deontic modals, i.e. modals that have a universal force. Existential (weak) deontic modals do not show the aspectual restriction. This difference is predicted by the analysis proposed in this paper, see the discussion around example (30).

It is important to note that in (1b) with ɪᴘғᴠ repeated in (4a) the deontic modal is interpreted below negation (i.e. there is no obligation for Ivan to leave), (4b). The interpretation where the deontic is interpreted above negation (i.e. Ivan has an obligation to stay/not to leave) is not available, (4c). When the verb is ᴘғᴠ in (1b), the deontic cannot be interpreted either below or above negation.

(4) a. Ivan ne dolžen   uxodit'.
        Ivan not must.ᴘᴛᴄᴘ leave.ɪᴘғᴠ
        'Ivan doesn't have to leave.'

    b. Available scope: 'There is no obligation for Ivan to leave.'
        Abbreviated as ✓ ¬ □ ɪᴘғᴠ

    c. Unavailable scope: 'Ivan has an obligation to stay/not to leave.'
        Abbreviated as ✗ □ ¬ ɪᴘғᴠ

(5) summarizes the first pattern that we need to account for. As can be seen in (5), there is only one available configuration where the verb is in ɪᴘғᴠ and the modal is interpreted below negation, (5a). Three other configurations are not well-formed, (5b)–(5d).

(5) Pattern 1 to explain (the aspectual restriction)

    a. ✓ ¬ □ ɪᴘғᴠ
    b. ✗ □ ¬ ɪᴘғᴠ
    c. ✗ ¬ □ ᴘғᴠ
    d. ✗ □ ¬ ᴘғᴠ

1.2 Relativization and obviation

Our second core observation is that the aspectual restriction is lifted when obligations are set on the addressee rather than the subject. In an out-of-the-blue context in examples (1)–(4), the obligations are set on the subject (= Ivan). Consider now the context in (6), in which police officers are ordered to stop Ivan from leaving.[3] In (6), the subject (and the agent) of leaving (= Ivan) does not hold any obligations. The obligations to stop Ivan from leaving are on the addressee (= police officers). In this context, ᴘғᴠ under negated deontic modals becomes available.[4]

---

[3]The interpretation of deontics in contexts like (6) is similar to that of imperatives. The connection between deontics and imperatives is discussed in §5.

[4]To the best of my knowledge, this observation has not been discussed in published work, although it seems to be common place for Russian speakers.

(6) *Situation:* Police arrive at a crime scene and see Ivan fleeing with the stolen *Mona Lisa.* A police chief to police officers:

Ivan ne dolžen   ujti!
Ivan not must.PTCP leave.PFV
'Ivan must not leave/escape.'

The relativization of the deontic modal base can also be illustrated using conjunctive sentences in which the second conjunct denies the prejacent. As discussed in Ninan (2005), strong deontic modals, (7a), unlike weak ones, (7b), are infelicitous when the prejacent of the modal is negated in the second conjunct.

(7) a. # Sam must go to confession, but he's not going to.
     b. Sam should/ought to go to confession, but he's not going to.

(Ninan 2005: 2)

The deviant (7a) has the following form: *Sam must p and ¬Sam is going to p.* Suppose that in (6) the modal can only be interpreted above negation (see below). If the modal base is relativized to Ivan, the configuration *Ivan must ¬p and Ivan is going to p* should be as infelicitous as (7a). But this is not the case:[5]

(8) In the situation described in (6):
Ivan ne dolžen   ujti,   xotja   on i sobiraetsja.
Ivan not must.PTCP leave.PFV although he FOC going.to
'Ivan must not escape, although he is going to.'

However, the counterpart of (8), in which the second conjunct denies that the police are going to retain Ivan (stop him from leaving) is deviant, (9). This suggests that *dolžen* in cases like (6), (8), and (9) is relativized to the addressee (= police officers) rather than the subject (= Ivan).[6]

(9) # Ivan ne dolžen   ujti,   no my ne sobiraemsja ego zaderživat'.
     Ivan not must.PTCP leave.PFV but we not going   him retain.INF
     'Ivan must not escape, but we are not going to retain him.'

---

[5]I thank an anonymous reviewer for pointing out problems with this example in the earlier version of the paper.

[6]The interpretation of (6), (8), and (9) is more involved. It resembles that of coercion constructions such as *The addressee must bring it about that Ivan doesn't leave.* This, however, does not mean that they are structurally different from non-relativized constructions as in (1). Grano (2017) has an informative discussion of coercion-free semantics for similar constructions. For reasons of space, I do not elaborate on this point here. However, I believe that something along the lines proposed in Grano (2017) can be adopted for Russian facts with *dolžen.*

With respect to scopal possibilities, it is important to note that cases like (6), (8), and (9) have only one reading, in which the modal is interpreted above negation, (10b). The surface scope reading, in which the modal is below negation, is not available, (10c).[7]

(10)   a.   In the situation described in (6):
          Ivan ne dolžen   ujti!
          Ivan not must.PTCP leave.PFV
          'Ivan must not leave/escape.'

     b.   Available scope: 'You, officers, must make Ivan not leave/stay.'
          Abbreviated as ✓ $\Box^{ADR} \neg$ PFV^{SBJ}

     c.   Unavailable scope: 'You, officers, don't have to make Ivan leave.'
          Abbreviated as ✗ $\neg \Box^{ADR}$ PFV^{SBJ}

The second pattern to be accounted for in this paper is summarized in (11). This pattern concerns only configurations with the verb in perfective. The summary shows that the only possible construal is (11d), in which the deontic is relativized to the addressee and scopes above negation.

(11)   Pattern 2 to explain (relativization and obviation)

     a.   ✗ $\neg \Box^{SBJ}$ PFV^{SBJ}
     b.   ✗ $\Box^{SBJ} \neg$ PFV^{SBJ}
     c.   ✗ $\neg \Box^{ADR}$ PFV^{SBJ}
     d.   ✓ $\Box^{ADR} \neg$ PFV^{SBJ}

2 What we already know

That deontic modality interacts with negation and aspect has been reported in the literature and these interactions have received some accounts (e.g. Han 1999; Ninan 2005; Iatridou & Zeijlstra 2013). In this section, I briefly show that the Russian facts discussed in §1 are not reducible to previously reported observations. The Russian facts constitute a superset; thus, previous analyses under-generate and their extension to the Russian data is hopeless.

---

[7]Here and below, I use superscripts ADR = addressee and SBJ = subject to mark whose obligations are relevant and who is the agent of the action described by the verb.

2.1 Deontics and negation

It has been claimed that some deontic modals in Dutch, English, German, Greek, and Spanish exhibit the behaviour characteristic of POSITIVE POLARITY ITEMS (PPIs), e.g. Homer (2011); Iatridou & Zeijlstra (2010; 2013). The examples in (12) illustrate PPI-like properties of English *must* under its deontic reading. (12a) shows that *must* can only be interpreted above negation. (12b) shows the rescuing effect when *must* appears under two (Strawson) downward entailing operators. (12c) is an intervention effect: *must* can be interpreted under negation when a universal quantifier intervenes.

(12) a. John mustn't leave. ✗ ¬ □ / ✓ □ ¬
   b. Only John must not work tonight. only ¬ □
   c. A student's mistake mustn't necessarily be hurled on the shoulders of his teachers. ¬ ∀ □

   (Iatridou & Zeijlstra 2013: 543, 539)

Russian *dolžen*, however, cannot be assimilated to PPI deontic modals such as English *must*. As we saw above, *dolžen* can (in fact, must) scope below negation when the verb is imperfective. The deviance under negation arises only when the prejacent has a perfective verb. The relevant examples are repeated in (13):

(13) a. Ivan ne dolžen uxodit'.
   Ivan not must.PTCP leave.IPFV
   'Ivan doesn't have to leave.' ✓ ¬ □ / ✗ □ ¬
   b. *Ivan ne dolžen ujti.
   Ivan not must.PTCP leave.PFV
   Intended: 'Ivan doesn't have to leave.' ✗ ¬ □ / ✗ □ ¬

2.2 Deontics and aspect

It has been also discussed in the literature (Han 1999; Ninan 2005; Hellie 2016, a.o.) that the deontic reading of *must* is unavailable when the verb is perfective, (14). This is attributed to the future-oriented nature of the deontic *must*, which is corroborated by examples like (15).

(14) John must$_{\text{epist/*deon}}$ have left.

(15) a. You must finish your homework tonight/now.
   b. # You must have finished your homework yesterday.   (Han 1999: 485)

Be as it may, this line of reasoning cannot be extended to the Russian data. As we already saw, in positive sentences, *dolžen* can have the deontic interpretation when the verb is perfective. The relevant example is repeated in (16):

(16) Ivan dolžen   ujti.
      Ivan must.PTCP leave.PFV
      'Ivan must$_{\text{epist/deon}}$ leave.'

To sum up, in this section we saw that the interaction between deontic modals and negation as well as the interaction between deontic modals and the perfective aspect, has already been discussed in the literature. The Russian data, however, are not reducible to either of these interactions. We should, therefore, seek an answer elsewhere.

3 Towards an account

3.1 Preliminaries

We saw in the previous section that accounts that explain the PPI-like behaviour of deontic modals and the unavailability of the deontic reading with the perfective aspect cannot be extended to the Russian facts. However, these studies are instructive providing us with the assumptions that we can use to develop our analysis of the Russian data.

In particular, we need two sets of assumptions to get off the ground. The first set of assumptions concerns the syntax of (sentential) negation and deontic modals, as well as their scopal relations. The second set of assumptions pertains to the interpretation of Russian aspect.

We start with the first set. Following Iatridou & Zeijlstra (2010; 2013), we make the three assumptions in (17). For justification of these assumptions, I refer the interested reader to Iatridou & Zeijlstra (2010; 2013).

(17) a. Negation never lowers at LF
      b. Deontic modals are base-generated lower than the inflectional head (I^0).
      c. Semantic scope has a corresponding configuration in the (overt or covert) syntax

Using these assumptions, the behaviour of "neutral" deontic modals (i.e. modals that do not show polarity sensitivity), such as English *have to*, is straightforwardly explained. The modal is base-generated below I^0 and therefore, below

sentential negation. The scopal relation between the modal and negation is determined by the surface structure in overt syntax. The reverse order is ruled out by the ban on negation lowering. This is illustrated in (18).[8] Note that Iatridou & Zeijlstra (2013) assume that modals are raising verbs and the subject reconstructs.

(18) Neutral deontic modals
    a. John doesn't have to leave.   ✓ ¬ □ / ✗ □ ¬
    b. Syntax: [John$_j$ [not [have-to [$_{vP}$ t$_j$ leave]]]]
    c. LF: [not [have-to [John leave]]]

PPI modals, such as English *must*$_{deon}$, involve head movement that does not reconstruct, as reconstruction results in ungrammaticality (due to the PPI nature of the modal), see Iatridou & Zeijlstra (2013: 549). This is illustrated in (19). Configurations in which PPI modals surface below negation, as in the Spanish example in (20a), are derived by a QR-like covert movement of the modal to the position above negation, see (20b) and (20c).

(19) PPI deontic modals (overt movement)
    a. John must not leave.   ✗ ¬ □ / ✓ □ ¬
    b. Syntax: [John$_j$ [must$_i$ [not [t$_i$ [$_{vP}$ t$_j$ leave]]]]]
    c. LF: [must$_i$ [not [t$_i$ [John leave]]]]

(20) PPI deontic modals (covert movement)
    a. Juan no debe ir.   (Spanish)
        Juan not must go
        'Juan must not go.'   ✗ ¬ □ / ✓ □ ¬
    b. Syntax: [Juan$_j$ [not [must [$_{vP}$ t$_j$ leave]]]]
    c. LF: [must$_i$ [not [t$_i$ [Juan leave]]]]

We now turn to the second set of assumptions, which concern the interpretation of Russian aspect. Following Zinova & Filip (2015), I assume that the perfective aspect in Russian asserts that the action has achieved the end-point and has an inference that the action has started. Moreover, this inference is generated as a scalar implicature (SI), (21).[9] The imperfective aspect asserts that the action has

---

[8]Iatridou & Zeijlstra (2010; 2013) do not commit themselves to particular syntactic projections; so, we will stay away from filling in these particulars as well.

[9]Much previous work on Slavic aspect erroneously claimed that the inference found with PFV is a presupposition (e.g. Bogusławski 1985, Rappaport 1985). For evidence that the presuppositional account cannot be on the right track see the text below and Zinova & Filip (2015). I also thank an anonymous reviewer for pointing out to me some additional data.

started and generates no SI, (22). For expository purposes, I abbreviate aspectual inferences as in (23) where EP = end-point and S = start.

(21) a. Ivan ne pročital   ėtu knigu.
          Ivan not read.PFV this book
          'Ivan didn't read this book completely through.'

       b. Assertion: 'Ivan did not finish reading this book.'

       c. SI: 'Ivan started reading/read a part of this book.'

                                                    (Zinova & Filip 2015: 383)

(22) a. Ivan ne čital   ėtu knigu.
          Ivan not read.IPFV this book
          'Ivan didn't read (any part of) this book.'

       b. Assertion: 'Ivan didn't start reading/read any part of this book.'

       c. no SI

(23) a. Perfective, (21):
          Assertion: ¬ Ivan-read-book EP
          SI: Ivan-read-book S

       b. Imperfective, (22):
          Assertion: ¬ Ivan-read-book S
          (no SI)

Zinova & Filip (2015) argue that evidence for treating the start-inference of the perfective, see (21c), as an SI rather than a presupposition comes from two observations. First, this inference is cancelable, (24):

(24) Ivan ne pročital   ėtu knigu. On daže ne otkryl eë.
      Ivan not read.PFV this book   he even not opened it
      'Ivan didn't read this book. He even didn't open it.'

                                                    (Zinova & Filip 2015: 391)

Second, the start-inference, (21c), shows the projective behaviour characteristic of SIs rather than presuppositions. Chemla (2009) shows that SIs project existentially under negated universal quantifiers, whereas presuppositions project universally in the same configuration. Zinova & Filip (2015) conducted an informal survey that showed that most Russian speakers prefer the existential inference of the perfective, (25b), to the universal one, (25c). The numbers in square brackets show mean acceptability judgments. These results strongly suggest that the start-inference of the perfective is an SI.

(25) a. Nikto iz nas ne pročital učebnik.
nobody of us not read.pfv textbook
'None of us read the textbook.'

b. Some of us started reading at least a part of the textbook. [3.11/4]

c. All of us started reading at least a part of the textbook [1.65/4]

(Zinova & Filip 2015: 396–398)

With these assumptions at hand, we are now ready to account for the two core observations of this paper.

3.2 Proposal

To the best of my knowledge, there is no formal analysis of the aspectual restriction in Russian, nor is there one for relativization and obviation. Below, I briefly discuss some intuitions in Rappaport (1985), which are repeated (with minor elaborations) in more recent accounts. Then, I present my own proposal.

According to Rappaport (1985), the use of perfective in negated strong deontic statements is pragmatically unjustified because it is weaker than a corresponding structure with the imperfective. In his own words:

> if the imperfective verb form in […] *on ne dolžen opravdyvat'sja* 'he need not justify himself' is replaced by a perfective form […], the result would be a statement saying that while there is no need for him to succeed in justifying himself, there may be a need for him to attempt to do so. There is nothing logically incoherent about such a state of affairs, but it makes little pragmatic sense […] (Rappaport 1985: 218–219)

I believe Rappaport's intuitions to be on the right track, although he does not formalize them and assumes that the start-inference of the perfective is presupposition-like. Rappaport (1985) also does not discuss the relativization and obviation facts.

The account I propose in this paper capitalizes on Rappaport's intuitions, but uses recent developments in formal semantics to formalize them. To start, let's see how an SI is generated in a simple perfective sentence like (26a). The assertion of (26a) using the abbreviations introduced above is shown in (26b). Suppose that (26a) competes (for informativity) with a corresponding imperfective statement, whose meaning is given in (26c) as an alternative to (26a). This imperfective alternative, (26c), is stronger than the original perfective statement, (26b), as shown

by the asymmetric entailment relation in (26d). Therefore, the use of (26a) is justified if the speaker supposes that the stronger alternative is not true. This derives the desired SI that Ivan started to leave, as shown in (26e).[10]

(26)   a.   Ivan ne   ušel.
              Ivan not leave.PFV
              'Ivan didn't leave.'

      b.   Assertion: ¬ Ivan-leave EP

      c.   Alternative: ¬ Ivan-leave S   (= imperfective)

      d.   Asymmetric entailment: ¬ Ivan-leave S $\Rightarrow$ ¬ Ivan-leave EP
                                     ¬ Ivan-leave S $\not\Leftarrow$ ¬ Ivan-leave EP

      e.   SI: ¬¬ Ivan-leave S ≡ Ivan-leave S

In a negated deontic sentence with PFV, as in (27a), with the meaning schematized in (27b), the generated SI is as shown in (27e).[11]

(27)   a.   * Ivan ne   dolžen   ujti.
              Ivan not must.PTCP leave.PFV
              Intended: 'Ivan doesn't have to leave.'

      b.   Assertion: ¬ □ Ivan-leave EP

      c.   Alternative: ¬ □ Ivan-leave S   (= imperfective)

      d.   Asymmetric entailment: ¬ □ Ivan-leave S $\Rightarrow$ ¬ □ Ivan-leave EP
                                       ¬ □ Ivan-leave S $\not\Leftarrow$ ¬ □ Ivan-leave EP

      e.   SI: ¬¬ □ Ivan-leave S ≡ □ Ivan-leave S

The asymmetric entailment in (27d) captures Rappaport's intuition that the use of the perfective in the negated deontic sentences is pragmatically unjustified. Note, however, that the derivation of the implicature in (27e) by itself does not explain the aspectual restriction. Combined together, the assertion in (27b) and the SI in (27e) give rise to the following interpretation: Ivan doesn't have an obligation to finish leaving, but he has an obligation to start leaving. (27a) does not have this reading; rather, the sentence is ungrammatical. Therefore, we need to strengthen the account in order to derive the ungrammaticality of (27a).

To achieve this, I propose a more elaborate semantics of deontic modals. I capitalize on the intuition that deontic worlds are *idealized* worlds (Kratzer 2012;

---

[10]The description of SI generation is deliberately vague. As far as I can tell, both Neo-Gricean and grammatical approaches to SIs are compatible with the proposal in this paper, with some adjustments.

[11]The reverse scope interpretation is discussed in §4.

Maribel Romero, p.c.). That is to say, in such worlds, if an action (with a defined telos) starts, it must achieve its end point.[12] To capture this intuition, I add the following conditional to the deontic modal base: 'if x-action S, then x-action EP', where x is an individual whose obligations are relevant (usually the subject). (28) shows a preliminary denotation of Russian *dolžen*.[13]

(28) Denotation of *dolžen* (to be revised)
       $[\![\text{dolžen}_x]\!]^{w_c}(p) = 1$ iff
       $\forall w[w$ is compatible with x's obligations in w_c and w is such that
       if x-action S, then x-action EP$][p$ is true in $w]$
       (where x is an individual whose obligations are relevant, usually the
       subject)

To see how the denotation in (28) helps accounting for the aspectual restriction, consider again the assertion and SI of (27a), now with the contribution of the modal spelled out:

(29) a. * Ivan ne dolžen   ujti.
             Ivan not must.PTCP leave.PFV
             Intended: 'Ivan doesn't have to leave.'

       b. Assertion:
           $\neg\forall w[w$ is compatible with Ivan's obligations in w_c and w is such
           that if Ivan-action S, then Ivan-action EP][Ivan-leave EP in $w]$

       c. SI:
           $\forall w[w$ is compatible with Ivan's obligations in w_c and w is such that
           if Ivan-action S, then Ivan-action EP][Ivan-leave S in $w]$

It is not difficult to see that combining (29b) with (29c) results in a contradiction. The modal base consists of worlds in which every action that Ivan starts is completed by him. (29c) states that in all wolds in the modal base Ivan starts leaving. Therefore, by Modus Ponens, all worlds in the modal base must be such that Ivan's leaving is completed. But (29b) requires there to be at least one world where Ivan's leaving is not completed. Following Gajewski (2002), we assume that contradiction results in ungrammaticality.[14]

---

[12]This is a simplification. The idealized nature of deontic worlds is connected to the fact that the deontic modal base is compatible only with intentional actions, as argued in Goncharov (2018).

[13]I assume the standard interpretation of modals in terms of a modal base and ordering source (Kratzer 1991). The conditional 'if x-action S, then x-action EP' restricts the modal base to worlds where the conditional is true of any action.

[14]This account assumes that there are situations in which implicatures are not easily cancelable, see, for example, Magri (2009; 2011).

This account makes an immediate prediction, namely, that weak/existential deontic/root modals are allowed with both ɪᴘꜰᴠ and ᴘꜰᴠ in Russian. This prediction is borne out:

(30)   *Context:* 'According to the prison regulations...'
       Ivan ne možet ottuda   {zvonit'   / pozvonit'}.
       Ivan not can   from.there call.ɪᴘꜰᴠ   call.ᴘꜰᴠ
       'Ivan is not allowed to call from there.'

To account for relativization and obviation, I extend Stephenson's (2007a) account of epistemic modals to deontic modals. More precisely, I propose that deontic modals, just like epistemics, take an individual/PRO argument, which determines whose obligations are relevant, see (31):

(31)   Denotation of *dolžen* (final)
       $[\![\text{dolžen}]\!]^{w_c}(x, p) = 1$ iff
       $\forall w[w$ is compatible with x's obligations in w_c and w is such that
       if x-action S, then x-action EP$][p$ is true in $w]$

PRO is syntactically present and is co-indexed with the closest referential nominal at LF. The ungrammatical (27a), repeated in (32a), has a simplified syntactic representation in (32b) and a corresponding LE in (32c). PRO is co-indexed with the subject (which reconstructs); thus the modal base is relativized to Ivan and the ungrammaticality arises, as discussed above.

(32)   a. * Ivan ne dolžen   ujti.
            Ivan not must.ᴘᴛᴄᴘ leave.ᴘꜰᴠ
            Intended: 'Ivan doesn't have to leave.'
       b.   Syntax: [Ivan$_j$ [not [[must PRO] [$_{vP}$ t$_j$ leave.ᴘꜰᴠ]]]]
       c.   LF: [not [[must PRO$_j$] [Ivan$_j$ leave]]]

To account for cases where *dolžen* is relativized to the addressee, such as (6), repeated in (33a), we assume following Speas & Tenny (2003), among others that a Speech Act (SA) projection is syntactically present. Among other projections, it contains the Addressee Phrase (AdrP), (33b). The modal covertly moves to AdrP and PRO is co-indexed with the addressee rather than the subject, (33c). Relativization of the modal base to the addressee does not derive a contradiction as the reader can verify by conjoining (33d) with (33e).

(33)   a. *Situation:* Police arrive at a crime scene and see Ivan fleeing with the
          stolen *Mona Lisa*. A police chief to police officers:

Ivan ne dolžen   ujti!
Ivan not must.PTCP leave.PFV
'Ivan must not leave/escape.'

b. Syntax: [_AdrP Adr ... [Ivan_j [not [[must PRO] [_vP t_j leave.PFV]]]]]

c. LF: [[must PRO_j]_k [Adr_j [not [t_k [Ivan leave]]]]]

d. Assertion:
$\forall w$[w is compatible with police' obligations in w_c and w is such that if police-action S, then police-action EP][¬ Ivan-leave EP in w]

e. SI:
$\forall w$[w is compatible with police' obligations in w_c and w is such that if police-action S, then police-action EP][Ivan-leave S in w]

Thus, making the modal base (including the conditional) relativizable to the addressee straightforwardly accounts for the absence of the aspectual restriction in cases like above. It also explains the absence of the parse where the modal is interpreted below negation, see §4.

In this section, we saw how the aspectual restriction and relativization and obviation are derived. More precisely, the aspectual restriction is due to a contradiction between the assertion and SI of the perfective. This account requires that the deontic worlds are idealized such that every action that starts achieves its end-point. Relativization and obviation are explained by two assumptions: (i) deontics, like epistemics, take an individual PRO argument in syntax and (ii) deontics can covertly move to the SA projection where PRO is co-indexed with the addressee. This section, however, did not explore all possible parses. This is the task for the next section.

4 Explaining the patterns

The first pattern we need to account for is in (5), repeated in (34). As can be seen in (34), the only configuration in which Russian deontic necessity modals can appear is when they are followed by the verb in the imperfective and are interpreted below negation, (34a). All other parses are ill-formed.

(34) Pattern 1 to explain (the aspectual restriction)

a. ✓ ¬ □ IPFV

b. ✗ □ ¬ IPFV

c. ✗ ¬ □ PFV

d. ✗ □ ¬ PFV

The contrast between (34a) and (34b) can be straightforwardly accounted for if we analyze Russian *dolžen* as a 'neutral' modal (similar to *have to*) in the system proposed in Iatridou & Zeijlstra (2013). Recall from the discussion in §2.1 that neutral modals (i.e. modals that are not polarity sensitive) have surface scope. That is to say, if they appear below negation, they scope below negation. (35) shows the representation of the well-formed parse in (34a) for Russian *dolžen*:

(35) Parse in (34a): ✓ ¬ □ IPFV

    a. Ivan ne dolžen   uxodit'.
    Ivan not must.PTCP leave.IPFV
    'Ivan doesn't have to leave.'

    b. Syntax: [Ivan$_j$ [not [[must PRO] [$_{vP}$ t$_j$ leave.IPFV]]]]

    c. LF: [not [[must PRO$_j$] [Ivan$_j$ leave]]]

Note that in (35), we continue to assume that modals are raising verbs and the subject reconstructs at LF (Iatridou & Zeijlstra 2013). We also represent the individual argument of *dolžen* as PRO co-indexed with the subject.[15] The imperfective does not generate any relevant implicature; thus, no contradiction arises.

    The parse in (34b), in which *dolžen* is interpreted above negation, is ruled out by the ban on negation lowering:

(36) Parse in (34b): ✗ □ ¬ IPFV

    a. Ivan ne dolžen   uxodit'.
    Ivan not must.PTCP leave.IPFV
    Available: 'Ivan doesn't have to leave.'
    Unavailable: 'Ivan has to stay/not leave.'

    b. Syntax: [Ivan$_j$ [not [[must PRO] [$_{vP}$ t$_j$ leave.IPFV]]]]

    c. LF: *[t$_k$ [[must PRO$_j$] [not$_k$ [Ivan$_j$ leave]]]]   (impossible)

The parse in (34c) was discussed in the previous section when I showed the derivation of the aspectual restriction, (29). For completeness, I repeat its syntactic and LF representations in (37):

---

[15]It is important to mention that PRO here is not indexical, bound, or controlled, see Stephenson (2007b: 500) for discussion. I leave it open how its interpretation is determined. What is important for the account proposed here is that PRO is co-indexed with the closest nominal at LF.

(37) Parse in (34c): ✗ ¬ □ PFV

    a. * Ivan ne dolžen   ujti.
       Ivan not must.PTCP leave.PFV
       Intended: 'Ivan doesn't have to leave.'

    b. Syntax: [Ivan$_j$ [not [[must PRO] [$_{vP}$ t$_j$ leave.PFV]]]]

    c. LF: [not [[must PRO$_j$] [Ivan$_j$ leave]]]

The configuration in (34d), in which negation is interpreted below the modal, can be ruled out by the ban on negation lowering, like in (36). Alternatively, it can be shown that (34d) results in a contradiction, like in (29). In the latter case, it can be supposed that the modal QRs at LF as in the Spanish example in (20). The second alternative is illustrated in (38). I leave it to the reader to verify that (38) gives rise to a contradiction.

(38) Parse in (34d): ✗ □ ¬ PFV

    a. * Ivan ne dolžen   ujti.
       Ivan not must.PTCP leave.PFV
       Intended: 'Ivan has to stay/not leave.'

    b. Syntax: [Ivan$_j$ [not [[must PRO] [$_{vP}$ t$_j$ leave.PFV]]]]

    c. LF: [[must PRO$_j$]$_k$ [not [t$_k$ [Ivan$_j$ leave]]]]

The second pattern to be accounted for is repeated in (39). The crucial point in deriving the aspectual restriction is that obligations (i.e. the modal base) are relativized to the subject. Thus, (34c) and (34d), with relativization information, can be represented as (39a) and (39b) respectively. The well-formedness of (39d) (obviation) is explained in (33). What remains to be account for is the unavailability of (39c).

(39) Pattern 2 to explain (relativization and obviation)

    a. ✗ ¬ □^{SBJ} PFV^{SBJ}
    b. ✗ □^{SBJ} ¬ PFV^{SBJ}
    c. ✗ ¬ □^{ADR} PFV^{SBJ}
    d. ✓ □^{ADR} ¬ PFV^{SBJ}

For completeness, I repeat the syntactic and LF representations of (39d):

(40) Parse in (39d): ✓ □^ADR ¬ PFV^SBJ

    a. Ivan ne dolžen   ujti!
       Ivan not must.PTCP leave.PFV
       'Ivan must not leave/escape!' (uttered in the situation described in (6))

    b. Syntax: [_AdrP Adr ... [Ivan_j [not [[must PRO] [_vP t_j leave.PFV]]]]]

    c. LF: [[must PRO_j]_k [Adr_j [not [t_k [Ivan leave]]]]]

The configuration in (39c), in which the modal is relativized to the addressee, but scopes below negation, is ruled out by the standard claim that (sentential) negation cannot scope above SA. This is schematically shown in (41):

(41) Parse in (39c): ✗ ¬ □^ADR PFV^SBJ

    a. Ivan ne dolžen   ujti!
       Ivan not must.PTCP leave.PFV
       Available: 'Ivan must not leave/escape!' ≈ 'You, officers, must make Ivan stay.'
       Unavailable: 'It's not the case that you, officers, must make Ivan leave.'

    b. Syntax: [_AdrP Adr ... [Ivan_j [not [[must PRO] [_vP t_j leave.PFV]]]]]

    c. LF: *[not [[must PRO_j]_k [Adr_j [t_k [Ivan leave]]]]]   (impossible)

5 In lieu of conclusion

This paper discussed the aspectual restriction of Russian strong deontic modals under negation. Capitalizing on previous work on Russian aspect and interaction between deontic modals and negation, I proposed a formalization of the aspectual restriction. I also showed that the aspectual restriction can be circumvented when the deontic modal is relativized to the addressee rather than the subject.

From a cross-Slavic perspective, it is important to mention that not many Slavic languages show the aspectual restriction discussed for Russian in this paper.[16] (42) shows that in Polish and Serbian negated strong deontic modals can be used with either imperfective or perfective.

---

[16]I thank an anonymous reviewer for asking about cross-Slavic variation with respect to the aspectual restriction in negated deontic constructions. Although I agree with the reviewer that this variation deserves thorough investigation, the claims made in this paper are independent from the cross-Slavic observations. Apart from cursory remarks in this section, I leave the question for future research.

(42) a. Ivan ne dolžen {uxodit' / *ujti}. (Russian)
        Ivan not must.PTCP leave.IPFV leave.PFV
        'Ivan doesn't have to leave.'

    b. Ivan nie musi {iść / wyjść}. (Polish)
        Ivan not must leave.IPFV leave.PFV
        'Ivan doesn't have to leave.'

    c. Ivan nije dužan {odlaziti / otići} kući kasno. (Serbian)
        Ivan not.be obliged go.IPFV go.PFV home late
        'Ivan doesn't have to go home late.'

Interestingly, most Slavic languages do show the aspectual restriction with negative imperatives. Compare (43) with (44):

(43) a. {Otkryvaj / otkroj} okno! (Russian)
        open.IPFV.IMP open.PFV.IMP window
        'Open the window!'

    b. {Jedz / zjedz} tego jabłka! (Polish)
        eat.IPFV.IMP eat.PFV.IMP that apple
        'Eat that apple!'

    c. {Jedi / pojedi} tu jabuku! (Serbian)
        eat.IPFV.IMP eat.PFV.IMP that apple
        'Eat that apple!' (Despić 2016: 2)

(44) a. Ne {otkryvaj / *otkroj} okno! (Russian)
        not open.IPFV.IMP open.PFV.IMP window
        'Don't open the window!'

    b. Nie {jedz / *zjedz} tego jabłka! (Polish)
        not eat.IPFV.IMP eat.PFV.IMP that apple
        'Don't eat that apple!'

    c. Ne {jedi / *pojedi} tu jabuku! (Serbian)
        not eat.IPFV.IMP eat.PFV.IMP that apple
        'Don't eat that apple!' (Despić 2016: 2)

The parallel behaviour of strong deontics and imperatives, as we see in Russian, is not unexpected. In many accounts, deontics and imperatives receive similar treatment (e.g. Han 1999; Ninan 2005; Kaufmann 2012). The challenge is to explain why some Slavic languages (like Russian) show the aspectual restriction with both strong deontic modals and imperatives, whereas other Slavic languages (like

Table 1: Deontics and imperatives across Slavic

	imperatives	deontics
Russian	imp ¬ {IPFV / *PFV }	¬ □$_{\text{deon}}$ {IPFV / *PFV }
Serbian, Polish	imp ¬ {IPFV / *PFV }	¬ □$_{\text{deon}}$ {IPFV / PFV }

Polish and Serbian) show the aspectual restriction only with imperatives. This challenge is summarized in Table 1.

In the remainder of this concluding section, I briefly outline how the challenge presented by cross-Slavic variation can be addressed. In particular, I would like to suggest that the observed difference between Slavic languages is due to the differences in their aspectual systems. Slavic aspect is a complex topic and I will not be able to do justice to the vast literature on this subject. However, I would like to point out that there are accounts that try to systematize aspectual phenomena across Slavic languages. One such account is the so-called East-West Theory of Slavic aspect. According to this Theory, there is a systematic difference between Eastern Slavic languages (Russian, Ukrainian, Belarus) and Western Slavic languages (Serbian, Czech, Slovenian, etc.), with some mixed cases (Polish, Bulgarian, Macedonian), see Fortuin & Kamphuis (2015) for a recent review. The difference can be summarized as follows:

> [In the Eastern group] the meaning of the [PFV] is made up of three "layers":
>
> (a) the event expressed by the predicate is terminative;
>
> (b) the event is seen as a totality [...] such that there is a change of situation;
>
> (c) the event expressed by the [PFV] verb is sequentially connected to a following and/or preceding situation.
>
> [In the Western group, perfective only needs to satisfy (a) and (b).]
>
> (Fortuin & Kamphuis 2015: 165)

The difference in use of imperfective/perfective between Western and Eastern groups can be seen in Table 2. Table 2 shows that the use of perfective in Eastern Slavic languages is more restricted. Arguably, this is due to the fact that perfective in the Eastern group has an additional condition: it must be sequentially connected (condition (c) above).

Table 2: The distribution of aspect in Western and Eastern Slavic languages (from Fortuin & Kamphuis 2015: 173, 182)

		Western group	Eastern group
Habitual	Non-past contexts	IPFV / PFV	IPFV (PFV)
	Past contexts	IPFV / PFV	IPFV
Narration	Present tense narration	IPFV / PFV	IPFV
	Past tense narration	IPFV / PFV	PFV

I would like to suggest that the difference between Russian, on the one hand, and Polish and Serbian, on the other hand, with respect to the aspectual restriction in deontic and imperative constructions is due to the same factor. In imperatives (by their nature) the sequential connection to a following situation is present in both Eastern and Western Slavic languages (Bogusławski 1985; Han 1999). This makes Western Slavic languages superficially look like Eastern Slavic languages with regard to imperatives. Deontics, on the other hand, do not require sequential connection, which creates the difference between Eastern and Western Slavic languages in negated deontic construction. This idea is illustrated in Table 3.

Table 3: Aspect in deontics and imperatives

imperatives with PFV		deontics with PFV	
Western group	Eastern group	Western group	Eastern group
(a)	(a)	(a)	(a)
(b)	(b)	(b)	(b)
seq. connected	(c) seq. connected		(c) seq. connected

Suppose that the sequential connection to a preceding or following situation goes hand-in-hand with SI generation in the aspectual system. Recall that in this paper I argued that SI of the perfective is responsible for the aspectual restriction. This line of reasoning will correctly account for the fact that the aspectual restriction with imperatives exists in both Eastern and Western Slavic groups, whereas the aspectual restriction with strong deontics is only active in the Eastern group. I leave further investigation of this line of reasoning for future research.

Abbreviations

ADR	addressee	LF	logical form
DAT	dative	PFV	perfective
deon	deontic	PPI	positive polarity item
EP	end-point	PTCP	participle
epist	epistemic	QR	quantifier raising
FOC	focus	S	start
IMP	imperative	SA	speech act
INF	infinitive	SBJ	subject
IPFV	imperfective	SI	scalar implicature

Acknowledgements

References

Bogusławski, Andrej. 1985. The problem of the negated imperative in perfective verbs revisited. *Russian Linguistics* 9(2-3). 225–239.   DOI:10.1007/BF03380061

Chemla, Emmanuel. 2009. Presuppositions of quantified sentences: Experimental data. *Natural Language Semantics* 17(4). 299–340.   DOI:10.1007/s11050-009-9043-9

de Haan, Ferdinand. 1997. *The interaction of modality and negation: A typological study.* New York: Garland.

Despić, Miloje. 2016. *Negative imperatives, aspect and agree: From locality to argument structure.* Presented at the LinG Colloquium, Linguistics in Göttingen, October 2016.

Forsyth, James. 1970. *A grammar of aspect: Usage and meaning in the Russian verb.* Cambridge: Cambridge University Press.

Fortuin, Egbert & Jaap Kamphuis. 2015. The typology of Slavic aspect: A review of the East–West Theory of Slavic aspect. *Russian Linguistics* 39(2). 163–208. DOI:10.1007/s11185-015-9144-7

Gajewski, Jon R. 2002. *L-analyticity in natural language*. Manuscript, MIT.

Goncharov, Julie. 2018. *Imperatives and intentions*. Presented at the Non-Canonical Imperatives workshop at Humboldt-Universität zu Berlin, May 2018.

Grano, Thomas. 2017. The logic of intention reports. *Journal of Semantics* 34(4). 587–632. DOI:10.1093/jos/ffx010

Han, Chung-Hye. 1999. Deontic modality, lexical aspect and the semantics of imperatives. In *Linguistics in the morning calm 4*, 479–495. Seoul: Hanshin Publications.

Hellie, Benj. 2016. Obligation and aspect. *Inquiry* 59(4). 398–449. DOI:10.1080/0020174X.2015.1042037

Homer, Vincent. 2011. *Polarity and modality*. Los Angeles, CA: University of California. (Doctoral dissertation).

Iatridou, Sabine & Hedde Zeijlstra. 2010. On the scopal interaction of negation and deontic modals. In Maria Aloni, Harold Bastiaanse, Tikito de Jager & Katrin Schulz (eds.), *Languages, logic and meaning*, 315–324. Berlin: Springer.

Iatridou, Sabine & Hedde Zeijlstra. 2013. Negation, polarity and deontic modals. *Linguistic Inquiry* 44(4). 529–568. DOI:10.1162/ling_a_00138

Kaufmann, Magdalena. 2012. *Interpreting imperatives*. New York: Springer.

Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: de Gruyter. DOI:10.1515/9783110126969.7.639

Kratzer, Angelika. 2012. *Modals and conditionals*. New York: Oxford University Press.

Magri, Giorgio. 2009. A thery of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics* 17(3). 245–297. DOI:10.1007/s11050-009-9042-x

Magri, Giorgio. 2011. Another argument for embedded scalar implicatures based on oddness in downward entailing environments. *Semantics and Pragmatics* 4(6). 1–51. DOI:10.3765/sp.4.6

Ninan, Dilip. 2005. Two puzzles about deontic necessity. In Valentine Hacquard, Jon R. Gajewski, Bernard Nickel & Seth Yalcin (eds.), *MIT Working Papers in Linguistics 51: New work on modality*, 149–178. Cambridge, MA: MITWPL.

Paducheva, Elena V. 2013. *Russkoe otricatel'noe predlozhenie*. Moscow: Jazyki slavjanskoi kul'tury.

Rappaport, Gilbert. 1985. Aspect and modality in contexts of negation. In Michael S. Flier & Alan Timberlake (eds.), *The scope of Slavic aspect*, 194–222. Columbus, OH: Slavica Publishers.

Rett, Jessica. 2016. On a shared property of deontic and epistemic modals. In Nate Charlow & Matthew Chrisman (eds.), *Deontic modality*, 200–229. Oxford: Oxford University Press.

Speas, Peggy & Carol L. Tenny. 2003. Configurational properties of point of view roles. In Anna Maria Di Sciullo (ed.), *Asymmetry in grammar, Vol. 1: Syntax and semantics*, 315–344. Amsterdam: John Benjamins. DOI:10.1075/la.57.15spe

Stephenson, Tamina. 2007a. A parallel account of epistemic modals and predicates of personal taste. In Estela Puig-Waldmüller (ed.), *Proceedings of Sinn und Bedeutung 11*, 583–597. Barcelona: Pompeu Fabra University.

Stephenson, Tamina. 2007b. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30(4). 487–525. DOI:10.1007/s10988-008-9023-4

Wiltschko, Martina. 2017. *The syntacticization of linguistic interaction: Evidence from confirmationals and response markers.* Presented at the Hebrew University of Jerusalem.

Zalizniak, Anna A. 2006. *Mnogoznachnost' v jazyke i sposoby ee predstavlenija.* Moscow: Jazyki slavjanskix kul'tur.

Zinova, Yulia & Hana Filip. 2015. Scalar implicatures of Russian verbs. In *Formal Approaches to Slavic Linguistics 23: The Berkeley Meeting 2014*, 382–401. Ann Arbor, MI: Michigan Slavic Publications.

Chapter 4

Epistemic comparatives and other expressions of speaker's uncertainty

Julie Goncharov

Hebrew University of Jerusalem / University of Göttingen

Monica Alexandrina Irimia

University of Modena and Reggio Emilia

Our study endeavors to further our understanding of the ways the speaker's perspective is expressed in natural language. We analyze a rarely discussed construction, namely epistemic comparatives and their interaction with inferential indirect evidentials and epistemic modals. We show that epistemic comparatives are incompatible with inferential indirect evidentials, but are well-formed with epistemic modals. We base our discussion on data from Bulgarian and we also show that similar facts hold in Romanian, thus strengthening the empirical coverage. On the theoretical side, we claim that inferential indirect evidentials are structurally distinguished from epistemic modals. This accounts for their different behavior with epistemic comparatives, thus providing further support to accounts which take indirect evidentials and epistemic modals to be separate categories.

Keywords: epistemic comparatives, indirect evidentials, epistemic modals, inferentials, Bulgarian, Romanian

1 Introduction

Recent research on epistemic modality has brought to the fore a previously neglected construction, namely EPISTEMIC COMPARATIVES (hereafter ECs; von Fintel & Kratzer 2014; Herburger & Rubinstein 2014). We illustrate this phenomenon with the example in (1) from Bulgarian. In this sentence, the speaker conveys that she believes the state of affairs where Ivan is in the office to be more plausible than the state of affairs where Ivan is at home.

(1) Ivan po-skoro   e v ofisa otkolkoto vkušti.
     Ivan more-soon is in office than   home
     'According to the speaker, it is more plausible that Ivan is at work than at
     home.'   (Bulgarian)

We contribute to the rising interest in ECs by examining some previously un-
reported properties of these constructions. Starting from Bulgarian, we address
an interesting puzzle raised by the interaction of ECs with other expressions
of speaker's uncertainty. This puzzle concerns the difference between epistemic
modals and grammaticalized evidentials. ECs are not possible with inferential
grammaticalized evidentials, while they are well-formed with epistemic modals.
The sentence in (2) contains a present perfect, which in Bulgarian obtains an IN-
DIRECT EVIDENTIAL (IE) interpretation. The curious observation is that the EC is
not well-formed under the inferential reading of the indirect evidential.

(2) ?? Ivan bil   po-skoro   v offisa otkolkoto vkušti.
      Ivan be.PST.PTCP more-soon in office than   home
      Intended: 'Apparently, it is more plausible that Ivan is in the office than
      at home.'   (Bulgarian)

However, ECs seem to be possible with epistemic modals, as seen below:

(3) Ivan po-skoro   moje da e v ofisa otkolkoto vkušti.
     Ivan more-soon can   DA is in office than   home
     'It is more plausible that Ivan might be at work than at home.' (Bulgarian)

These facts are not restricted to Bulgarian. We show that similar patterns obtain
in yet another language that has ECs and grammaticalized evidentiality, namely
Romanian. Parallel observations from outside the Slavic family make available
a clearer perspective into the nature of ECs and evidentiality, highlighting the
cross-linguistic uniformity of these phenomena. We provide a comprehensive
analysis for both Bulgarian and Romanian.

   More precisely, we propose to explain the puzzle by arguing that the ill-formed-
ness of ECs with grammaticalized evidentials boils down to competition for the
same position. We build on a decompositional account of ECs (Herburger & Ru-
binstein 2014) and an analysis of the Assert operator as a degree relation (Green-
berg & Wolf 2018). We show that our proposal makes a number of correct predic-
tions, including the difference between inferential and hearsay evidentials when
combined with ECs.

   The paper is organized as follows: §2 provides some background on ECs and
evidentiality, §3 presents the core observation, §4 spells out our analysis of the

facts described in §3, §5 contains some concluding remarks and avenues for future research.

2 Background on epistemic comparatives

ECs compare two propositions with respect to the probabilities assigned (usually) by the speaker.[1,2] As we already mentioned, in the Bulgarian example in (1), the speaker communicates that she believes (or is committed to act as if she believes) the state of affairs where Ivan is in the office to be more plausible than the state of affairs where Ivan is at home.

One of the defining properties of ECs is that they employ a comparative form of a temporal adverb. A fuller definition of ECs should be in terms of their meaning and distribution. For reasons of space, we only mention this salient characteristic here and refer the readers to the works on ECs mentioned in this paper. (4) illustrates the Bulgarian *(po-)skoro* with its usual temporal meaning.[3]

(4) Toj dojde po-skoro   otkolkoto očakvah.
     he came more-soon than   expected.1SG
     'He came sooner than I expected.'   (Bulgarian)

ECs have not been thoroughly investigated. There are four studies we are aware of: von Fintel & Kratzer (2014); Herburger & Rubinstein (2014); Goncharov (2014); Goncharov & Irimia (2018). von Fintel & Kratzer (2014) look at the properties of

---

[1]Propositional comparatives can, in fact, be classified into three types: (i) metalinguistic comparatives, expressing appropriateness (Bresnan 1973; McCawley 1988; Embick 2007; Morzycki 2011; a.o.), (ii) preference comparatives, ranking choices (Giannakidou & Stavrou 2009; Giannakidou & Yoon 2011), and (iii) epistemic comparatives, ordering speaker's probabilities (von Fintel & Kratzer 2014; Herburger & Rubinstein 2014). There is no consensus in the literature on whether these three types of comparatives should be treated uniformly or not (see Morzycki 2011 for discussion). We stay away from this debate here, as we focus on ECs.

    (i)   a. She is more fit than slender.   (metalinguistic)
          b. I would rather die than marry him.   (preference)

[2]Relativization to the speaker is true only of unembedded ECs. When ECs are embedded under an attitude predicate, they are relativized to the attitude holder and when ECs occur in a question, they are relativized to the hearer, see Herburger & Rubinstein (2014: 562).

[3]How modal uses of temporal adverbs like *skoro* relate to their temporal uses is an important question. Although we do not address this issue here, we believe the diachronic account of the development of modal uses of adverbs like *skoro* (and *rather*) from their temporal uses presented in Gergel (2016) is on the right track.

ECs in German without attempting a formal analysis. Goncharov (2014) describes ECs in Russian. Herburger & Rubinstein (2014) use ECs to argue that possibility modals in German are not gradable. We believe the analysis proposed by Herburger & Rubinstein (2014) is on the right track. We modify it slightly in §4 to align it better with our assumptions. Goncharov & Irimia (2018) discuss the cross-linguistic variation of ECs, and provide an account for it.

In the next section we introduce the puzzle that is the focus of the present paper. The facts are based on novel data related to ECs in Bulgarian. We also show that similar facts obtain in Romanian. Although not a Slavic language, Romanian proves very useful for deepening our understanding of the Bulgarian facts and for making cross-linguistic generalizations. Romanian has ECs, as seen in (5a). Similarly to Bulgarian, this sentence is constructed with an adverb (*mai degrabă* 'more soon'), which also has a temporal meaning, illustrated in (5b).[4]

(5) a. Ion este mai   degrabă   la birou decât acasă.
         Ion is   more ADV.early at office than home
         'According to the speaker, it is more plausible that Ion is in the office than at home.'

     b. Trebuie să   vii   mai degrabă.
         need   SBJV come.2SG more ADV.early
         'You need to come sooner.'   (Romanian)

Importantly, Romanian is part of the Balkan Sprachbund, just like Bulgarian. Thus, it exhibits several features that are characteristic to this geographical area, among which: (i) presence of suffixed definiteness; (ii) lack of sequence-of-tense phenomena; (iii) an analytic future constructed from the auxiliary *want* (see especially Tomić 2006a). Another feature that Bulgarian and Romanian share, which is most important for us here, is the existence of grammaticalized evidentiality. Looking at two languages – Bulgarian (Slavic) and Romanian (Romance) – in a typological contact situation can give us relevant hints into the nature of the phenomena discussed in this paper. In the next section we show that ECs are not well-formed with grammaticalized indirect evidentials, under inferential interpretations. We use data from both Bulgarian and Romanian.

---

[4]ECs built from temporal adverbs can be absent in certain languages, for example present day English, as seen in (i). ECs with temporal adverbs existed in Middle/Early Modern English, as discussed in Gergel (2016). ECs formed as 'it is more likely to … than' arguably have similar meaning, but for the purposes of this paper we take them to be a different constructions.

  (i)   * John is {sooner/more/rather} at work than at home.

3 ECs and grammaticalized evidentials

Both Romanian and Bulgarian have verbal forms that express (indirect) evidentiality. We take the existence of these forms to support the claim that (indirect) evidentiality is grammaticalized in these languages.[5]

We provide two relevant examples in (6a) and (6b):

(6) a. Ivan bil   vkušti.
         Ivan be.PST.PTCP=EVID home
         'Apparently, Ivan is at home.'   (Bulgarian)

     b. Ion o   fi acasă.
         Ion PRESMP be at.home
         'Apparently, Ion is at home.'   (Romanian)

As mentioned in the introduction and illustrated in (6a), the Bulgarian past participle can carry IE interpretations, see Jakobson (1956); Comrie (1976); Palmer (1986); Izvorski (1997); Alexiadou et al. (2003); Tomić (2006b); a.o. In Bulgarian, the present perfect is ambiguous. It can have a regular temporal interpretation or function as an evidential. The present perfect is formed from a present auxiliary and a past participle. With the 3rd person the auxiliary can be omitted. In such cases, the present perfect is unambiguously interpreted as evidential, see Izvorski (1997: fn. 7). In this paper, we use this strategy to rule out the temporal interpretation (*bil* is 3SG.PAST.PART).[6]

In the Romanian example in (6b), IE is conveyed by the presumptive mood (PRESMP), as discussed especially by Slave (1956); Coşeriu (1976); Reinheimer Râpeanu (2000); Squartini (2001); Irimia (2010; 2018); a.o.[7]

The puzzle we discuss in this paper is that ECs are not well-formed in sentences with IEs, whereas they are grammatical with modal auxiliaries.[8] Compare (7) with (8):

(7) a. ?? Ivan bil   po-skoro   v offisa otkolkoto vkušti.
            Ivan be.PST.PTCP more-soon in office than   home
            Intended: 'Apparently, it is more plausible that Ivan is in the office
            than at home.'   (Bulgarian)

---

[5]In this paper we make a distinction between grammaticalized evidentials and what we will later call phrasal/lexical evidentials. See §5 for some discussion.

[6]We are grateful to Roumyana Pancheva for clarifying this point.

[7]The Romanian presumptive form *o* 'PRESMP.3SG' is a modal auxiliary which shows ambiguity between a modal future reading and an evidential reading. The presumptive mood in Romanian can also be constructed from other modal auxiliaries, e.g. conditional, subjunctive.

[8]We limit our investigation to indirect evidentiality because it has been convincingly demonstrated that ECs are unacceptable with direct evidence, just like regular epistemic modals (see Herburger & Rubinstein 2014).

b. ?? Ion o   fi mai degrabă la birou decât acasă.
    Ion PRESMP.FUT be more ADV.early at office than home
    Intended: 'Apparently, it is more plausible that Ion is in the office
    than at home.'   (Romanian)

(8)   a. Ivan po-skoro moje da e v ofisa otkolkoto vkušti.
      Ivan more-soon can DA is in office than   home
      'It is more plausible that Ivan might be at work than at home.'
                                                   (Bulgarian)

  b. Ion poate fi   mai degrabă la birou decât acasă.
      Ion can.3SG be.INF more ADV.early at office than home
      'It is more plausible that Ion might be in the office than at home.'
                                                  (Romanian)

One important observation is that the incongruence between the indirect evidential and the EC only affects the INFERENTIAL INDIRECT EVIDENTIALS (IIE).[9]

It is well known that indirect evidentials come in two broad categories: inferential (the statement is based on the inference the speaker draws from available evidence) and hearsay (the statement is based on somebody else's report). In Bulgarian the present-perfect-as-an-evidential in (6a) is ambiguous. It can obtain both inferential and hearsay readings (see especially Smirnova 2013; Koev 2017). Ill-formedness only arises under the inferential reading. The hearsay interpretation ('based on what I heard, it is more plausible that Ivan is in the office than at home') is not deviant.[10] The Romanian example in (7b) is not ambiguous. Hearsay readings of the PRESMP are normally constructed from conditional morphology on the auxiliary (e.g. *ar*=COND.3), see Irimia (2018).[11]

---

[9]We are grateful to Roumyana Pancheva for the illuminating discussion on this point.

[10]Koev (2017: fn. 2) notes that the inferential reading of the present-perfect-evidential in Bulgarian is more restricted than its hearsay reading. There is also speaker variation in this respect. Therefore, speakers who only have the hearsay reading will not perceive the contrast we are interested in.

[11]One way to disambiguate the hearsay readings from the conditional ones is by embedding them under overt hearsay marking ('they say that'), as in the following example. Note that this sentence also illustrates the phenomenon of evidential concord.

   (i)   (Se spune că) Ion ar   fi fost la birou mai degrabă decât acasă.
       SE says that Ion COND=PRESMP.3 be been at office more ADV.early than home
      '(They say that) based on hearsay, it is more plausible that Ion was in the office rather than at home.'

To summarize, the novel observation is that ECs are compatible with epistemic modals, but are deviant with grammaticalized IIEs. The observation is based on Bulgarian and Romanian, two languages which grammaticalize evidentials using the present perfect and the presumptive mood respectively.

4 Proposal

In a nutshell, our proposal is that *-er* in ECs and grammaticalized (inferential) evidentials are both degree modifiers of the (gradable) SPEECH ACT (SA) operator Assert. Thus, they computationally compete for the same position. Therefore, the deviance of the examples in (7) is similar to **John is more that tall*, in which *more* and *that* specify the degree of tallness. To flesh out our account we build on the insights in Greenberg & Wolf (2018), who propose that Assert contains a degree argument and thus has to compose with a degree modifier (similarly to gradable adjectives). We combine this insight with the idea in Davis et al. (2007) that evidentials reset the degree of credence of the speaker.

4.1 Preliminary remarks

We begin this subsection by discussing Greenberg & Wolf (2018) and then, spell-out some details of Herburger & Rubinstein's (2014) analysis of ECs, on which we build our account.

Greenberg & Wolf (2018) base their proposal on the idea that Assert is modifiable. The evidence they use comes from the difference in distribution between MODAL ADVERBS (MAdvs) and MODAL ADJECTIVES (MAdjs). It has been previously noticed in the literature that MAdvs and MAdjs differ in (at least) the following respects.

First, speaker orientation is stronger with MAdvs than MAdjs, as seen in the contrast between (9a) and (9b), cited from Greenberg & Wolf (2018).

(9) a. A: It is probable that they have run out of fuel.

 B: Whose opinion is this?

 b. A: They have probably run out of fuel.

 B: # Whose opinion is this?

Second, only MAdjs are possible in the antecedent of conditionals that do not have an assertive force by themselves (e.g. Kratzer 1991), as shown in (10):

(10) a. # If John is possibly/probably/definitely in the office, I will call the office.

     b. If it is possible/probable/certain that John is in the office, I will call the office.   (Greenberg & Wolf 2018)

Finally, in confirmational tag-questions, *yes* picks up the content of the proposition when MAdv is used, see (11a). On the other hand, MAdj confirms the degree of credence, see (11b).

(11) a. A: John is possibly/probably/certainly in the office, eh?/right?

         B: Yes. (John is in the office.)

     b. A: It's possible/probable/certain that John is in the office, eh?/right?

         B: Yes. (It's possible/probable/certain that John is in the office.)

Greenberg & Wolf (2018) use these differences to suggest that MAdvs function as Assert modifiers. To implement this idea compositionally they use the degree-semantics framework. More specifically, they propose the denotation of Assert in (12).[12]

    In (12), informally speaking, CR_x is a measure function from propositions to degrees on the credence scale (see Herburger & Rubinstein 2014 and references cited there for some discussion). Additionally, we assume that 'x's credence in p' implies x's commitment to behave as if (x believes that) p. This is important for explaining the difference between IIE and hearsay.

(12) $[\![\text{Assert}_x]\!]^c = \lambda p \lambda d_d . \text{CR}_x(p) \geq d$
     where CR_x is a function that takes a proposition p and returns a degree to which x is committed to behave as if x believes that p

According to this system, simple sentences like (13a) have the LF in (13b), where the degree argument is saturated by a contextually set POS(itive) operator, defined in (13c). This is similar to the standard treatment of gradable adjectives in simple sentences like *John is tall* in degree-semantics (e.g. Kennedy & McNally 2005). (13d) computes the truth-conditions of (13a). (13a) is true iff there is a contextually

---

[12]Greenberg & Wolf (2018) also propose a more complex denotation of Assert in terms of context update. Although we believe that their implementation of the gradable Assert is more adequate and can be easier incorporated into the existing SA and evidential systems, for the purpose of this paper we use a simplified denotation. This denotation is sufficient to demonstrate the interaction between ECs and IIE.

set degree d such that the speaker has credence in the proposition that John is home at least to d.

(13)   a.   John is at home.

       b.   LF: [[POS [Assert]] [John is at home]]

       c. $[\![\text{POS}]\!]^c = \lambda G \lambda p \exists d [d \geq \text{STANDARD}_c \wedge G(p)(d)]$

       d. $\big[[[\![\text{POS}]\!]^c (\![\text{Assert}_x]\!]^c)]([\![\text{J. is at home}]\!]^c)\big] =$
          $= [[\lambda G \lambda p \exists d [d \geq \text{STANDARD}_c \wedge G(p)(d)]](\lambda p \lambda d . \text{CR}_x(p) \geq d)]([\![\text{J. is at}$
          $\text{home}]\!]^c) =$
          $= \exists d [d \geq \text{STANDARD}_c \wedge \text{CR}_x([\![\text{J. is at home}]\!]^c) \geq d]$
          (where x = the speaker in c)

In more complex sentences with modal adverbs as in (14a), the degree argument of Assert is saturated by the MAdv. Informally, we take MAdvs to set degrees on the credence scale, e.g. *possibly p* holds iff $\text{CR}_x(p) > 0$, *probably p* holds iff $\text{CR}_x(p) > 0.5$, and *certainly p* holds iff $\text{CR}_x(p) > 0.98$, see also Greenberg & Wolf (2018) for a more formal discussion. (14a) has the LF in (14b). The denotation of *probably* is given in (14c) and the truth-conditions for (14a) in (14d). (14a) is true iff the degree of speaker's credence in the proposition that John is home is greater than 0.5 on the credence scale.

(14)   a.   John is probably at home.

       b.   LF: [[probably [Assert]] [John is at home]]

       c. $[\![\text{probably}]\!]^c = \lambda G \lambda p \exists d [d > 0.5 \wedge G(p)(d)]$

       d. $[[\![\text{probably}]\!]^c (\![\text{Assert}_x]\!]^c)] ([\![\text{J. is at home}]\!]^c) =$
          $= [[\lambda G \lambda p \exists d [d > 0.5 \wedge G(p)(d)]](\lambda p \lambda d . \text{CR}_x(p) \geq d)] ([\![\text{J. is at}$
          $\text{home}]\!]^c)$
          $= \exists d [d > 0.5 \wedge \text{CR}_x ([\![\text{J. is at home}]\!]) \geq d]$
          (where x = the speaker in c)

A potential objection to the idea of gradable Assert could be that SA operators are not normally part of the compositional derivation and cannot be embedded. However, various contributions have shown that under certain conditions SA operators can be embedded, see Greenberg & Wolf (2018) for references.

Turning now to ECs, we follow the decompositional account of Herburger & Rubinstein (2014), who analyze German ECs of the type in (15). For Herburger & Rubinstein (2014) *eher* is decomposed into a comparative head *-er* with the regular denotation in (16a) and an epistemic component *eh-*, which they take to be a believe-type predicate with a degree argument, see (16b).

(15) Hans ist eher   auf der Arbeit als   zu Hause.
Hans is sooner at   the work   than at home
'According to the speaker, it is more plausible that Hans is at work than
at home.'   (German)

(16) a. $[\![\text{-er}]\!] = \lambda P \lambda Q \,.\, \text{MAX}(Q) > \text{MAX}(P)$
      b. $[\![\text{eh-}]\!]^z = \lambda p \lambda d \,.\, z$ is d-ready to believe p
          (defined only if z doesn't have direct evidence for p)

According to this system, the example in (15) has the LF in (17a) and the truth-
conditions in (17b). (16) and (17) are from Herburger & Rubinstein (2014: 564–
565). With angle brackets ($\langle . . . \rangle$), we signal the material that is not phonologically
present.

(17) a. LF: [[-er [than ⟨eh- Hans is⟩ at home]] [eh- Hans is at work]]
      b. $[\![\text{-er}]\!]^z([\![\text{than eh- Hans is at home}]\!]^z)([\![\text{eh- Hans is at work}]\!]^z) =$
          $\text{MAX}(\lambda d \,.\, z$ is d-ready to believe that Hans is at work) $> \text{MAX}(\lambda d \,.\, z$ is
          d-ready to believe that Hans is at home)
          where z is the speaker

Interesting support for the decompositional analysis of *eher* comes from the fact
that in Austrian and Bavarian German there is a discourse particle *eh-* with a sim-
ilar epistemic interpretation, see (18) from Herburger & Rubinstein (2014: ex.32).
See also Zobel (2017) for a detailed investigation of *eher*.

(18) Das ist auf regionaler Ebene eh möglich.
that is on regional   level   eh possible
'That is anyways possible on a regional level.'   (Austrian German)

4.2 Analysis

We begin our analysis by discussing the interpretation of INFERENTIAL INDIRECT
EVIDENTIALS (IIEs). We propose that they function as degree modifiers of Assert
on a par with MAdvs (as discussed above). This claim is limited to IIEs; in this
paper, we remain agnostic with respect to other types of evidentials, apart from
hearsay evidentials that, as we show below, are not modifiers of Assert. Further-
more, building on Davis et al. (2007), we assume that IIEs reset the threshold
of the credence function from a contextually set value (set by POS) to the ev-
idential value, see (19). This is illustrated for Romanian in (20). Bulgarian IIEs

receive a similar account. For reasons of space, we provide only the LF and the truth-conditions.[13]

(19) $[\![\text{Evid}_{\text{IIE}}]\!]^c = \lambda G \lambda p \exists d[d = \mu^c(\textit{evid}) \wedge G(p)(d)]$
where μ maps the strength of evidence to a degree on the credence scale in c

(20) a. Ion o   fi acasă.
         Ion PRESMP=IIE be at.home
         'Apparently, Ion is at home.'   (Romanian)

      b. LF: [[Evid$_{\text{IIE}}$[Assert]] [Ion is at home]]

      c. $\exists d[d = \mu_c(\textit{evid}) \wedge \text{CR}_x([\![\text{Ion is at home}]\!]) \geq d]$
         (where x = the speaker in c)

Our proposal for IIEs makes the immediate prediction that IIEs are incompatible with MAdvs, as they compete for the same position. This prediction is borne out for Romanian *posibil*, as shown in (21):

(21) ?? Ion (posibil) o   fi (posibil) acasă.
          Ion possibly PRESMP=IIE be (possibly) at.home
          Intended: 'Possibly, Ion is apparently at home.'   (Romanian)

We now account for the core observation, namely that ECs are incompatible with IIEs. We propose that the epistemic component in ECs, expressed by *eh-* in German (see the observations above) can be assimilated to Greenberg & Wolf's (2018) Assert. We generalize Herburger & Rubinstein's (2014) analysis of German to Bulgarian and Romanian and represent *eh-* abstractly as Epist below. Both Assert and Epist are gradable and both manipulate (usually) the speaker's degree of credence in the proposition expressed by the prejacent. There is, however, an important difference between the two: Epist is presuppositional, i.e. it is undefined if the speaker has direct evidence (see the discussion in Herburger & Rubinstein 2014).

(22) $[\![\text{Epist}_x]\!]^c = \lambda p \lambda d \, . \, \text{CR}_x(p) \geq d$
       (defined only if x doesn't have direct evidence for p)

---

[13]We gloss over the mechanics of how the evidential meaning comes about and how the source of evidence is encoded. These details are orthogonal to the point made in this paper, but see the discussion in Koev (2017).

To simplify the computation of comparatives and make it parallel to modal adverbs, we slightly modify the structure advocated by Herburger & Rubinstein (2014) for eher-comparatives. We assume that EC in (23a) has the LF in (23b). We further assume that *-er* has the denotation in (23c), where the *than*–clause is a definite description of degrees (as assumed for gradable adjectives), see (23d). For reasons of space, we omit the details of how the meaning of the than-clause is obtained. The truth-conditions for (23a) (if defined) are given in (23e) and paraphrased in (23f). We show this using Bulgarian, but the same holds for Romanian.

(23) a. Ivan po-skoro   e v ofisa otkolkoto vkušti.
          Ivan more-soon is in office than   home
          'According to the speaker, it is more possible that Ivan is at work than
          at home.'   (Bulgarian)

   b. [[[-er [than ⟨Epist Ivan is⟩ at home]] Epist] [Ivan is in the office]]

   c. $\llbracket\text{-er}\rrbracket = \lambda d \lambda G \lambda p \exists d'[d' > d \wedge G(p)(d')]$

   d. $\llbracket\text{than Epist}_x \text{ Ivan is at home}\rrbracket^c = \text{MAX}(\{d : \text{CR}_x(\llbracket\text{Ivan is at home}\rrbracket) \geq d\})$

   e. $\llbracket\text{-er}\rrbracket(\llbracket\text{than Epist}_x \text{ Ivan is at home}\rrbracket^c) =$
       $= \lambda G \lambda p \exists d'[G(p)(d') \wedge d' > \text{MAX}(\{d : \text{CR}_x(\llbracket\text{Ivan is at home}\rrbracket) \geq d\})]$
       $\llbracket\text{-er than Epist}_x \text{ Ivan is at home}\rrbracket^c(\llbracket\text{Epist}_x\rrbracket^c) =$
       $= \lambda p \exists d'[\text{CR}_x(p) \geq d' \wedge d' > \text{MAX}(\{d : \text{CR}_x(\llbracket\text{Ivan is at home}\rrbracket) \geq d\})]$
       $\llbracket\text{-er than Epist}_x \text{ Ivan is at home Epist}_x\rrbracket^c(\llbracket\text{Ivan is in the office}\rrbracket^c) =$
       $= \exists d'[\text{CR}_x(\llbracket\text{Ivan is in the office}\rrbracket) \geq d' \wedge d' > \text{MAX}(\{d : \text{CR}_x(\llbracket\text{Ivan is at home}\rrbracket) \geq d\})]$

   f. In prose: There is a degree to which x believes Ivan is in the office is
       plausible and this degree is higher than the maximal degree to which
       x believes that Ivan is at home is plausible (where x is the speaker)

Given these assumptions, ECs are deviant with IIEs for the same reason *posibil* is deviant with IIEs in (21) above. That is to say, IIE competes with *-er* for the degree modifier position. (24b) shows a simplified LF for the ill-formed (24a) repeated from above (the underlined part shows the competition).

(24) a. * Ivan bil   po-skoro   v ofisa otkolkoto vkušti.
          Ivan be.PST.PTCP more-soon in office than   home
          Intended: 'Apparently, it is more plausible that Ivan is in the office
          than at home.'   (Bulgarian)

   b. LF: *[[{ [-er than ⟨Ivan is⟩ at home] / Evid } Epist] [Ivan is in the
       office]]

Our account also explains why hearsay evidentials are well-formed with ECs. Several researchers, among which Faller (2002) and Smirnova (2013) have pointed out that hearsay evidentials do not require the speaker's commitment. In our system, this can be implemented by saying that hearsay evidentials are not Epist/Assert modifiers. Therefore, they do not compete with *-er* in ECs for the degree modifier position.

As epistemic modals are not degree modifiers of Epist/Assert, they are felicitous with ECs, see (25a) repeated from above and its simplified LF in (25b).[14] The same holds for Bulgarian in (26).

(25) a. Ion poate   fi   mai degrabă   la birou decât acasă.
         Ion can-3SG be-INF more ADV.early at office than home
         'It is more plausible that Ion might be in the office than at home.'
         (Romanian)

      b. LF: [[[-er than ⟨Ion be⟩ at home] Epist] [might [Ion be in the office]]]

(26) Ivan po-skoro   moje da e v ofisa otkolkoto vkušti.
      Ivan more-soon can   DA is in office than   home
      'It is more plausible that Ivan might be at work than at home.' (Bulgarian)

Independent support for our proposal comes from the fact that ECs are also ill-formed with MAdvs. Recall that according to Greenberg & Wolf (2018), MAdvs are degree modifiers of the gradable Assert. Thus, they are expected to compete with *-er* in ECs, see (27).

---

[14]Independent support for this comes from Irimia (2018) who has shown that there are important structural differences between the IIE reading and the non-IIE modal reading of Romanian PRESMP. Modal interpretations are obtained when the modal features are merged in Mod^0 and raised to T^0. IIE interpretations are obtained by the merge of features related to the speaker's deictic location 'now' in the Sentience projection in the CP layer above the modal in T^0. Note that according to this account IIE features are interpreted higher than modal features. One question would be why examples like (7b) are not well-formed under the future reading of the relevant morpheme. The situation with this auxiliary is more complex. First, not many speakers accept an interpretation of this morpheme which is purely future. For those speakers, though, for which the unmarked future reading is possible, no ill-formedness arises with EC. For the majority of the other speakers, the question is what type of epistemic future this auxiliary encodes that is distinct from both IIE, as well as from a more well-behaved future, but at the same time is also ill-formed with ECs. We leave this issue for further research, as the exact readings need further attention (see also Mihoc 2013).

(27) * Ion posibil   este la birou mai   degrabă decât acasă.
      Ion possibly is   at office more soon   than home
      Intended: 'According to the speaker, it is more plausible that Ion is
      possibly in the office rather than at home.'   (Romanian)

To summarize, by assimilating the epistemic component in ECs to gradable Assert (Greenberg & Wolf 2018), we derive the incompatibility of ECs and IIEs as a result of the competition for the degree modifier position. This proposal assumes that (some) evidentials function as degree modifiers. This correctly predicts the difference between inferential and hearsay indirect evidentials, assuming that the latter does not involve speaker's commitment. We, thus, identify three (overt) elements that can function as degree modifies for Epist/Assert: MAdvs, IIEs, and *-er* in ECs.

4.3 Predictions

Our account makes a number of correct predictions. The first prediction is that the IIE is compatible with regular comparatives. In regular comparatives, IIE scopes above *-er* and the structure is grammatical, as shown in (28a) for Romanian and in (28b) for Bulgarian. In (28a) and (28b), *-er* merges low as it compares degrees of tallness/happiness, rather than degrees of belief as in ECs. The simplified LF for (28a) and (28b) is illustrated in (28c).

(28) a. Ion o   fi mai   înalt decât Maria.
         Ion PRESMP=IIE be more tall   than Mary
         'Apparently, John is taller than Mary.'   (Romanian)

      b. Ivan bil   po-stastliv ot   Maria.
         Ivan be-PST.PTCP=IIE more-happy from Maria
         'Apparently, Ivan is happier than Maria.'   (Bulgarian)

      c. LF: [[Evid$_{IIE}$(Assert)] [[-er [than Mary is d-tall/happy]] [John is d'-tall/happy]]]

The second prediction is that IIE can co-occur with epistemic attitude predicates like 'believe'. This is illustrated in (29a) for Romanian and in (29b) for Bulgarian. We give the simplified LF for these examples in (29c).

(29) a. Ion o   fi crezând   toate minciunile.
         Ion PRESMP=IIE be believe.GER all   lie.the.PL
         'Apparently, Ion believes all the lies.'   (Romanian)

b. Ivan bil   vjarval   na vsički l'ži.
Ivan be.PST.PTCP=IIE believe.PFV on all   lies
'Apparently, Ivan believes all the lies.'   (Bulgarian)

c. [[Evid$_{IIE}$(Assert)] [Ion believes all the lies]]

These data support our account of the ill-formedness of ECs with IIEs. They also rule out alternative analyses according to which the deviance is due either to the incompatibility of evidentials and comparatives or to a potential conflict between evidentials and epistemic attitudes.

5 Concluding remarks and future research

We have analyzed some previously unnoticed facts related to epistemic modals and evidentials when they occur with epistemic comparatives in Bulgarian and Romanian. We showed that ECs are incompatible with IIEs and explained this pattern by claiming that the two categories compete for the same position. Given that the ill-formedness does not arise with epistemic modals, the data examined here argue for accounts under which inferential evidentials are separated from epistemic modals (Faller 2002, Aikhenvald 2014, Murray 2010; a.o.).

From a broader perspective, the observation presented in this paper and its account give rise to several questions. In the remainder of the conclusion we briefly touch on three of them, leaving the detailed investigation for future research.

First, one expectation is that ECs should be ill-formed with indirect evidentials across-the-board. However, there appear to be cases in which ECs are well-formed with expressions that could be analyzed as having evidential meaning.[15] We illustrate some examples below. In (30) and (31) we see that evidential-like adverbials like *vidimo* and *aparent* 'apparently' are well-formed with the EC.[16]

---

[15]We thank Sergei Tatevosov for this observation.

[16]As expected, adverbials with hearsay semantics are well-formed, see below. Recall that hearsay evidentials are not Assert modifiers, thus do not compete with ECs.

(i) Kazvat Ivan po-skoro   e v offisa otkolkoto vkušti.
they.say Ivan more-soon is in office than   home
'As they say, Ivan is at work rather than at home.'   (Bulgarian)

(ii) Cică   Ion este la birou mai   degrabă decât acasă.
they.say John is   at office more soon   than   home
'As they say, John is in the office rather than at home.'   (Romanian)

(30)   ? Vidimo,   Ivan po-skoro   moje da e v ofisa otkolkoto vkušti.
       apparently, Ivan more-soon can   DA is in office than   home
       'It is more plausible that Ivan might be at work than at home.' (Bulgarian)

(31)   Aparent,   Ion este la birou mai   degrabă decât acasă.
       apparently Ion is   at office more soon   than home
       'Apparently, Ion is in the office rather than at home.'   (Romanian)

Examples of this type touch on an important issue, namely the difference between grammaticalized and phrasal evidentials. We take the former to be expressed by means of (inflectional) morphology on the verb. In the latter class we include adverbial evidentials (like *apparently*, etc.) and other phrasal units (like *in my opinion*, etc.), which have evidential semantics, see for example Aikhenvald 2014, a.o. We follow standard accounts for phrasal evidentials as having different syntax from grammaticalized evidentials (Faller 2002, Aikhenvald 2014; a.o.). Thus, the well-formedness of (30) and (31) is not problematic for our account, as lexical evidentials do not compete with -*er* for the ASSERT modifier position.

Second, we also observe that ECs can be embedded under expressions like *I guess*, etc, that are sometimes claimed to have evidential interpretations. Two examples from Romanian are given in (32).

(32)   a. Bănuiesc că   Ion este   la birou mai degrabă decât acasă.
           guess.1SG that Ion be.3SG at office more soon   than home
           'I guess Ion is in the office rather than at home.'

       b. Cred   că   Ion este   la birou mai degrabă decât acasă.
           believe.1SG that Ion be.3SG at office more soon   than home
           'I believe Ion is in the office rather than at home.'   (Romanian)

However, for cases like (32), there is independent evidence that they are biclausal (for example the presence of overt complementizers like *că* 'that'). Therefore, competition does not arise. It is also well known that ECs can be embedded under attitude predicates like *believe, hope*, etc. (see Herburger & Rubinstein 2014, as well as the discussion in footnote 2). We assume that the embedding under *I guess* is amenable to a parallel analysis.[17]

More surprisingly, embedding improves the ungrammaticality of grammaticalized evidentials with ECs. See the contrast in (33a) vs. (33b) and (33c) from

---

[17]We thank an anonymous review for bringing to our attention Czech data that support the same conclusion. We are also grateful to another anonymous reviewer who pointed out to us the connection between embedding under *I guess* and attitude reports.

Romanian. This contrast deserves detailed attention and we leave it for further research.

(33) a. ?? Ion o   fi la birou mai   degrabă decât acasă.
           Ion PRESM.3.SG be at office more soon   than home
           'I guess Ion is in the office rather than at home.'

      b. Bănuiesc că   Ion o   fi la birou mai   degrabă decât acasă.
          guess.1.SG that Ion PRESM.3.SG be at office more soon   than home
          'I guess Ion is in the office rather than at home.'

      c. Cred   că   Ion o   fi la birou mai   degrabă decât
          believe.1.SG that Ion PRESM.3.SG be at office more soon   than
          acasă.
          home
          'I believe Ion is in the office rather than at home.'   (Romanian)

Finally, one of the anonymous reviewers makes the interesting observation that Polish ECs are impossible in negated future contexts. The same point can be made using Romanian data, as seen below:

(34) ?? Ion nu va   fi la birou mai   degrabă decât acasă.
          Ion not FUT be at office more soon   than home
          Intended: 'Ion will not be in the office rather than at home.' (Romanian)

In Bulgarian similar examples seem to be well-formed, see (35).

(35) Ivan ne šte da byde na rabota, a   po-skoro   v kušti.
      Ivan not FUT DA be   at work   but more-soon at home
      'Ivan will not be at work rather than at home.'   (Bulgarian)

However, the future marker *šte* in Bulgarian has been shown to be a versatile category with various types of interpretations (Rivero & Simeonova 2014). Thus, more refined diagnostics are needed to settle this problem. We leave a detailed account of this observation for further research.

Abbreviations

ADV	adverb	INF	infinitive
COND	conditional	MAdv	modal adverbs
DA	modal particle	MAdjs	modal adjectives
EC	epistemic comparative	PFV	perfective
Epist	epistemic	PL	plural
Evid/*evid*	evidential	PRESMP	presumptive
FUT	future	SA	speech act
GER	gerund	PST	past
IE	indirect evidential	PTCP	participle
IIE	inferential indirect evidential	SBJV	subjunctive
		SG	singular

Acknowledgements

References

Aikhenvald, Alexandra Y. 2014. *Evidentiality*. Oxford: Oxford University Press.

Alexiadou, Artemis, Monika Rathert & Arnim von Stechow. 2003. *Perfect explorations*. Berlin: Mouton de Gruyter.

Bresnan, Joan. 1973. The syntax of the comparative clause construction in English. *Linguistic Inquiry* 4(3). 275–343. https://www.jstor.org/stable/4177775.

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

Coşeriu, Eugen. 1976. *Das romanische Verbalsystem*. Tübingen: Narr.

Davis, Christopher, Christopher Potts & Margaret Speas. 2007. The pragmatic values of evidential sentences. In Tova Friedman & Masayuki Gibson (eds.), *SALT 17: Proceedings from the 17th Conference on Semantics and Linguistic Theory*, 71–88. Ithaca, NY. DOI:10.3765/salt.v17i0.2966

Embick, David. 2007. Blocking effects and analytic/synthetic alternations. *Natural Language & Linguistic Theory* 25(1). 1–37. DOI:10.1007/s11049-006-9002-9

Faller, Martina. 2002. *Semantics and pragmatics of evidentials in Cuzco Quechua*. Stanford, CA: Stanford University. (Doctoral dissertation).

Gergel, Remus. 2016. Modality and gradation: Comparing the sequel of developments in *rather* and *eher*. In Elly van Gelderen (ed.), *Cyclical change continued*, 319–350. Philadelphia/Amsterdam: John Benjamins. DOI:10.1075/la.227.11ger

Giannakidou, Anastasia & Melita Stavrou. 2009. On metalinguistic comparatives and negation in Greek. In Claire Halpert, Jeremy Hartman & David Hill (eds.), *MIT Working Papers in Linguistics 57: Proceedings of the 2007 Workshop on Greek Syntax and Semantics at MIT*, 57–74. Cambridge, MA: MITWPL.

Giannakidou, Anastasia & Suwon Yoon. 2011. The subjective mode of comparison: Metalinguistic comparatives in Greek and Korean. *Natural Language & Linguistic Theory* 29(3). 621–655. DOI:10.1007/s11049-011-9133-5

Goncharov, Julie. 2014. Comparing propositional attitudes. In Brian Buccola, Michael D. Hamilton, Alanah McKillen & James Tanner (eds.), *McGill Working Papers in Linguistics 24(1): Proceedings from the 6th Toronto–Ottawa–Montreal (TOM) Workshop in Semantics*, vol. 24, 17–36.

Goncharov, Julie & Monica A. Irimia. 2018. Epistemic comparatives: A cross-linguistic investigation. In Wm. G. Bennett, Lindsay Hracs & Dennis Ryan Storoshenko (eds.), *WCCFL 35: Proceedings of the 35th West Coast Conference on Formal Linguistics*, 169–176. Somerville, MA: Cascadilla Proceedings Project. http://www.lingref.com/cpp/wccfl/35/paper3387.pdf.

Greenberg, Yael & Lavi Wolf. 2018. Gradable assertion speech acts. In Sherry Hucklebridge & Max Nelson (eds.), *NELS 48: Proceedings of the Forty-Eighth Annual Meeting of the North East Linguistic Society*, vol. 1, 271–280. GLSA Publications.

Herburger, Elena & Aynat Rubinstein. 2014. Is 'more possible' more possible in German? In Todd Snider, Sarah D'Antonio & Mia Weigand (eds.), *SALT 24: Proceedings of the 24th Semantics and Linguistic Theory Conference*, 555–576. Ithaca, NY: CLC Publications. DOI:10.3765/salt.v24i0.2717

Irimia, Monica A. 2010. Some remarks on the evidential nature of the Romanian presumptive. In Reineke Bok-Bennema, Brigitte Kampers-Mange, Walter De Moulder & Patrick Dendale (eds.), *Romance languages and linguistic theory 2008: Selected papers from Going Romance*, 125–144. Amsterdam/Philadelphia: John Benjamins.

Irimia, Monica A. 2018. Pragmatics or morpho-syntax? The encoding of indirect evidentiality in Romanian. *Journal of Pragmatics* 128. 148–160. DOI:10.1016/j.pragma.2017.11.015

Izvorski, Roumyana. 1997. The present perfect as an epistemic modal. In Aaron Lawson (ed.), *SALT 7: Proceedings of the 7th Semantics and Linguistic Theory Conference*, 222–239. Ithaca, NY: CLC Publications. DOI:10.3765/salt.v7i0.2795

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

Kennedy, Chris & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–81. DOI:10.1353/lan.2005.0071

Koev, Todor. 2017. Evidentiality, learning events and spatio-temporal distance: The view from Bulgarian. *Journal of Semantics* 34(1). 1–41. DOI:10.1093/jos/ffv014

Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: de Gruyter. DOI:10.1515/9783110126969.7.639

McCawley, James D. 1988. *The syntactic phenomena of English*. Chicago, IL: University of Chicago Press.

Mihoc, Teodora. 2013. *The Romanian presumptive mood: Inferential evidentiality and upper-end degree epistemic modality*. Ottawa: University of Ottawa. (MA thesis).

Morzycki, Marcin. 2011. Metalinguistic comparison in an alternative semantics for imprecision. *Natural Language Semantics* 19(1). 39–86. DOI:10.1007/s11050-010-9063-5

Murray, Sarah E. 2010. *Evidentiality and the structure of speech acts*. New Brunswick, NJ: Rutgers University. (Doctoral dissertation).

Palmer, Frank R. 1986. *Mood and modality*. Cambridge: Cambridge University Press.

Reinheimer Râpeanu, Sanda. 2000. Le présomptif roumain: Marquer evidentiel et épistemique. In Liliane Tasmowski-De Ryck, Matine Coene, Walter De Moulder & Patrick Dendale (eds.), *Traiani Augusti vestigia pressa sequamur: Studia lingvistica in honorem Lilianae Tasmowski*, 481–491. Padova: Unipress.

Rivero, María Luisa & Vesela Simeonova. 2014. An inferential modal in Bulgarian: The inferential future. In Laura Teddiman (ed.), *Proceedings of the 2014 Annual Conference of the Canadian Linguistics Association*, 1–15. Canadian Linguistics Association.

Slave, Elena. 1956. Presumptivul (the presumptive). *Studii de gramatică* 11. 53–60.

Smirnova, Anastasia. 2013. Evidentiality in Bulgarian: Temporality, epistemic modality, and information source. *Journal of Semantics* 30(4). 479–532. DOI:10.1093/jos/ffs017

Squartini, Mario. 2001. The internal structure of evidentiality in Romance. *Studies in language* 25(2). 297–331.   DOI:10.1075/sl.25.2.05squ

Tomić, Olga Mišeska. 2006a. *Balkan Sprachbund morpho-syntactic features*. Dordrecht: Springer.

Tomić, Olga Mišeska. 2006b. The perfect and the evidential. In Olga Mišeska Tomić (ed.), *Balkan Sprachbund morpho-syntactic features*, 333–411. Dordrecht: Springer.

von Fintel, Kai & Angelika Kratzer. 2014. Modal comparatives: Two dilletantes in search of an expert. In Luka Crnić & Uli Sauerland (eds.), *MIT Working Papers in Linguistics 70: The art and craft of semantics: A festschrift for Irene Heim, Vol. 1*, 175–179. Cambridge, MA: MITWPL.

Zobel, Sarah. 2017. "*Eh* ist eh anders" – *eh* and *sowieso* in Federal German and Austrian German. In Clemens Mayr & Edwin Williams (eds.), *Wiener Linguistische Gazette 82: Festschrift für Martin Prinzhorn*, 323–330. Wien: Uni Wien.

Chapter 5

Czech modal complement ellipsis from a comparative perspective

Hana Gruet-Skrabalova

Université Clermont Auvergne,
Laboratoire de recherche sur le langage

This paper deals with modal complement ellipsis in Czech from a comparative perspective. I show that Czech modal complement ellipsis displays a mixed behaviour in comparison with languages like English, Dutch and French. Like English, it allows for various extractions from the ellipsis site and for different subjects in antecedent-contained deletion constructions. Like French and Dutch, it does not allow for intervening elements between the modal verb and the ellipsis site and it requires voice identity of the elided VP and its antecedent. Adopting a deletion approach to ellipsis, I propose to account for this behaviour by parametrizing the syntactic properties of a presumably universal ellipsis feature [E], initially proposed by Lobeck (1995). In my proposal, the syntax of [E] includes the head-licensing ellipsis and the ellipsis site. I argue that the type of licensing head (T, V or Mod) and the type of ellipsis site (VP, TP or VoiceP) induce the properties of modal complement ellipsis that I observe at the surface.

Keywords: modal complement ellipsis, VP ellipsis, modal verbs, auxiliary verbs, ellipsis parameter

1 Introduction

This paper deals with verb-phrase ellipsis that occurs in context of modal verbs, as in (1a). Following Aelbrecht (2008), I will refer to this phenomenon as MODAL COMPLEMENT ELLIPSIS (MCE) in order to distinguish it from a well-known phenomenon of VP ellipsis in (1b) (Ross 1969b; Sag 1976; Merchant 2001; among others). When it comes to other languages than English, it seems relevant to set

apart ellipsis after auxiliary verbs and ellipsis after modal verbs, because the former is not always available in languages that display the latter, as e.g. Romance languages (Busquets & Denis 2001; Depiante 2001; Dagnac 2008; 2010, a.o.) and Germanic languages like German or Dutch (Lobeck 1995; Aelbrecht 2008), see (2) and (3) respectively.[1] The contrast between English on the one hand and Romance and other Germanic languages on the other has been argued to follow from the properties of modal verbs (see §2).

(1) a. John helped them, but Mary could not.

    b. John helped them, but I did not.

(2) a. Jean les   a   aidés,   mais Marie n'a   pas pu.
        J.   them has helped.PL but M.   NEG.has NEG could
        'Jean helped them, but Marie could not.'   (French)

    b. *Jean les   a   aidés,   mais je n'ai   pas.
        J.   them has helped.PL but   I NEG.have NEG
        Intended: 'Jean helped them, but I did not.'   (French)

(3) a. Jan heeft ze   geholpen, maar Maria mocht niet.
        J. has   them helped   but M.   could not
        'Jan helped them, but Maria could not.'   (Dutch)

    b. *Jan heeft ze   geholpen, maar Ik heb niet.
        J. has   them helped   but   I have not
        Intended: 'Jan helped them, but I did not.'   (Dutch)

Interestingly, Czech behaves differently from both English-like and Romance-like languages in that: (i) ellipsis is only partially available after auxiliary verbs, compare (4b) with the past auxiliary and (4c) with the future auxiliary, and (ii) ellipsis after modal verbs in (4a) does not behave entirely like either VP ellipsis in English or like MCE in French or Dutch, as we will see in detail in §4.[2] While the possibility of ellipsis after auxiliary verbs in Czech can be claimed to depend on their morphosyntactic status (see §3.1), I will argue that the mixed properties of Czech MCE follow from both the properties of modal verbs (see §3.2) and the structure targeted by ellipsis. Adopting a deletion approach to ellipsis, I will propose that we can account for (not only) Czech MCE by parametrizing the syntactic properties of a presumably universal ellipsis feature [E], initially proposed by Lobeck (1995) and formalized in Merchant (2001), which determines the licensing head and the selection of the ellipsis site in each language.

---

[1]VP ellipsis after auxiliary verbs is not available in Romance languages, except for Portuguese, see Cyrino & Matos (2002).

[2]See §3.1 for different forms of the auxiliary verb *být* 'be' in Czech.

(4)   a.   Jan jim   pomohl, ale Marie bohužel   nemohla.
           J.   them.DAT helped   but Marie unfortunately NEG.could
           'Jan helped them, but unfortunately Marie could not.'   (Czech)

     b. * Jan jim   pomohl, ale já bohužel   nejsem.
           J.   them.DAT helped   but I   unfortunately NEG.AUX.1SG
           Intended: 'Jan helped them, but unfortunately I did not.'   (Czech)

     c.   Jan jim   bude   pomáhat, ale Marie bohužel
           J.   them.DAT AUX.3SG help   but Marie unfortunately
           nebude.
           NEG.AUX.3SG
           'Jan will help them, but unfortunately Marie won't.'   (Czech)

The paper is organized as follows. §2 briefly presents the main evidence for a deletion approach to MCE. §3 presents the properties of auxiliary and modal verbs in Czech. I discuss here ellipsis after auxiliary verbs and argue that modal verbs behave syntactically neither like T nor like V heads. §4 focuses on the properties of MCE in Czech in comparison with English, French and Dutch. I show that Czech MCE resembles English VP ellipsis in that it allows for various extractions from the ellipsis site and for different subjects in ANTECEDENT-CONTAINED DELETION (ACD) constructions. By contrast, Czech MCE is similar to French and Dutch in that it does not allow for intervening elements between the modal verb and the ellipsis site and it requires voice identity of the elided phrase and its antecedent. §5 proposes a syntactic analysis of this variation based on the mechanism of ellipsis and the syntax of the feature [E], as developed in Aelbrecht (2010). §6 sums up the paper.

2 Assumptions about the syntax of ellipsis

There are two main approaches to ellipsis in the literature, the deletion approach and the null-proform approach, both of which have been applied to VP ellipsis and to MCE. Within the first approach, ellipsis is considered to be deletion or not spelling-out of a fully specified verbal phrase. This analysis is generally assumed for VP ellipsis in English after both auxiliary and modal verbs (Ross 1967; Sag 1976; Hankamer & Sag 1976; Merchant 2001; 2008a, a.o.), see (5a), but it has also been recently argued for Dutch (Aelbrecht 2008; 2010) and Romance (Dagnac 2010). The second type of analysis sees ellipsis as involving a null verbal proform, so-called null-complement anaphora, represented by *e* in (5b). This analysis has been in particular proposed by Depiante (2001) for Spanish and Italian, and by Lobeck (1995) for Dutch.

(5)   a. You can help me, but she can't [VP ~~help me~~].

       b. Je   kan me wel helpen, maar ze   kan niet [VP *e*].
           you can me PRT help   but   she can not
           'You can help me, but she cannot.'   (Dutch)

The main argument in favour of a deletion approach that I will adopt here consists in the possibility of extraction from the ellipsis site. Extraction of the wh-object from the elided VP is possible in English, while it seems impossible in Dutch and Spanish, compare (6a) and (6b)–(6c). However, even if Dutch does not behave exactly like English, Aelbrecht (2008; 2010) shows that at least subject extraction from the elided VP in Dutch in (7) is possible, contrary to object extraction. She argues that this is because MCE in Dutch targets a larger string than VP ellipsis in English, namely VoiceP, which constitutes a phase blocking the object extraction (i.e. when the licensing head is merged, the ellipsis site is sent to PHONETIC FORM (PF) and the site is frozen for extraction).

(6)   a. I don't know who Sue invited, but I know who she couldn't invite.

       b. *Ik weet   wie Katrien moet uitnodigen maar ik weet niet wie ze
            I   know who K.   must invite   but   I   know not who she
            moet niet.
            must not
            Intended: 'I know whom Katrien must invite, but I don't know
            whom she must not.'   (Dutch; Aelbrecht 2008)

       c. *José no   sabe   qué   libro Maria quiere leer, pero Pedro sabe
            J.   not knows which book M.   wants read but   P.   knows
            qué   revisto Anna no pudo.
            which revue   A.   not can
            Intended: 'José doesn't know which book Maria wants to read, but
            Pedro knows which revue Anna can't.'   (Spanish; Depiante 2001)

(7)   a. Deze broek moet niet gewassen worden, maar die   rok   moet wel.
           this pants must not washed become but   that skirt must PRT
           'These pants don't need to be washed, but that skirt does.'
                                                        (Dutch; Aelbrecht 2010)

       b. … maar die rok*ᵢ* moet wel [TP t*ᵢ* [VoiceP t*ᵢ* ~~gewassen worden~~]]

Likewise, Dagnac (2010) shows that even object extraction is possible in Romance if subjects in both clauses are identical, as in (8) and (9). She calls this constraint The Same Subject Constraint. Assuming that modal verbs in Romance are raising

verbs selecting a TP, Dagnac argues that the ellipsis after modal verbs in Romance is not a VP deletion but a TP deletion. This allows to explain e.g. why ellipsis in ACD constructions requires subjects of both elided TP and its TP antecedent to be identical.

(8) a. Maintenant, je sais à qui je peux confier mon fils, mais je ne
         now   I know to who I can entrust my son but I NEG
         sais toujours pas à qui je ne peux pas.
         know still   NEG to who I NEG can NEG
         'Now I know to whom I can entrust my son, but I still don't know to
         whom I can't.'   (French; Dagnac 2010)

     b. … à qui$_i$ je$_j$ ne peux pas [$_{TP}$ t$_j$ [$_{vP}$ t$_j$ ~~confier mon fils~~ t$_i$]]

(9) Ahora, ya   sé   a quién puedo confiar mi hijo, pero todavía
     now   already know.1SG to who can.1SG confide my son   but still
     no sé   a quién no puedo.
     NEG know.1SG to who   not can.1SG
     'Now I know to whom I can confide my son, but I still don't know to
     whom I can't.'   (Spanish; Dagnac 2010)

In addition, an overt pronoun and the verbal anaphor *le faire* 'do it' are ungrammatical with wh-extraction and relativization from the VP in French, see (10a) and (10b) respectively (Dagnac 2008).[3] This also supports the claim that there is a movement from an elided structure, which cannot be reduced to a null pronoun.

(10) a. Il embrasse {qui il peut / *qui il peut le faire / *qui il le peut}.
         he kisses   who he can   who he can it do   who he it can
         'He kisses who he can.'   (French; Dagnac 2008)

---

[3]French modal verbs may combine with an overt pronoun in contexts without extraction. Here, the pronoun can be analysed as a pronominalization of the overt TP complement of the modal verb. These contexts thus constitute arguments neither for deletion, nor for null anaphora. See also §4.1.

    (i)   Jean peut te   répondre, mais moi, je ne (le) peux pas.
          J.   can you answer   but me I NEG it can NEG
          'Jean can answer you, but I can not.'   (French)

    (ii) Je vais résister aussi longtemps que je (le) peux.
          I will resist   as   long   that I it can
          'I will resist as long as I can.'   (French)

b. Léa lit   tous les livres {qu'elle peut / *qu'elle peut le faire   /
    Lea reads all   the books that.she can   that.she can it do
    *qu'elle   le peut}.
     that.she it can
    'Léa reads all the books she can.'   (French; Dagnac 2008)

Dagnac's and Aelbrecht's arguments thus make it very reasonable to assume that there is an underlying syntactic structure in contexts involving MCE, but they also suggest that we need to specify for each language:

1. the type of head licensing ellipsis,

2. the ellipsis site.

I will propose in §5 that these two micro-parameters can be encoded in the syntax of the ellipsis feature [E] responsible for the distribution of ellipsis throughout languages.

3 Auxiliary and modal verbs in Czech

3.1 The auxiliary *být*

Czech is a West-Slavic language with a rich morphology in both its nominal system (number, gender, case) and its verbal system (tense, voice, aspect). It also differs from English, French and Dutch in that (i) it is a subject pro-drop language, (ii) it has second position clitics (2PCl) including pronominal and verbal (auxiliary) clitics, and (iii) it has – despite its basic SVO order – a relatively free word order that reflects the information structure of the clause. Like many other languages, it shows various elliptical constructions, such as gapping and sluicing (Gruet-Skrabalova 2016).

Czech uses only the auxiliary verb *být* 'be', in three series of forms: (i) forms in *js-* in the past tense, (ii) forms in *by-* in the conditional mood, (iii) forms in *bud-* in the future tense. Past and conditional forms are 2PCl, which combine with lexical *-l* participles in the active voice and with the (non-clitic) passive auxiliary *byl* 'been' and lexical *-n* participles in the passive voice, see (11a)–(11b).[4] The future forms are autonomous words and combine with non-finite imperfective verbs in the active voice and with *-n* participles in the passive voice, see (11c)–(11d).

---

[4]Czech morphologically distinguishes active *-l* participles and passive *-n* participles. The former are considered as tensed forms (see Veselovská 1995; 2008), which may bear sentential negation *ne-*.

Veselovská (1995; 2008) argues that Czech clitic auxiliaries bear agreement, but not tense features, and thus that they are generated above T (cf. Roberts 2010, who places Slavic 2PCL in the C-system). By contrast, the non clitic future auxiliary should be merged within the extended VP since it is sensitive to aspect on the lexical verb (cf. Kyncl 2008), as shown in (11c). For the purpose of this paper, I assume that 2PCl auxiliaries are generated in a low C-head, while non clitic future and passive auxiliaries are generated below T (Aspect and Voice respectively, see Cinque 2004), as indicated in (12). I also assume the finite auxiliaries move to T to check their T-feature.

(11) a. Já jsem   {(ne-) četl   / (ne-) přečetl} všechny tyhle knihy.
         I PAST.1SG NEG read.IMPF   NEG PF.read all   these books
         'I (have) (not) read all these books.'   (Czech)

      b. Já jsem   (ne-) byl   pozván.
         I PAST.1SG NEG been invited.PASS
         'I was (not) invited.'   (Czech)

      c. Já budu   {číst   / *přečíst} všechny tyhle knihy.
         I FUT.1SG read.IMPF   PF.read all   these books
         'I will read all these books.'   (Czech)

      d. Já (ne-) budu   pozván.
         I   NEG FUT.1SG invited.PASS
         'I will (not) be invited.'   (Czech)

(12) [CP ... CL [TP ... [AspP bud- [VoiceP byl [VP ...]]]]]

As has already been shown in §1, ellipsis is not available with clitic auxiliaries, see (13b). Gruet-Skrabalova (2012) argues that this follows precisely from their clitic status: 2P clitics cannot license VP ellipsis because they appear too high in the structure with respect to the VP domain. I will return to the analysis of ellipsis after the future auxiliary as in (13a) in §5.2.

(13) a. Já budu   číst nahlas, a   ty budeš   taky.
         I FUT.1SG read aloud   and you FUT.2SG too
         'I will read aloud, and you will too.'   (Czech)

      b. * Já {jsem   / bych}   četl nahlas, a   ty {jsi   / bys}
         I   PAST.1SG COND.1SG read aloud   and you PAST.2SG COND.2SG
         taky.
         too
         Intended: 'I read aloud, and you did too.' / 'I would read aloud, and
         you would too.'   (Czech)

3.2 Modal verbs

There are five strictly modal verbs in Czech: *moci/moct* 'can/be able to', *smět* 'may/be allowed to', *muset* 'must/have to', *nemuset* 'need not', and *mít* 'have to'. These verbs have mixed morphosyntactic properties, as shown in Kyncl (2008): like functional verbs, they have no imperative (**mus*) and no passive (**musen, *mocen*) and they do not combine with aspectual affixes (**domuset, *musívat*). They are not sensitive to the aspectual makeup of the lexical verbs either, contrary to the future auxiliary *budu* requiring imperfective verbs in (11c) above. Like lexical verbs, modal verbs combine with the auxiliary *být* 'be', see (14b)–(14c), and bear the prefix *ne-* expressing sentential negation when they are finite, as we can see in (14a)–(14b). They can be followed by active or passive infinitival verbs, see (14d).

(14) a. Nemůžu   přece   {přečíst / číst}   všechny tyhle knihy.
          NEG.can.1SG though PF.read   read.IMPF all   these books
          'I cannot read all these books.'   (Czech)

      b. Nemohl   jsem   přece {přečíst / číst}   všechny tyhle knihy.
          NEG.could PAST.1SG though PF.read   read.IMPF all   these books
          'I could not read all these books.'   (Czech)

      c. Nikdy nebudu   moci {přečíst / číst}   všechny tyhle knihy.
          never NEG.FUT.1SG can   PF.read   read.IMPF all   these books
          'I will not be able to read all these books.'   (Czech)

      d. Já budu   muset být pozván.
          I FUT.1SG must be invited.PASS
          'I will have to be invited.'   (Czech)

Although modal verbs can occur with auxiliary verbs, modal verbs cannot co-occur, just like in English, and contrary to Romance or Dutch, see (15). The co-occurrence of French and Dutch modal verbs in (15c) and (15d) respectively can be explained if we assume, as has been argued in the literature (Ruwet 1972; Wurmbrand 1999; 2001), that they are raising verbs selecting not a VP, but a TP complement.[5]

(15) a. *Já musím moci přečíst ty   knihy.
          I must   can PF.read these books
          Intended: 'I must be able to read these books.'   (Czech)

---

[5]Traditionally (e.g. Ross 1969a), deontic verbs have been claimed to be control predicates and epistemic verbs to be raising predicates. For Wurmbrand (1999), however, this semantic difference is not represented in syntax.

b. * You must can read these books.

c. Vous devez pouvoir lire ces   livres.
   you must can   read these books
   'You must be able to read these books.'   (French)

d. Hij moet goed kunnen koken.
   he must well can   cook
   'He must be able to cook well.'   (Dutch; Lobeck 1995)

Czech modal verbs also systematically require climbing of pronominal clitic complements of the lexical verb, and thus behave obligatorily like restructuring verbs (Medová 2000). Although clitic climbing is typical for Romance languages (e.g. Rizzi 1978; Roberts 1991), it is no longer true for modal verbs in French.

(16) a. Petr musí přečíst ty   knihy. / Petr (je)   musí (*je)
         P.   must PF.read these books   P.   CL:them must   CL:them
         přečíst.
         PF.read
         'Petr must read these books / them.'   (Czech)

      b. Pierre doit lire ces   livres. / Pierre (*les)   doit (les)   lire.
         P.   must read these books   P.   CL:them must CL:them read
         'Pierre must read these books / them.'   (French)

The properties of Czech modal verbs discussed above suggest that they behave neither like T heads, as in English (Sag 1976 a.o.), nor like V heads, as in French or Dutch (Aelbrecht 2008; Dagnac 2010). Given their restructuring properties, I assume that modal verbs are heads of a specific functional projection ModP between V and T (cf. Cinque 2004) selecting an extended VP projection as complement.

As already said, ellipsis is available after modal verbs, as shown in (17a). Interestingly, ellipsis may occur even if the modal verb follows the future auxiliary and is therefore non-finite, as in (17b). This suggests that MCE is not licensed by the head T. Note that the modal verb does not appear in the first sentence and thus constitutes new information in the elliptical clause.

(17) a. I   když já čtu   nahlas, ty nemusíš.
         even if   I read.1SG aloud   you NEG.must.2SG
         'Although I read aloud, you do not have to.'   (Czech)

      b. I   když já budu   číst nahlas, ty nebudeš   muset.
         even if   I FUT.1SG read aloud   you NEG.FUT.2SG must
         'Although I will have to read aloud, you will not have to.'   (Czech)

4 Modal complement ellipsis in Czech

4.1 English-like properties

This section focuses on Czech MCE in comparison with English, Dutch and French. We will see that Czech MCE looks like English VP ellipsis with respect to extraction and subjects in ACD constructions, and like French and Dutch MCE with respect to the size of the elided string and voice properties of the elided VP and its antecedent.

Starting with non-elliptical constructions, we can see that the verbs 'can' and 'must' in the languages under discussion can have two readings, a deontic (root) reading and an epistemic reading. Czech is similar to English in that both readings are also acceptable in elliptical constructions, although ellipsis appears most frequently with the deontic reading. In contrast, it has been observed for Romance and Dutch that MCE is only available with deontic reading of these modal verbs:[6]

(18) Deontic reading

    a. John couldn't come to the party, and Peter was not allowed.

    b. Jan na večírek přijít nemohl   a   Petr nesměl.
        J.   to party   come NEG.could and P.   NEG.could
        'Jan was not able to come to the party, and Petr was not allowed.'
        (Czech)

    c. Jean a   pu   venir à la soirée, mais Pierre n'a   pas pu.
        J.   has could come to the party but P.   NEG.has NEG could
        'Jean could come to the party, but Pierre couldn't.'   (French)

    d. Jan kon   naar het feest kommen, maar Piet mocht niet.
        J.   could to   the party come   but P.   could not
        'Jean could come to the party, but Piet was not allowed.'   (Dutch)

(19) Epistemic reading

    a. It can be true, but it doesn't have to.

    b. Může to být pravda, ale nemusí.
        can it be true   but NEG.must
        'It can be true, but it doesn't have to.'   (Czech)

---

[6]Cf. Barbiers (1995); Lobeck (1995); Aelbrecht (2008). For my informants, the verb *moeten* 'must' would be ruled out in (18d), the verb *hoefen* 'should' being more acceptable.

    c. * Cela peut être vrai, mais cela ne   doit   pas.
         this can be   true but   this NEG must NEG
         Intended: 'It can be true, but it doesn't have to.'   (French)

    d. ? Het zou   waar kunnen zijn, maar het hoeft   niet.
         it   should true can   be but it   should not
         'It should be true, but it shouldn't have to.'   (Dutch)

Another property Czech shares with English concerns the possibility of pronominalizing the elided string. Actually, missing material after modal verbs in Czech is not in complementary distribution with an overt pronoun, as shown in (20). French and Dutch behave differently except for contexts with extraction like in (21); see §2, footnote 3. This might be not completely surprising if both pronominalization and MCE in these languages target a TP, as proposed by Dagnac (2010); see §5.1.

(20) a. John will answer you, but I can't (*it).

      b. Jan ti   odpoví, ale já (*to) nemůžu.
         J. you.DAT answers but I   it NEG.can
         'Jan will answer you, but I can't.'   (Czech)

      c. Jean te   répondra,   mais moi, je ne   (le) peux pas.
         J.   you answer.FUT.3SG but me I NEG it can NEG
         'Jean will answer you, but I can't.'   (French)

      d. Jan zal je   antwoorden, maar ik kan (het) niet.
         J.   will you answer   but I can it   not
         'Jan will answer you, but I can't.'   (Dutch)

(21) a. Jean lit   tous les livres qu'il   (*le) peut.
         J.   reads all   the books that he   it can
         'Jean reads all the books he can.'   (French)

      b. Joris leest elk   boek dat hij (*het) kan.
         J.   reads every book that he it   can
         'Joris reads every book he can.'   (Dutch)

Elliptical relative clauses, so-called antecedent contained deletion (ACD), display another property in which languages may differ. In Czech and English, a relative clause containing ellipsis and its matrix clause may have different subjects. In

contrast, Romance and Dutch require both subjects to be identical (the Same Subject Constraint, see §2):[7]

(22) a.   John reads all the books that Mary can't.

    b.   Jan čte   všechny knihy, které Marie nesmí.
        J.   reads all   books that M.   NEG.can
        'Jan reads all books that Marie is not allowed to read.'   (Czech)

    c.   *Jean lit   tous les livres que Marie ne   peut pas.
        J.   reads all   the books that M.   NEG can NEG
        Intended: 'Jean reads all the books that Marie can't.'
                                                (French; Dagnac 2008)

    d.   *Joris leest elk   boek dat Monika moet niet.
        J.   reads every book that M.   must not
        Intended: 'Joris reads every book that Monika doesn't have to.'
                                            (Dutch; Aelbrecht 2008)

Finally, recall that the main argument for the deletion approach of MCE is based on extraction from the elided string. In Czech, several types of extraction are possible, to both A- and A′- positions. Extraction to subject position can be seen with the inaccusative verb přijít 'come' in the example (16a) above, repeated in (23).

(23) *Extraction to subject position:*

    Jan na večírek přijít nemohl   a   Petr nesměl.
    J. to party   come NEG.could and P.   NEG.could
    'Jan couldn't come to the party and Petr was not allowed.'   (Czech)

The example (24) shows regular wh-object extraction of the dative wh-word *komu* 'to whom' to the CP domain (cf. §2, ex. (7)), and the example (25) shows topicalization of an accusative DP-object. It must, however, be noted that extraction from VP in English, like in (25a), usually requires a specific contrastive focus

---

[7]Here is an attested and more natural example for (22b):

(i)   Helenku   bolelo břicho a   tak Elizabetka snědla všechny bonbóny který
     Helenka.ACC ached stomach and so Elizabetka ate   all   sweets   that
     Helenka nemohla.
     Helenka NEG.could
     'Helenka had a stomach ache, so Elizabetka ate all sweets that Helenka could not.'
                                              (Czech)

(Schuyler 2001; Lasnik 2001; Merchant 2008b). This is also true for French, where topicalization is acceptable provided the context is contrastive enough. In Czech, no specific prosody is required to accompany extraction in the examples below. I assume that this is because the word order in Czech is much is much freer than in English and French and serves to mark a specific information status of a phrase. In English, the information focus or topic status is generally marked by prosody. With respect to extraction, Czech may thus seem even more permissive than English.

(24) *Wh-object extraction:*

Vím   komu   můžu děti   svěřit   a   komu   nemůžu.
know.1SG who.DAT can.1SG children confide and who.DAT NEG.can.1SG
'I know to whom I can confide my children and to whom I can't.' (Czech)

(25) *Object topicalization:*

   a. The toy she had already bought, but the book she really couldn't.

   b. Hračky už   jsem   koupila, ale knížky   jsem   ještě
toys.ACC already PAST.1SG bought   but books.ACC PAST.1SG yet
nemohla.
NEG.could
'I have already bought the/some toys, but I was not yet able to buy the/some books.'   (Czech)

   c. Les jouets, je les   ai   déjà   achetés, mais les livres, je
the toys   I   them have already bought but   the books I
n'ai   pas encore pu.
NEG.have NEG yet   could
'I have already bought the toys, but I haven't yet been able to buy the books.'   (French)

   d. * Het speelgoed had   ik al   gekocht, maar het boekje kon   ik
the toy   have I already bought but   the book   could I
niet.
not
Intended: 'I have already bought the toy, but I haven't yet been able to buy the book.'   (Dutch)

The last example in this section presents a more questionable extraction: in English, (26a) is generally considered as a case of pseudogapping, but some also analyse it as involving movement (Aelbrecht 2008; Gengel 2013). Whether or

not we assume that (26b) involves an object extraction, Czech would again be similar to English rather than to French or Dutch.

(26)   a.   I will vote for him, and you can for her.

   b.   Já budu   volit pro něj a   ty   můžeš pro ni.
        I   FUT.1SG vote for him and you can   for her
        'I will vote for him and you can for her.'   (Czech)

   c.   *Je voterai   pour lui, et   tu   peux pour elle.
        I   vote.FUT.1SG for   him and you can for   her
        Intended: 'I will vote for him and you can for her.'
                                                    (French; Dagnac 2010)

   d.   *Ik zal   voor hem stemmen, en   je   kan voor haar.
        I   will for   him vote   and you can for   her
        Intended: 'I will vote for him and you can for her.'
                                                    (Dutch; Aelbrecht 2008)

4.2 Differences from English

Despite several similarities with English VP ellipsis presented in the previous subsection, Czech MCE is also (at least to some extent) similar to French and Dutch MCE.

First, we observe that verbal elements intervening between modals and the VP must be elided, see the passive auxiliary 'be' in (27). In French and Dutch, these elements also include negation and aspect morphemes, which in Czech appear directly on the verb stem.[8]

(27)   a.   This text can be read aloud but this poetry really cannot (be).

---

[8]See also these two examples:

   (i)   *Paul peut avoir fini   en juin, et   Luc peut aussi avoir [fini   en juin].
         Paul can have finished in June and Luc can too   have finished in June
         Intended: 'Paul can have finished in June and Luc can have.'   (French; Dagnac 2008)

   (ii) *Paul peut repasser LM01 et   Luc peut ne pas [repasser LM01].
         Paul can take   LM01 and Luc can not   take   LM01
         Intended: 'Paul is allowed to take LM01 and Luc is allowed not to.'
                                                    (French; Dagnac 2008)

b. Tento text může být čten   nahlas, ale tato báseň skutečně
this   text can   be   read.PASS aloud   but this poetry really
nemůže (*být).
NEG.can   be
'This text can be read aloud but this poetry really cannot be.' (Czech)

c. Ce texte peut être lu   à voix haute, mais ce   poème ne   peut
this text can be   read in voice high   but this poetry NEG can
vraiment pas (*être).
really   NEG   be
'This text can be read aloud but this poetry really cannot be.'

(French; Dagnac 2010)

d. Deze tekst kan hardop gelezen worden, maar deze poëzie kan echt
this   text can aloud read   be   but   this poetry can really
niet (*zijn).
not   be
'This text can be read aloud but this poetry cannot be.'   (Dutch)

By contrast, second position clitic auxiliaries are obligatory with MCE in Czech, see (28a), which is not surprising since they occur very high in the clause (cf. §3.1).[9] Note also that second position pronominal clitics that are complements of the lexical verb are excluded. These clitics normally appear on modal verbs, see (16a) above and (28b). This suggests that pronominal clitics must be elided before clitic climbing.[10]

(28) a. Já jsem   to   musela podepsat, ale ty   *(jsi)   (*to)
I PAST.1SG it.CL had.to sign   but you PAST.2SG   it.CL
nemusel.
NEG.had.to
'I had to sign it, but you didn't have to.'   (Czech)

b. Já jsem   to   musela (*to) podepsat (*to).
I PAST.1SG it.CL had.to   it.CL sign   it.CL
'I had to sign it.'   (Czech)

---

[9]Contrary to MCE, auxiliary clitics cannot escape sluicing, as has been noted by Merchant (2001) in his Sluicing-Comp Generalization. This follows if MCE targets a smaller structure than sluicing does.

[10]For Roberts (2010), pronominal clitics consistently escape the interior of the low *v*-cycle. If this is true, MCE in Czech targets a larger structure than the low *v*-cycle, which is compatible with my analysis in §5.

Finally, it has been pointed out (Hardt 1993; Merchant 2008a; 2013) that the voice of the elided VP and that of its VP antecedent in English may differ. In (29a), the elliptical clause is in the active voice, while the clause with the VP antecedent is in the passive voice. The example (29b) shows the opposite distribution: passive elliptical clause and active clause containing the antecedent.[11] Assuming that voice is encoded on the head Voice and that Voice is distinct from the head *v*, Merchant (2008a) argues that VP ellipsis in English targets a verbal phrase (*v*P/VP) below Voice head. Ellipsis therefore does not include VoiceP.

(29) a. This problem had to be solved long ago, but obviously nobody could (solve it).   (Merchant 2008a)

    b. The janitor must remove the trash whenever it is apparent that it should be (removed).   (Merchant 2008a)

Contrary to English, we can see in (30) that Czech, French and Dutch require the same active or passive morphosyntax in both the elliptical clause and the clause containing the VP antecedent.[12] Ellipsis is excluded here because the elliptical clause is presumably in the active voice, while the clause containing the antecedent is in the passive voice. This suggests that VoiceP is included in the ellipsis site.

(30) a. Ten problem měl   být dávno   vyřešen   ale nikdo zřejmě
       this problem had.to be   longtime solved.PASS but nobody obviously
       nemohl   *(ho vyřešit).
       NEG.could   it solve
       'This problem had to be solved long ago, but nobody could solve it.'
                                     (Czech)

    b. Ce problème aurait   déjà dû   être résolu, mais personne
       this problem   has.COND yet   had.to be   solved but   nobody
       n'a   pu   *(le résoudre).
       NEG.has could   it solve
       'This problem had to be solved long ago, but nobody could solve it.'
                         (French; Dagnac 2010)

    c. Dit probleem had al   lang geleden opgelost moeten worden
       this problem   had already long ago   solved   must   be

---

[11]Apparent counter-examples in English are reanalysed by Merchant (2013) as cases of pseudogapping.

[12]Voice identity also applies to ellipsis after the auxiliary future verb, see §5.2.

maar niemand kon   *(het opgelessen).
but   nobody could   it   solve
'This problem had to be solved long ago, but nobody could solve it.'

(Dutch)

Given the properties discussed above, I conclude that Czech MCE seems to target a larger structure that VP ellipsis in English, but probably a smaller structure than MCE in French or Dutch.

4.3 Summary

Table 1 summarizes the properties discussed in this section. Rows 1 to 3 indicate for each language whether it allows ellipsis after auxiliary verbs (Aux + ellipsis), co-occurrence of an auxiliary and a modal verb (T + Mod), and co-occurrence of two modal verbs (Mod + Mod). With respect to MCE, rows 4 to 8 indicate whether it is compatible with deontic and epistemic reading of modal verbs (Deont/Epist reading), with subject extraction, object topicalization, wh-object extraction and object scrambling (if there is any). Finally, rows 9 to 12 show whether MCE requires identical subjects in ACD constructions (Same Subject Constraint) and identical voice on the verb (Voice identity) and whether it allows a passive auxiliary to occur after the modal verb (Passive Aux). We can see that Czech shares most but not all the examined properties with English.

5 Proposal

Before proposing an analysis of MCE in Czech, I present here my assumptions about the general mechanism licensing ellipsis, following in particular Aelbrecht (2010), and to some extent Lobeck (1995), Merchant (2001) and Craenenbroeck & Lipták (2013). I thus assume that (i) ellipsis is triggered in a checking relation (Agree) between the licensing head X and the ellipsis site YP, (ii) there is a feature [E] that occurs on the head of Y and indicates that YP will not be spelled out (non-pronunciation at PF) once the feature is checked out by the head X.[13]

The feature [E] has a specific syntax consisting of two properties: (i) selection of the head on which the feature may occur, i.e. the head of the constituent that will be elided (SEL X), (ii) uninterpretable features that must be checked against the features of the head licensing ellipsis (uY). I will propose that parametrizing

---

[13]This assumption is at variance with Merchant (2001), for whom it is the complement of the head bearing [E] that is elided.

Table 1: Properties related to MCE

		English	Czech	French	Dutch
1	Aux + ellipsis	Yes	Yes/No (2PCl)	No	No
2	T + Mod	No	Yes	Yes	Yes
3	Mod + Mod	No	No	Yes	Yes
4	Deont/Epist reading	Yes/Yes	Yes/Yes	Yes/No	Yes/No?
5	Subject extraction	Yes	Yes	Yes	Yes
6	Object topicalization	Yes	Yes	Yes	No
7	Wh-object extraction	Yes	Yes	Yes	No
8	Object scrambling (?)	Yes	Yes	No	No
9	Same Subject Constraint	No	No	Yes	Yes
10	Overt pronoun	No	No	Yes/No	Yes/No
11	Voice identity	No	Yes	Yes	Yes
12	Passive Aux	Yes	No	No	No

these two properties accounts for the behaviour of (not only) Czech MCE. The elided YP must be given (Barbiers 1995; Lobeck 1995; Merchant 2001) and syntactically/structurally isomorphic with its antecedent (cf. Fiengo & May 1994).

5.1 Specifying the properties of the [E]-feature in MCE

We have seen in §3 that the behaviour of modal verbs in Czech suggests that they are neither T nor V head, but rather head of a specific functional projection between T and V, which I call Mod. They can therefore co-occur with T but not with other modal verbs.

To account for ellipsis licensed in the contexts of modal verbs in Czech, I propose in (31a) that the feature [E] is merged on the head Voice, i.e. that it selects as ellipsis site the phrase headed by Voice (sᴇʟ Voice). Moreover, the feature [E] must have its uninterpretable features (uMod) checked out by the head Mod, i.e. it is licensed by Mod. The properties of the [E] feature in Czech would differ from the properties of the [E] feature in English, French and Dutch respectively, as shown in (31b)–(31d). In English, ellipsis targets *v*P and is licensed by T (see Merchant 2008a). In French, ellipsis targets TP and is licensed by modal V selecting a TP (see Dagnac 2008). In Dutch, ellipsis targets VoiceP but it is licensed by a deontic V (see Aelbrecht 2008). Contrary to Dutch, however, we do not need to

postulate that the VoiceP in Czech constitutes a phase blocking object extraction, since both subject and object extractions may take place before VoiceP is sent to PF.

(31) The syntax of [E] feature in MCE: (cf. Merchant 2008a)
    a. Czech: E_{MCE} [INFL [uMod], SEL [Voice]]
    b. English: E_{MCE} [INFL [uT], SEL [v]]   (see Merchant 2008a)
    c. French: E_{MCE} [INFL [uV], SEL [T_{nf}]]   (see Dagnac 2008)
    d. Dutch: E_{MCE} [INFL [uV_{deon}], SEL [Voice]]   (see Aelbrecht 2008)

(32) a. Czech: [$_{ModP}$ může ~~[$_{VoiceP}$ [$_{vP}$ t$_{subj}$ [$_{VP}$ …]]]~~]
    b. English: [$_{TP}$ can [$_{AspP}$ (have) [$_{VoiceP}$ (been) [$_{vP}$ t$_{subj}$ ~~[$_{VP}$ …]~~]]]]
    c. French: [$_{VP}$ peut ~~[$_{TP}$ t$_{subj}$ [$_{AspP}$ [$_{vP}$ t$_{subj}$ [$_{VP}$ …]]]]~~]
    d. Dutch: [$_{VP}$ kan [$_{TP}$ t$_{subj}$ ~~[$_{VoiceP}$ [$_{vP}$ t$_{subj}$ [$_{VP}$ …]]]~~]]
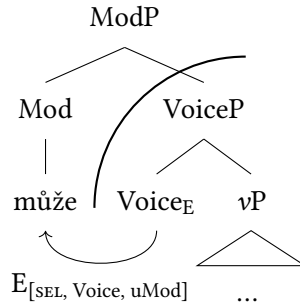
Figure 1: Syntactic structure for (32a)

The proposed analysis can account for the properties of Czech MCE as follows. First, MCE requires voice identity in both the elliptical clause and its antecedent, i.e. both clauses must be either active or passive. Assuming that the parallelism requirement on ellipsis includes voice features, postulating that the feature [E] targets VoiceP guarantees that ellipsis takes place only if elided and antecedent VoiceP are identical. Furthermore, since VoiceP is neither a nominal nor a clausal phrase, it follows that it cannot be pronominalized by an overt pronoun.

Second, MCE does not target the clitic and the future auxiliaries, but it cannot leave aside the passive auxiliary. Since clitic auxiliaries are generated high in the structure, the analysis predicts that they will not be included in the ellipsis site. Likewise, the future auxiliary generated above modal verbs will not be elided,

see §3.2, ex. (17b). In contrast, the passive auxiliary located in the VoiceP will be elided along with the VoiceP.

Third, MCE allows extraction of focused (wh-object) and contrastively focused XPs (contrastive topics). Since elided elements are informationally given, it follows that only focused XPs can escape ellipsis and undergo extraction. This is especially visible in the case of pronominal object clitics, which cannot be focused, and will thus never be allowed to escape the ellipsis site. Extraction of non-identical XPs from the ellipsis site could, however, be viewed as problematic for parallelism constraints assumed for deletion, although these constraints do not mean full morphophonological identity. I thus propose to assume with Merchant that focus overrides "identity condition" in deletion (Merchant 2001). In the case of subject extraction from vP to TP, for instance, the identity required for deletion reduces to the type of argument (referential DP), but it does not concern the meaning or the reference of the DP subjects themselves. In the case of ACD, we can consider that the subject of the relative clause must escape deletion precisely because it is contrasted with the subject on the main clause.

(33) a. Jan čte   všechny knihy, které$_i$ Eva nesmí   (číst t$_i$).
         Jan reads all   books that   Eva NEG.can read
         'Jan reads all the book that Eva can't.'   (Czech)

      b. Jan čte [$_{VoiceP}$ [$_{VP}$ t$_{sub}$ t$_v$ všechny knihy [$_{CP}$ které$_i$ [$_{TP}$ Eva [$_{ModP}$ nesmí [$_{VoiceP}$ [$_{vP}$ t$_{sub}$ [$_{VP}$ … t$_i$]]]]]]]]]

The observation that extraction is relatively easy in Czech can be related to the monoclausal structure of sentences with modal verbs. In the case of an intermediate extraction (if we assume objects scrambling instead of pseudogapping), like in (34), we can suppose that extracted elements are hosted by a TP-internal Focus position (following Belletti 2004) between the modal verb and the elided VoiceP. This kind of extraction would not be available in French or Dutch, where ellipsis targets the TP complement of the modal V.

(34) Já třeba   napíšu   Heleně básničku, a   ty   zas   můžeš písničku.
      I maybe PF.write H.DAT poem.ACC and you then can   song.ACC
      'I might write a little poem for Helen, and you, you can write a little song for her.'   (Czech)

5.2 Extending the analysis to ellipsis after the future auxiliary

Since English modal and auxiliary verbs represent the same kind of head (Ross 1967), it is not surprising that they behave alike with respect to ellipsis. We can
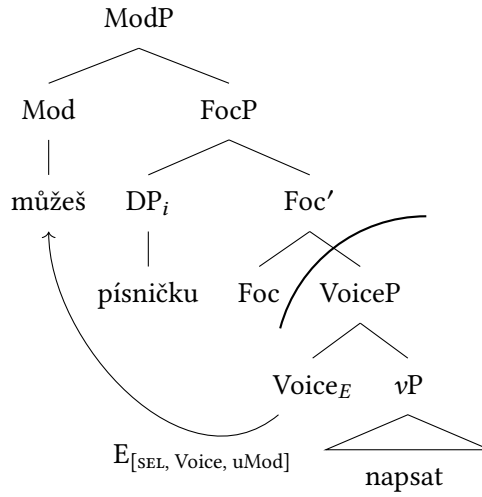
Figure 2: Syntactic structure for (34)

thus reduce the analysis of MCE to the analysis of VP ellipsis, by having the same [E] feature for both. In French and Dutch, modal and auxiliary verbs are syntactically different heads and behave differently with respect to ellipsis. If these languages only possess the [E] feature with the syntax given above, ellipsis after auxiliary verbs will be excluded since auxiliary verbs are not V heads and do not have a non-finite TP complement.

As for Czech, I propose that the analysis in terms of VoiceP-ellipsis can be extended the future auxiliary because: (i) the future auxiliary is a functional verbal head between V and T, (ii) its complement is an extended VP, and (iii) the ellipsis also requires voice identity:

(35) a. Udělali to, kdykoliv {museli   (~~to udělat~~) / *to muselo   (
         did.PL it whenever had.to.PL it do   it had.to
         ~~být uděláno~~)}.
         be done.PASS
         'They did it whenever they did have to (do it).'   (Czech)

     b. Měl   být operován,   ale {nebude   (~~operován~~)   /
         had.to be operated.PASS but NEG.FUT.3SG operated.PASS
         *nebudou   (~~ho operovat~~)}.
          NEG.FUT.3PL him operate
         'He had to be operated but he will not (be operated).'   (Czech)

To allow both the future auxiliary and modal verbs to license ellipsis, I suggest defining the [E] feature as follows:

1. the licensing head is a functional verbal head F_v, that can be realized as Mod and Asp,

2. the uninterpretable features of [E] are uF_v, see Figure 3.

(36) Měl   být operován,   ale pro$_i$ [$_{AspP}$ nebude
      had.to be operated.PASS but   NEG.FUT.3SG
      [$_{VoiceP}$ PASS [$_{VP}$ operován t$_i$]]].
                  operated.PASS
      'He had to be operated but he will not (be operated).'   (Czech)
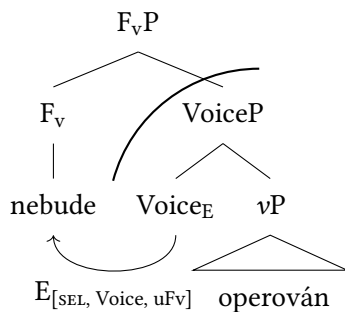
Figure 3: Syntactic structure for (36)

6 Conclusion

In this paper, I have dealt with MCE in Czech from a comparative perspective. I have shown that Czech modal verbs and Czech MCE exhibit a mixed behaviour with respect to languages like English, French, and Dutch. Like English, it allows for various extractions from the ellipsis site and for different subjects in ACD constructions. Like French and Dutch, it does not allow for intervening elements between modal verb and ellipsis site and it requires voice identity of the elided VP and its antecedent.

I have argued that the properties of MCE that we observe at the surface are induced by the head licensing ellipsis (F_v, V, or T) and the ellipsis site (VoiceP,

TP, or VP). Adopting a deletion account of MCE based on a presumably universal ellipsis feature [E], I have undertaken to parametrize this feature [E], whose properties include precisely the licensing head and the ellipsis site. In addition, the properties of this [E] feature also imply whether ellipsis is available with auxiliary verbs. As for Czech, I have proposed that [E] is licensed by a functional verbal head F_v, which can be realized by modal verbs or the future auxiliary and targets VoiceP.

The question remains whether we can relate the parametrization of the [E] feature to other language properties. One hypothesis to explore can be found in Cyrino & Matos (2002), who claim that there is a correlation between the possibility of verbal ellipsis (after both auxiliary and modal verbs) and the structure of the extended verbal projection, in particular the realization of aspect. This issue is however outside the scope of this paper and must be left to further investigation.

Abbreviations

1, 2, 3	first, second, third person	Mod	modal verb
2PCl	second position clitics	MCE	modal complement ellipsis
ACC	accusative		
ACD	antecedent-contained deletion	NEG	negation
AUX	auxiliary	PASS	passive
CL	clitic	PAST	past tense
COND	conditional	PF	phonetic form
DAT	dative	PF	perfective
deont/epist	deontic/epistemic	PL	plural
FUT	future tense	PRT	particle
IMPF	imperfective	SG	singular

Acknowledgements

References

Aelbrecht, Lobke. 2008. Dutch modal complement ellipsis. In Olivier Bonami & Patricia Cabredo-Hofherr (eds.), *Empirical issues in syntax and semantics 7*, 7–33. http://www.cssp.cnrs.fr/eiss7/aelbrecht-eiss7.pdf.

Aelbrecht, Lobke. 2010. *The syntactic licensing of ellipsis.* Amsterdam: John Benjamins.

Barbiers, Sjef. 1995. *The syntax of interpretation.* Leiden: University of Leiden. (Doctoral dissertation).

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: The cartography of syntactic structures*, 16–51. Oxford: Oxford University Press.

Busquets, Joan & Pascal Denis. 2001. L'ellipse modale en français: le cas de *devoir* et *pouvoir. Cahiers de grammaire* 26. 55–74. https://halshs.archives-ouvertes.fr/halshs-01648372.

Cinque, Guglielmo. 2004. Restructuring and functional structure. In Adriana Belletti (ed.), *Structures and beyond*, 132–191. New York: Oxford University Press.

Craenenbroeck, Jereon van & Aniko Lipták. 2013. What sluicing can do, what it can't, and in which Language: On the cross-linguistic syntax of ellipsis. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Diagnosing syntax*, 486–502. Oxford: Oxford University Press.

Cyrino, Sonia & Gabriela Matos. 2002. VP-ellipsis in European and Brazilian Portuguese - a comparative analysis. *Journal of Portuguese Linguistics* 1(2). 177–195. DOI:10.5334/jpl.41

Dagnac, Anne. 2008. L'ellipse modale en français: Arguments pour ellipse du TP. In Jacques Durand, Benoît Habert & Bernard Laks (eds.), *Actes du CMLF'08*, 2453–2465. Paris: ILF. DOI:10.1051/cmlf08211

Dagnac, Anne. 2010. Modal ellipsis in French, Spanish and Italian: Evidence for a TP-deletion analysis. In Karlos Arregi (ed.), *Romance linguistics 2008: Interactions in Romance*, 157–170. Amsterdam: John Benjamins.

Depiante, Marcela. 2001. On null complement anaphora in Spanish and Italian. *Probus* 13(2). 193–221. DOI:10.1515/prbs.2001.003

Fiengo, Robert & Robert May. 1994. *Indices and identity.* Cambridge, MA: MIT press.

Gengel, Kirsten. 2013. *Pseudogapping and ellipsis.* Oxford: Oxford University Press.

Gruet-Skrabalova, Hana. 2012. VP-ellipsis and the Czech auxiliary *být* ('to be'). *Journal XLinguae* 5(4). 2–16.

Gruet-Skrabalova, Hana. 2016. VP-Elipsa. In Petr Karlík, Marek Nekula & Jana Pleskalová (eds.), *Nový encyklopedický slovník češtiny* [New encyclopedic dictionary of Czech]. Prague: Lidové noviny.

Hankamer, Jorge & Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7(3). 391–428. https://www.jstor.org/stable/4177933.

Hardt, Daniel. 1993. *Verb phrase ellipsis: Form, meaning and processing*. Philadelphia, PA: University of Pennsylvania. (Doctoral dissertation).

Kyncl, Jaroslav. 2008. *Semi-lexical heads in Czech modal structures*. Wolverhampton: University of Wolverhampton. (Doctoral dissertation).

Lasnik, Howard. 2001. When can you save a structure by destroying it? In Min-yoo Kim & Uri Strauss (eds.), *NELS 31: Proceedings of the 31st Annual Meeting of the North East Linguistic Society*, 301–320. Amherst, MA: GLSA Publications.

Lobeck, Anne. 1995. *Ellipsis: Functional heads, licensing and identification*. Oxford: Oxford University Press.

Medová, Lucie. 2000. *Transparency phenomena in Czech syntax*. Tromsø: University of Tromsø. (MA thesis).

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.

Merchant, Jason. 2008a. An asymmetry in voice mismatches in VP ellipsis and pseudogapping. *Linguistic Inquiry* 39(1). 169–179. DOI:10.1162/ling.2008.39.1.169

Merchant, Jason. 2008b. Variable island repair under ellipsis. In Kyle Johnson (ed.), *Topics in ellipsis*, 132–153. Cambridge: Cambridge University Press.

Merchant, Jason. 2013. Voice and ellipsis. *Linguistic Inquiry* 44(1). 77–108. DOI:10.1162/LING_a_00120

Rizzi, Luigi. 1978. A restructuring rule in Italian syntax. In Samuel J. Keyser (ed.), *Recent transformational studies in Eropean languages*, 113–158. Cambridge, MA: MIT Press.

Roberts, Ian. 1991. Restructuring, head movement, and locality. *Linguistic Inquiry* 28(3). 423–460. https://www.jstor.org/stable/4178986.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incroporation, and defective goals*. Cambridge, MA: MIT Press.

Ross, John R. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT. (Doctoral dissertation).

Ross, John R. 1969a. Auxiliaries as main verbs. In William Todd (ed.), *Studies in philosophical linguistics*, 77–102. Evanston, IL: Great Expectations Press.

Ross, John R. 1969b. Guess who? In Robert Binnick, Alice Davison, Georgia Green & James Morgan (eds.), *Papers from the 5th Meeting of Chicago Linguistic Society*, 252–286. Chicago: Chicago Linguistic Society.

Ruwet, Nicolas. 1972. *Théorie syntaxique et syntaxe du français*. Paris: Editions du Seuil.

Sag, Ivan. 1976. *Deletion and logical form*. Cambridge, MA: MIT. (Doctoral dissertation).

Schuyler, Tamara. 2001. Wh-movement out of the site of VP ellipsis. In Séamas Mac Bhloscaidh (ed.), *Syntax and Semantics at Santa Cruz, Volume 3*, 1–20. Santa Cruz: University of California.

Veselovská, Ludmila. 1995. *Phrasal movement and X° morphology: Word order parallels in Czech and English nominal and verbal projections.* Olomouc: Palacký University. (Doctoral dissertation).

Veselovská, Ludmila. 2008. The extended verbal projection in Czech: Three variants of the verb *být*. In Gerhild Zybatow, Luka Szucsich, Uwe Junghanns & Roland Meyer (eds.), *Formal Description of Slavic Languages: The Fifth Conference, Leipzig 2003*, 555–569. Frankfurt am Main: Peter Lang.

Wurmbrand, Susanne. 1999. Modal verbs must be raising verbs. In Sonya Bird, Andrew Carnie, Jason D. Haugen & Peter Norquest (eds.), *WCCFL 18: Proceedings of the 18th West Coast Conference on Formal Linguistics*, 599–612. Somerville, MA: Cascadilla Press.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure.* Berlin: Mouton de Gruyter.

Chapter 6

Czech infinitival conditionals

Uwe Junghanns
University of Göttingen

Hagen Pitsch
University of Göttingen

The paper deals with complex Czech sentences that consist of a finite matrix clause and an infinitival subordinate clause. The latter receives a conditional interpretation although there is no item to signal that function. Interestingly, an overt nominative subject can be realized within the infinitival clause. Conditional sentences of this type occur in Czech and Slovak only. The following issues are addressed in the paper: size of the infinitival clause, connection with the matrix, surface order, nominative case licensing, and interpretation. The proposal builds on the idea that finiteness is a conspiracy of tense and agreement marking (see Stowell 1982; 1995; Wurmbrand 2001). Infinitival structures come in two variants: tensed vs. untensed. Nominative subjects are realized in the former but not the latter. We assume an economical division of labour between null and overt subjects. The proposal invokes un/interpretable un/valued T-/Φ-features for an explanation (see Pesetsky & Torrego 2001). The issues of how the conditional interpretation comes about and what might possibly help to reach it are addressed as well.

Keywords: infinitives, conditionals, complex sentences, syntax, interpretation, nominative case licensing, non/finiteness, Czech

1 Introduction

Within the Slavic branch, Czech and Slovak are the only languages to exhibit a peculiar non-finite, clause-like structure. It is used and interpreted as a conditional clause. Therefore, we will call it INFINITIVAL CONDITIONAL (IC); see the following example:

(1) [Já   bý-t   tebou],   še-l   by-ch   k lékař-i.
     I.ɴᴏᴍ be-ɪɴꜰ you.ɪɴꜱ.ꜱɢ go-ʟᴘᴛ.ꜱɢ.ᴍ ᴄᴏɴᴅ-1ꜱɢ to doctor-ᴅᴀᴛ.ꜱɢ
     'If I were you, I would go to the doctor.'   (Czech; Trávníček 1951: 683)

Syntactically, ICs are peculiar in that they are headed by an infinitive, but may nevertheless contain an overt subject in the nominative case (ɴᴏᴍ). This pattern is suprising not only in Slavic languages, but in Indo-European in general.[1] Still, similar structures (though not conditional) exist, for instance in Tamil; see (2):

(2) Raman   [(Vasu)   puuri-jæ porikk-æ] maavŭ vaaŋ-in-aan.
     Raman.ɴᴏᴍ Vasu.ɴᴏᴍ puuri-ᴀᴄᴄ fry-ɪɴꜰ   flour.ᴀᴄᴄ buy-ᴘꜱᴛ-ᴍ.3ꜱɢ
     'Raman bought flour (for Vasu) to fry puuris.'
                                          (Tamil; McFadden & Sundaresan 2018: 467)

A number of authors have addressed the phenomenon of Czech ICs descriptively, among them Svoboda (1959; 1960a,b; 1962); Poldauf (1959); Porák (1959); Kiparsky (1960; 1967); Dunn (1982); Karlík (2007); Meyer (2010); Milotová (2011; 2012). Apart from that, more or less detailed notes on ICs can be found in Křížková (1972); Macháčková (1980), the *Mluvnice spisovné češtiny* (Trávníček 1951), Bauer & Grepl (1981), and Grepl & Karlík (1998). To the best of our knowledge, the present paper is the first attempt of a formal theoretical account of these structures.

The remainder of the paper is structured as follows: §2 characterizes properties of the structure and provides relevant data. In §3, we discuss both the internal and external syntax of ICs, while §4 deals with ɴᴏᴍ-licensing. In §5, we tackle the question of how the conditional interpretation of ICs arises. §6 summarizes the paper.

2 Description

We provide only a brief characterization of Czech IC structures (see Junghanns & Pitsch 2019 for a more detailed description).

1. ICs minimally consist of an infinitive (ɪɴꜰ) which may be of all possible valency classes and both aspects.[2] Typically, but not obligatorily, the ɪɴꜰ occupies the initial syntactic position in the IC; see (3), (4), and (5):

---

[1]Several Indo-European languages feature non-finite – but not infinitival – structures that allow ɴᴏᴍ-subjects. Well-known examples are gerundial clauses in English or Spanish. Russian exhibits non-canonical imperatives that combine with subjects of any person/number (see, a. o., Xrakovskij 2009); it is, however, an open question whether the imperative (in this use) is a non-finite form. The same question pertains to the "inflected infinitive" in European Portuguese (see Raposo 1987), as the relevant paradigm involves person/number markers.

[2]Svoboda (1959) mentions the example in (i) with an IC consisting of an infinitive only:

(3) [Ne-mí-t   sv-ou   hudb-u],   tak se   tu
     NEG-have-INF own-ACC.SG.F music-ACC.SG.F so   REFL here
     z-blázní-m!
     PF-become.insane-1SG
     'If I had not my music, I would become insane here.'

     (Czech; Milotová 2012: 4)

(4) [Dozvědě-t se   to   Němc-i],   tak nás   […]
     learn;PF-INF REFL this.ACC.SG.N German-NOM.PL so   we.ACC
     po-stříle-l-i.
     PF-shoot-LPT-PL.M
     'If the Germans had learnt that, they would have shot us.'

     (Czech; Meyer 2010: 371)

(5) [Já   mí-t   peníz-e],   koupí-m   to   /
     I.NOM have-INF money-ACC.PL buy;PF-1SG this.ACC.SG
     koupi-l   jsem   to.
     buy;PF-LPT.SG.M be.1SG this.ACC.SG
     'If I had money, I would buy this.'/'If I had had money, I would have
     bought this.'   (Czech; Trávníček 1951: 683)

In addition to the INF, ICs may contain sentential negation (*ne*-), objects, and adverbials. On the other hand, ICs *never* contain a (conditional) subjunction or wh-pronouns.[3]

2. ICs can have an overt NOM-subject which may be a full nominal expression or a pronoun; see (6) and (7), respectively:

(i) [Prše-t], zůsta-l-i   bych-om doma.
     rain-INF stay-LPT-PL.M COND-1PL at.home
     'If it rained, we would stay at home.'   (Czech; Svoboda 1959: 167)

[3]The absence of a subjunction is a major difference in comparison with infinitival conditionals in Polish, which obligatorily contain a conditional subjunction; see (i):

(i) [Jeśli się przyjrze-ć bliżej], to   widać na dole małe   literk-i.
     if   REFL look.at-INF closer then visible at bottom little-ACC.PL letter-ACC.PL
     'If one takes a closer look, one can see little letters at the bottom.'

     (Polish; Engel & Rytel-Kuc 1999: 455)

Apart from this difference, Polish infinitival conditional clauses exclude NOM-subjects.

(6) [Bý-t otec doma], by-l-o by se
    be-INF father.NOM at.home be-LPT-SG.N COND.3SG REFL
    to ne-sta-l-o.
    this.NOM.SG.N NEG-happen-LPT-SG.N
    'If father had been at home, this would not have happened.'
    (Czech; Svoboda 1960a: 65)

(7) [U-děla-t to někdo dnes], ne-by-l-o
    PF-make-INF this.ACC somebody.NOM today NEG-be-LPT-SG.N
    by na tom nic výjimečn-ého.
    COND.3SG on this.LOC nothing extraordinary-GEN.SG.N
    'If somebody did this today, there would not be anything
    extraordinary about it.' (Czech; Milotová 2012: 1)

3. The subject of ICs can be coreferential or non-coreferential with the subject
   of the matrix clause; see, e. g., (4), (5) and (6), (7), respectively.

4. ICs either precede or follow their matrix clause. Examples (1) and (3) to (7)
   illustrate the former case, while the latter one is exemplified by (8):

(8) Ži-l-a by-s úplně jinak, [ne-bý-t
    live-LPT-SG.F COND-2SG completely differently NEG-be-INF
    válk-y]?
    war-GEN.SG
    'Would you have led a completely different life, if there had not
    been the war?' (Czech; Milotová 2012: 7)

5. ICs refer to specific situations in the extralinguistic world (Svoboda 1960a:
   77); see the finite conditional *-li*-clause in (9a) with a generic interpretation
   as opposed to the IC in (9b) which refers to a specific situation of coal
   burning:

(9) a. [Shoří-li 1 kg uhl-í na ohništ-i
       burn.down;PF.3SG=if 1 kg.NOM coal-GEN.SG on grate-LOC.SG
       kotl-e parn-ího stroj-e], vydá 7000
       boiler-GEN.SG steam-GEN.SG engine-GEN.SG emit;PF.3SG 7000
       velk-ých kalori-í.
       big-GEN.PL calory-GEN.PL
       'If 1 kg of coal burns down on the grate of a boiler of a steam
       engine, it emits 7000 kilocalories.'

    b. [Shoře-t   1 kg   uhl-í   na ohništ-i
      burn.down;PF-INF 1 kg.NOM coal-GEN.SG on grate-LOC.SG
      kotl-e   parn-ího   stroj-e],   vydá   7000
      boiler-GEN.SG steam-GEN.SG engine-GEN.SG emit;PF.3SG 7000
      velk-ých   kalori-í.
      big-GEN.PL calory-GEN.PL
      'If 1 kg of coal was burning down on the grate of a
      steam engine, it would emit 7000 kilocalories.'

                                                    (Czech; Svoboda 1960a: 77)

6. ICs express conditions and are therefore interpreted like finite conditional clauses marked by one of the subjunctions *jestli(že)*, *kdyby*, *když*, *-li*, *pak-liže*, or *pokud* 'if'. This is illustrated in (10), where the IC-example (8) from above is re-used and contrasted with a *kdyby*-clause with the same meaning:

(10) a. Ži-l-a   by-s   úplně   jinak,   [ne-bý-t
         live-LPT-SG.F COND-2SG completely differently NEG-be-INF
         válk-y]?
         war-GEN.SG
         'Would you have led a completely different life, if there had not been the war?'

      b. Ži-l-a   by-s   úplně   jinak,   [kdy-by
         live-LPT-SG.F COND-2SG completely differently if=COND.3SG
         ne-by-l-a   válk-a]?
         NEG-be-LPT-SG.F war-NOM.SG.F
         'Would you have led a completely different life, if there had not been the war?'

Finally, it should be noted that, unlike the above-mentioned types of conditional clauses with explicit conditional subjunctions, ICs never express real conditions, but are restricted to irreal (hypothetical and counterfactual) conditions.

3 Syntax of ICs

In this section, we discuss and analyze the internal (§3.1) and external (§3.2) syntax of Czech ICs. Assumptions are kept as minimal as possible.

3.1 Internal syntax

A number of facts provide evidence for the size of ICs.

A first thing to note is that given the INF is a transitive predicate, ICs contain objects; see (11). Objects are base-generated in VP, so ICs are at least VPs.

(11) [Ne-mí-t   sv-ou   hudb-u],   tak se   tu
       NEG-have-INF own-ACC.SG.F music-ACC.SG.F so   REFL here
       z-blázní-m!
       PF-become.insane-1SG
       'If I had not my music, I would become insane here.'

<div align="right">(Czech; Milotová 2012: 4)</div>

A second piece of evidence comes from the fact that ICs can be headed by infinitives of causative verbs. An example is *udělat* 'make' in (7) above. As causativity is commonly associated with the presence of the (non-deficient) head *v* (see, e. g., Marantz 1999), ICs must at least be *v*Ps.

Furthermore, ICs can be negated. Standard assumptions link sentential negation to the presence of a functional NegP (see, e. g., Błaszczak 2009: 432). It follows that ICs are at least NegPs/PolPs.

Another crucial property of ICs is that they can contain overt NOM-subjects, hence involve the option of NOM-licensing. Examples above with NOM-subjects are (1), (4), (5), (6), (7), and (9b). Since it is the functional head T which is often considered the locus of NOM-licensing (see, e. g., Pesetsky & Torrego 2001), this speaks in favor of analyzing ICs as TPs. The analysis of ICs as TPs is corroborated by the fact that these structures come with a propositional interpretation. By standard assumptions, TP is the syntactic equivalent of a proposition.

A final fact to observe is that ICs can be paraphrased by full-fledged (finite) conditional clauses. This might be regarded as evidence that, much like conditional clauses, ICs are full CPs. But the paraphrasability with conditional clauses is a weak *syntactic* argument, as it merely restates the fact that ICs are *interpreted* as conditions (see §5). What is more, ICs never contain subjunctions or wh-pronouns. Due to these facts and the lack of substantial evidence for more syntactic structure, we analyze ICs as TPs.[4] Table 1 summarizes the preceding considerations.

---

[4]We do not in principle exclude the possibility of a CP-layer in ICs. Under recent (phasal) assumptions, the C-head might be present but empty, submitting its features (and phase-hood) to T, thus effectively conflating with it. Under these assumptions, ICs would turn out as CPs and TPs at the same time.

Table 1: Size of infinitival conditionals

	VP	*v*P	NegP	TP	CP
availability of objects	✓				
availability of causative verbs		✓			
availability of sentential negation			✓		
NOM-licensing and propositional interpretation				✓	
paraphrasability with conditional clauses					?

3.2 External syntax

There are two questions immediately arising with respect to the external syntax of ICs, namely (i) how and where ICs are connected with the matrix clause, and (ii) how the position of ICs relative to the matrix – preposition vs. postposition – can be explained.

3.2.1 Connection between ICs and the matrix

The general question here is how and where ICs attach to the matrix structure. Our starting point to answer this question is a comparison between ICs and other clause types in Czech.

A particular feature of Czech syntax is that (at least a subset of) adverbial clauses behave differently from relative clauses and subject clauses. More precisely, whereas relative and subject clauses act as the syntactic host for second-position clitics in the matrix clause (see Dokulil 1956; Fried 1994), adverbial clauses do not (see, a. o., Dokulil 1956: 109; Lenertová 2004: 150).

A subject clause example is cited in (12). It shows that the second-position clitic (in italics) is syntactically hosted by the subject clause (in square brackets):[5]

(12) [Že nikdo   ne-protestova-l],   *ho*   ne-překvapi-l-o.
       that nobody.NOM NEG-protest-LPT-SG.M he.ACC NEG-surprise-LPT-SG.N
       'The fact that nobody spoke up against it did not surprise him.'

                                                            (Czech; Fried 1994: 168)

An adverbial clause example is given in (13). The second-position clitic cannot be syntactically hosted by the adverbial clause. Instead, it has to occur in a posi-

---

[5]Note that the direction of prosodic cliticization of Czech second-position clitics may change; see Toman (1996).

tion following a stressed constituent of the matrix clause (here: *odmlčel* '(he) fell silent'):

(13) a. * [Když domluvi-l],   *se*   odmlče-l.
          when finish-speaking-ʟᴘᴛ-sɢ.ᴍ ʀᴇꜰʟ fall.silent-ʟᴘᴛ-sɢ.ᴍ

     b.   [Když domluvi-l],   odmlče-l   *se.*
          when finish-speaking-ʟᴘᴛ-sɢ.ᴍ fall.silent-ʟᴘᴛ-sɢ.ᴍ ʀᴇꜰʟ
          'When he had finished speaking, he fell silent.'
                                          (Czech; Junghanns 2002: 130–131)

How do ICs behave with respect to second-position clitics in the matrix clause? The following data are based on example (6) above and reveal a parallel of ICs and adverbial clauses:

(14) a. * [Bý-t   otec   doma], *by*   se   to
          be-ɪɴꜰ father.ɴᴏᴍ at.home ᴄᴏɴᴅ.3sɢ ʀᴇꜰʟ this.ɴᴏᴍ.sɢ.ɴ
          by-l-o   ne-sta-l-o.
          be-ʟᴘᴛ-sɢ.ɴ ɴᴇɢ-happen-ʟᴘᴛ-sɢ.ɴ

     b.   [Bý-t   otec   doma], by-l-o   *by*   se
          be-ɪɴꜰ father.ɴᴏᴍ at.home be-ʟᴘᴛ-sɢ.ɴ ᴄᴏɴᴅ.3sɢ ʀᴇꜰʟ
          to   ne-sta-l-o.
          this.ɴᴏᴍ.sɢ.ɴ ɴᴇɢ-happen-ʟᴘᴛ-sɢ.ɴ
          'If father had been at home, this would not have happened.'
                                              (Czech; Svoboda 1960a: 65)

The contrast in (14) makes it clear that ICs are not like relative and subject clauses, but rather like adverbial clauses in that they do *not* syntactically host second-position clitics. This leads us to the conclusion that ICs adjoin clause-externally to the matrix CP projection; see the structure in (15):[6]

(15) [_CP **IC** CP]

Given this CP-adjunction structure, the ungrammaticality of sentence (14a) with the second-position clitic immediately following the IC is due to the fact that

---

[6]CP-adjunction is also proposed by Reis (1997: 138) for all "non-integrated" adverbial clauses in German, while Reis & Wöllstein (2010: 167 ff) argue for a CP-adjunction analysis to capture German V1-conditional clauses. More generally, Haegeman (2004: 71) claims all non-integrated adverbial clauses to be clause-external CP-adjuncts. On the other hand, integrated conditional clauses are usually analyzed as clause-internal TP- or *v*P/VP-adjuncts; see, a. o., Reis (1997: 138), Reis & Wöllstein (2010: 144–145, 168), and Bhatt & Pancheva (2006: 647).

the IC is a clause-external adjunct. Assuming that the (split or non-split) CP is the domain of second-position cliticization in Czech (see, a. o., Junghanns 2002; Lenertová 2004), the clause-external position of the IC renders it "invisible" to the second-position clitic in the matrix clause. The conclusion is that ICs cannot function as syntactic host for second-position clitics.[7] However, the structure in (15) does not immediately capture data like (16) where the IC is apparently located *within* the matrix CP.

(16) Když   jsme   se   tu   by-l-i   zjara   projí-t   s   Pauline,
     when   be.1PL REFL   here   be-LPT-PL in-spring walk-INF   with   Pauline
     prohodi-l-a,   [že [ne-bý-t   mě],   vůbec by   sem   na
     remark-LPT-SG.F that NEG-be-INF me.GEN at-all COND.3SG   here on
     ty   smutn-é   skál-y   ne-doš-l-a],   protože
     these.ACC.PL gloomy-ACC.PL rock-ACC.PL   NEG-go-LPT-SG.F   because
     není   romantik   jako já.
     NEG.be.3SG   romantic.NOM.SG.M   like I.NOM
     'When we went for a walk here in spring with Pauline, she remarked that,
     if it had not been for me, she would not at all have climbed these gloomy
     rocks, since she is not a romantic like me.'
                       (Czech; Miloš Urban: *Přišla z moře.* Praha: Argo 2014, 178–179.)

Examples like (16) are likely to involve CP-doubling as depicted in (17) (see Kaspar 2016 on the relevant syntactic structure).[8]

(17)   [CP *že* [CP **IC** CP]]

---

[7]Further clause-external expressions that do not host second-position clitics in Czech are gerundive expressions, vocatives, and external topics (see, a. o., Dokulil 1956, Trávníček 1959, Junghanns 2002: 130–131).

[8]Note that in (16) neither the IC nor the complementizer can function as a syntactic host for the second-position clitic *by* (which is hosted by *vůbec*). The CP-doubling analysis in (17) captures this because both the IC and the complementizer are clause-external relative to *by*.

Iatridou & Kroch (1992) observe that CP-doubling occurs under non-negative, non-irreal bridge verbs only. A reviewer remarks that the former condition does not seem to be correct; see (i):

   (i)   Ne-vědě-l,   že [ne-bý-t   mě],   vůbec by   tu   ne-by-l.
         NEG-know-LPT-SG.M that NEG-be-INF me.GEN at-all COND.3SG here NEG-be-LPT-SG.M
         'He didn't know that, if it was not for me, he would not be here at all.'

3.2.2 Position of ICs relative to the matrix clause

If the CP-adjunction analysis in (15) is on the right track, we are faced with two competing theoretical options concerning the position(ing) of the IC – pre- or postposition – relative to the matrix clause:

(i)   There is only one base position for ICs: initial or final.
This might be characterized as the "asymmetrical view" (see, a. o., Kayne 1994) according to which all phrasal structures are uniformly linearized. It follows that deviating surface orders must be derived.

(ii)   The initial and the final position of ICs are likewise base positions and thus syntactically on a par.
This might be characterized as the "symmetrical view" (see, a. o., Chomsky 1986; 2004) according to which both left and right adjunction are equally permitted. It follows that statistically prevalent and/or marked orders must be explained by specific syntactic restrictions or so-called third factors.

As it does not require any derivations to take place, minimalist principles favor (ii). Assuming this, (18a) and (18b) are equally possible base positions of ICs:

(18)   a. $[_{CP}$ **IC** CP $]$
         b. $[_{CP}$ CP **IC** $]$

Irrespective of which of these options is chosen, it is necessary to explain the empirical fact that ICs in initial position are more frequent than ICs in final position (see Svoboda 1960a: 74–75, Milotová 2011; 2012). A first hint to a possible answer comes from Svoboda (1960a: 75) who suggests that the position of ICs correlates with the theme/rheme (or information) structure of the relevant sentences. We build on his intuition, but will expand it: Indeed, the possible relative positions of ICs have to do with information structure, but information structure alone is not sufficient to explain the facts.

Our proposal is that a "third factor" needs to be considered, namely the logical structure underlying the relevant utterances; see (19). To put it more precisely: It is the relation between the underlying *logical structure* and the actual constituent order as following from *information structure* which determines if an IC is left- or right-adjoined to the matrix CP.

(19)   condition → consequence   (logical structure)

According to Greenberg's universal 14 , the conditional clause precedes the consequence clause (Greenberg 1963: 66). This is the normal order in conditional statements in all languages (see also Diessel 2001: 445–446). Given this, the default or unmarked surface order of statements with ICs is one where the surface ordering of IC and matrix clause iconically reflects the underlying logical structure; see (18a). The reverse ordering in (18b) is (relatively) more marked.

In order to explain the relation between logical structure and information structure, it is necessary to introduce at least basic information-structural concepts: Our theoretical framework in this respect is the "Leipzig model of information structure" (see, a. o., Junghanns & Zybatow 2009). According to this model, there are two coexisting, but not necessarily coextensive structures, namely the TOPIC/COMMENT and the FOCUS/BACKGROUND structure. Another crucial point to note in the current context is that (at least) two types of focus have to be distinguished, namely (a) NATURAL FOCUS and (b) CONTRASTIVE FOCUS. While constituents associated with (a) appear in sentence-final position, constituents associated with (b) show a particular intonational contour and may occupy any syntactic position.

With these few theoretical concepts, the following observations can be made with respect to conditional sentences containing ICs:

- Irrespective of its actual surface position in the sentence, the IC is *always* background material. This status is logically determined: Since conditions are logical premises, they correspond to background information in propositional utterances, viz. conditional clauses.

- Depending on the context, various information-structural analyses are possible for sentences with the "logically unmarked" ordering in (18a), i. e. with the IC in initial position.

- By contrast, the matrix CP is always contrastively focused in sentences showing the "marked" order in (18b), i. e. with the IC in final position; see (20a) and its information-structural analysis in (20b):[9]

(20) a. Part-a   by   se   rozpad-l-a,   [ne-bý-t
          group-NOM.SG.F COND.3SG REFL fall-apart-LPT-SG.F NEG-be-INF

---

[9]It should be noted that contrastively focused constituents in sentence-initial position convey special emphasis as compared with contrastively focused constituents in other positions. This seems to be the reason why the initial positioning of contrastively focused constituents is the preferred strategy in spoken language.

        televiz-e].
        television-GEN.SG
        'The group would (have) fall(en) apart, if there would not (have)
        be(en) television.'   (Czech; Milotová 2012: 4)

  b. Parta by se R/OZpadla, [nebýt televize].
       CONTRASTIVE FOCUS   BACKGROUND

Our conclusions concerning the relative position(ing) of ICs are as follows:

- ICs are right- or left-adjoined to the matrix CP projection, with neither of these possibilities being "more basic" than the other one (symmetrical view on adjunction).

- Right-adjoined ICs – see (18b) – are the (relatively more) marked option for two interrelated reasons:

    (i) Conditional sentences with ICs in final position do not match with the logical structure underlying conditional utterances (anti-iconicity).

   (ii) When the IC appears in sentence-final position, the matrix CP is always contrastively focused (information structure).

4 Nominative case licensing

A syntactic peculiarity of ICs is the availability of overt NOM-subjects. Not only for Slavic languages is this fact surprising, given the common view that there is a 1-to-1 correspondence between NOM-licensing and finiteness, where the latter term usually refers to clausal structures that exhibit both verb-subject agreement and a tense specification (see, e. g., Cowper 2002: 4).

    Czech is a consistent *pro*-drop language, so – again by standard assumptions – the following biconditionals (should) hold:

(21)   a. finite structure ↔ NOM   → *pro* or DP
        b. non-finite structure ↔ *NOM   → PRO

However, the picture drawn is too simple as it cannot explain the licensing of overt NOM-subjects in Czech ICs. To find a possible solution, we discuss three potential theoretical alternatives in §4.1. Building on one of these alternatives, we will present our proposal in §4.2. Finally, we will discuss the conditions for NOM-licensing in Czech ICs in §4.3.

4.1 Theoretical possibilities

We can come up with three theoretical possibilities challenging the standard view on NOM-licensing mentioned in (21) so as to explain the availability of overt NOM-subjects in ICs.

4.1.1 Finiteness in C

According to Kayne (1994), finiteness requires the incorporation of T[ense] into C[omp].[10] In a similar vein, recent minimalist theory (a. o., Chomsky 2007 or Richards 2007) assumes feature inheritance from C to T. Even more plainly, finiteness is analyzed as a part of the C-system in the cartographic framework of Rizzi (1997).

The assumptions just mentioned have in common that finiteness originates not in T, but in C. One theoretical advantage of this view is that the TP is furnished with finiteness "from outside", so that there is no necessity to directly connect finiteness or non-finiteness to respective morphological markers appearing on the verb in v/V. This, in turn, gives rise to the theoretical possibility for clauses headed by a non-finite verb to be finite.

Although the theory itself is quite appealing, it has one crucial disadvantage when applied to ICs: As discussed in §3.1, there is no empirical evidence that ICs comprise a CP-layer. Assuming the category of finiteness to originate in C would force one to stipulate the presence of a CP in ICs exclusively for theory-internal reasons. Since such an approach contradicts our general aim of keeping the body of assumptions as minimal as possible, other theoretical possibilities need to be taken into consideration.

4.1.2 "UPro" in non-finite structures

Sundaresan (2014) and McFadden & Sundaresan (2018) observe that the standard theory of pronominal null subjects in non-finite structures faces severe empirical problems. They convincingly show that there are a number of unrelated languages that exhibit non-finite clauses with case-marked *pro*-subjects (as well as there are languages with finite clauses that have PRO-subjects). To account for these data, they propose that *pro* and PRO are based on one and the same lexical item – an underspecified null pronoun dubbed "UPro" –, the concrete realization of which is determined by its actual syntactic context.

---

[10]Already Raposo (1987) argues for "infl-to-comp" movement.

Although the authors remain silent as regards case-licensing, their proposal allows for the possibility of NOM-licensing in non-finite clauses. Moreover, their theory is sufficiently restrictive in that NOM-licensing in non-finite environments can only take place under specific (syntactic) conditions, which effectively means that PRO remains the default type of subject pronouns in non-finite clauses.

Clearly, the crucial advantage of this approach lies in its flexibility concerning the distribution of null subject pronouns in (not only) non-finite structures. With respect to ICs, however, there is also at least one severe disadvantage, namely the unclear connection between the realization of either PRO, *pro*, or a full DP and the syntactic mechanism of case-licensing (since the cited authors pursue a derivational approach, they do not attribute the appearance of the elements in question to case-licensing at all). Moreover, the theory relates to null subject pronouns, leaving aside the possibility of overt DP subjects in non-finite structures, which is exactly what we find in Czech ICs. This is why, despite its obvious theoretical advantages and rich empirical coverage, the "UPro" theory cannot explain ICs with NOM-subjects, so a third theoretical option must be considered.

4.1.3 Tensed non-finite structures

The final theoretical possibility to be discussed goes back to Stowell (1982; 1995) and Wurmbrand (2001). According to it, there is no 1-to-1 correspondence between NOM-licensing and finiteness as suggested in (21). The crucial proposal is that non-finite clauses are not necessarily untensed, but may be tensed (which is taken to be the prerequisite for the NOM to be licensed). In other words, being non-finite is not the same as being untensed, which gives rise to a much more flexible system as compared to standard assumptions.

To demonstrate this flexibility, we use the two features [AGR] ('agreement') and [TENSED] ('tensed') for a cross-classification; see Table 2.[11]

The prototypical constellations (at least in most Indo-European languages) are the classes (i) and (ii). While class (i) represents prototypical finite structures, class (ii) represents prototypical non-finite structures. Classes (i) and (ii) correspond to (21). We should also like to point out that class (ii) is likely to capture Czech ICs that do not contain an overt NOM-subject: Assuming syntax to be a parsimonious device, one would clearly state that, if there is no necessity for an

---

[11]We prefer the feature [AGR] to [FIN] ('finite') because the grammatical concept of "finiteness" seems to be a conspiracy of two distinct properties, namely NON-/AGREEMENT (showing up on the verb) and UN-/TENSEDNESS (reflected in the non-/availability of a NOM-subject), rather than a single grammatical property.

Table 2: Classes of finiteness

	[AGR]	[TENSED]
(i)	✓	✓
(ii)		
(iii)	✓	
(iv)		✓

overt subject in an IC, NOM-licensing in terms of [TENSED] should not take place either, leaving a PRO-subject as the optimal choice.

However, two further options arise:

Class (iii) represents what might be called non-canonical (or apparent) finite structures. Since it does not immediately relate to our topic, we leave it to future empirical and theoretical research to find out whether or not structures of this kind exist.[12]

Finally, class (iv) represents what might be called non-canonical (or apparent) non-finite structure. This featural combination is of special importance for the present discussion, since it has the capacity to capture Czech ICs with an overt NOM-subject.

The overall advantage of this theoretical alternative is that, while the feature [AGR] relates to the inflectional morphology showing up on the verb form in v/V (which is always an infinitive in case of ICs), the feature [TENSED] relates to T. The two features are thus independent of one another (only for the two "prototypical" classes (i) and (ii) in Table 2 are they in "harmony").

We suspect that it is this splitting up of finiteness into two distinct grammatical properties – namely (i) being marked for agreement (verb form in v/V) and (ii) being tensed (functional head T) – which allows a consistent syntactic analysis of non-canonically behaving non-finite structures, among them Czech ICs with overt NOM-subjects. Essentially, we argue that ICs are underspecified as concerns the feature [UN/TENSED] and can, hence, belong to two different classes shown among those in Table 2: class (ii) or class (iv).

---

[12]Arguably, class (iii) captures Bulgarian non-finite *da*-clauses (see Krapova & Petkov 1999 for the proposal that Bulgarian *da*-clauses come in two varieties, [+T] or [−T]; see also Pitsch 2018). Here, NOM-licensing seems to be impossible despite the presence of a ("finite") verb agreeing in person and number with the (implicit) subject. It seems that this class is restricted to languages that lack an infinitive, hence do not explicitly mark non-finiteness (or, rather, non-agreement).

4.2 Proposal

Building on ideas by Stowell (1982; 1995) and Wurmbrand (2001) (see §4.1.3), we argue that the concept of finiteness is composed of two distinct grammatical properties, representable by the features [AGR] and [TENSED], respectively. Our specific claim with respect to Czech ICs is that, due to their being headed by an infinitive, these structures are always specified as lacking [AGR], but are variable as concerns the specification of T with or without [TENSED].

While this section concerns the technical details of our analysis, we will discuss the conditions under which T is (not) specified with [TENSED] in §4.3.

An adequate means to formalize our ideas is the twofold distinction between INTERPRETABILITY and VALUATION as proposed by Pesetsky & Torrego (2001). Accordingly, syntactic features can be interpretable or uninterpretable: Interpretable features are associated with semantic content and are thus relevant to interpretation, which is why they are transferred to the conceptual-intentional system. By contrast, uninterpretable features are purely formal and have to be deleted before spell-out. On the other hand, features of whatever grammatical category need a certain value. While some features come with a value, others need to be valued in the course of syntactic derivation. Crucially, interpretable features can only be sent to the interface(s) if they have (received) a value.

Above, we used two simple features, [AGR] and [TENSED]. In the framework just described, these features translate as follows into more complex features with a position for a value (in square brackets):

- [AGR] translates into {i,u}Φ[ ]

- [TENSED] translates into {i,u}T[ ]

For our analysis, the functional head T, overt NOM-subjects, PRO, and the infinitive in v/V are the relevant syntactic objects.[13] They have the following features:

1. Since it is the locus of semantic tense, the functional head T is equipped with an interpretable T-feature which can have one of two possible values, namely TENSED or UNTENSED.[14] On the other hand, the Φ-features of T are uninterpretable and unvalued.

---

[13]Our analysis implies that Czech ICs never contain *pro*-subjects.

[14]In finite clauses, the T-value TENSED T is further specified as present, past, or future. By contrast, tensed ICs lack further specification due to the absence of a tense marker on the infinitive. It follows that the relative temporal orientation of ICs can only be inferred from the context.

2. Overt subjects have an uninterpretable T-feature, the morphological reflex of which is the NOM (see Pesetsky & Torrego 2001). Additionally, they enter syntax equipped with inherent interpretable and valued Φ-features.

3. The null pronoun PRO bears an uninterpretable T-feature with the value UNTENSED (in other words, PRO has no case). As to Φ-features, PRO enters syntax without any value – only in the course of syntactic derivation is it supplied with values by its controller (usually a matrix argument).

4. The infinitive in v/V is equipped with an uninterpretable and unvalued T-feature. Its Φ-features are uninterpretable and unvalued, too.[15] This means that the infinitive is completely void of grammatical information. Crucially, it does *not* inherently bear the T-value UNTENSED, but is unspecified in this respect (the value is set in the course of derivation against the respective value of T).[16]

Table 3 gives an overview of the features of the relevant syntactic objects.

Table 3: Syntactic features

	T-feature	Φ-feature
T-head	iT[{TENSED,UNTENSED}]	uΦ[ ]
NOM-subjects	uT[TENSED]	iΦ[Φ]
PRO	uT[UNTENSED]	iΦ[ ]
infinitive (v/V)	uT[ ]	uΦ[ ]

From these assumptions it follows that Czech infinitives in ICs may either receive the value TENSED or UNTENSED. This in turn means that ICs come in two variants.[17] We give detailed analyses of both of them in the sections to follow.

---

[15]T- and Φ-features of verb forms in v/V are generally uninterpretable, their interpretable counterparts being T and the subject, respectively. The basic idea behind this is that verbal inflectional morphology is but a formal (semantically vacuous) reflex of the semantically relevant features of T and the subject.

[16]The unspecification of the Czech/Slovak infinitive with respect to its T-value seems to be the crucial difference as compared to infinitives of other (non-)Slavic languages that lack structures with infinitives and NOM-subjects.

[17]A reviewer suggests that, unless there is good evidence for PRO, it seems easier to postulate only one, *tensed* structure for ICs (*pro*). We agree that the status of PRO is controversial, but the claim of *pro* in ICs without overt subjects strikes us as theoretically disadvantageous, because if such ICs are tensed and contain *pro*, this implies that potentially all non-finite structures are really tensed and contain *pro* – a rather bold claim. Moreover, untensed ICs with PRO comply with the economy principle, as they can be derived and interpreted with minimal effort.

4.2.1 Non-finite ICs

The first possible variant of Czech ICs is untensed, hence the T-head involved enters syntax with the T-feature iT[UNTENSED]. The infinitive in v/V then receives the same value. From the lack of tensedness, it follows that the subject can only be PRO. PRO receives Φ-values (which are subsequently transmitted to T and v/V) from its controller. In (22), the line in (a) shows the situation as it is after base-merge, the line in (b) shows the relevant valuations, and the line in (c) shows the deletion of uninterpretable features before spell-out.

(22) a. $[_{TP} \; T_{iT[\text{UNTENSED}],u\Phi[\;]} \; [_{v/VP} \; \text{PRO}_{uT[\text{UNTENSED}],i\Phi[\;]} \; v/V_{uT[\qquad],u\Phi[\;]} \;]]$

 b. $[_{TP} \; T_{iT[\text{UNTENSED}],u\Phi[\Phi]} \; [_{v/VP} \; \text{PRO}_{uT[\text{UNTENSED}],i\Phi[\Phi]} \; v/V_{uT[\text{UNTENSED}],u\Phi[\Phi]} \;]]$

 c. $[_{TP} \; T_{iT[\text{UNTENSED}],\cancel{u\Phi[\Phi]}} \; [_{v/VP} \; \text{PRO}_{\cancel{uT[\text{UNTENSED}]},i\Phi[\Phi]} \; v/V_{\cancel{uT[\text{UNTENSED}]},\cancel{u\Phi[\Phi]}} \;]]$

4.2.2 Finite ICs

The second possible variant of ICs is tensed, hence the T-head involved enters syntax with the T-feature iT[TENSED].[18] The infinitive in v/V then receives the same value. The presence of a tense specification licenses the NOM on the subject (the subject's Φ-features subsequently value the Φ-features of T and v/V). At the surface, the tensedness of this type of ICs is reflected by the morphological NOM-marking of the subject, see (23):

(23) a. $[_{TP} \; T_{iT[\text{TENSED}],u\Phi[\;]} \; [_{v/VP} \; \text{DP}_{uT[\text{TENSED}],i\Phi[\Phi]} \; v/V_{uT[\qquad],u\Phi[\;]} \;]]$

 b. $[_{TP} \; T_{iT[\text{TENSED}],u\Phi[\Phi]} \; [_{v/VP} \; \text{DP}_{uT[\text{TENSED}],i\Phi[\Phi]} \; v/V_{uT[\text{TENSED}],u\Phi[\Phi]} \;]]$

 c. $[_{TP} \; T_{iT[\text{TENSED}],\cancel{u\Phi[\Phi]}} \; [_{v/VP} \; \text{DP}_{\cancel{uT[\text{TENSED}]},i\Phi[\Phi]} \; v/V_{\cancel{uT[\text{TENSED}]},\cancel{u\Phi[\Phi]}} \;]]$

4.3 Conditions for NOM-licensing in ICs

This section concerns the conditions that require ICs to be tensed, hence under which NOM-licensing in ICs takes place (see §4.2.2).

---

[18]A reviewer asks what it means interpretationally for a T to be tensed, and suggests that tensed T's might be not (as) dependable in their interpretation on the matrix T. Krapova & Petkov (1999) tackle this issue wrt. Bulgarian *da*-clauses which seem to come in two varieties, tensed and untensed, thus resembling ICs. However, *da*-clauses are not restricted to conditionals, but occur in many contexts, which allows linguistic tests to reveal differences between their variants. The conditional interpretation of ICs excludes such tests. As an example, it is impossible to use different temporal adverbials in the IC and the main clause, since conditionals imply a temporal order that cannot be altered. Hence our belief that the un/tensedness of ICs only shows in the un/availability of an overt NOM-subject.

Two different scenarios – (a) and (b) – have to be distinguished. They can be characterized by two criteria:

  (i) Is the subject of the IC overt or null?

  (ii) Are the subjects in the compound conditional sentence in coreference or in disjoint reference?

(a) Null subject, coreference

In this scenario, derivational economy forces the insertion of PRO, as it does not require NOM-licensing and automatically yields coreference with the matrix subject; see (24):

(24) [PRO$_i$ Ne-mí-t   sv-ou   hudb-u],   tak se   tu
           NEG-have-INF own-ACC.SG.F music-ACC.SG.F SO   REFL here
      pro$_i$ z-blázní-m!
           PF-become.insane-1SG
      'If I had not my music, I would become insane here.'

(b) Overt subject, coreference or disjoint reference

In this scenario, T has to bear the feature iT[TENSED] to license the NOM on the subject; see (25):

(25) [Bý-t otec$_j$ doma], by-l-o   by   se
       be-INF father.NOM at.home be-LPT-SG.N COND.3SG REFL
      to$_i$ ne-sta-l-o.
      this.NOM.SG.N NEG-happen-LPT-SG.N
      'If father had been at home, this would not have happened.'

Note that in (25), the subject is background material. However, subjects may also be contrastively focused. If so, they typically occur in clause-initial position as illustrated by (26):

(26) [Otec$_j$ bý-t   doma], by-l-o   by   se
       father.NOM be-INF at.home be-LPT-SG.N COND.3SG REFL
      to$_i$ ne-sta-l-o.
      this.NOM.SG.N NEG-happen-LPT-SG.N
      'If F/ATHER had been at home, this would not have happened.'

A special subcase of this scenario involves overt subject pronouns, since these, in Czech, are always contrastively focused (as opposed to non-emphatic null pronouns); see (27a) illustrating coreference and (27b) illustrating disjoint reference:

(27) a. [Já$_i$ bý-t váš syn], tak se *pro$_i$*
I.NOM be-INF your.NOM.SG.M son.NOM.SG.M so REFL
ne-žení-m.
NEG-marry-1SG
'If /I were your son, I would not marry.'

(Czech; Milotová 2012: 12)

   b. [Já$_j$ bý-t doma], by-l-o by se
I.NOM be-INF at.home be-LPT-SG.N COND.3SG REFL
to$_i$ ne-sta-l-o.
this.NOM.SG.N NEG-happen-LPT-SG.N
'If /I had been at home, this would not have happened.'

The following generalizations can be made:

1. If the subject of the IC is contrastively focused, it has to be overt, so the NOM must be licensed and T has to be equipped with iT[TENSED]; see (27a).[19]

2. If the subject of the IC is not contrastively focused, PRO is sufficient and the NOM does not need to be licensed, hence T can enter syntax with iT[UNTENSED]; see (24).

3. Irrespective of being contrastively focused or not, the subject of the IC has to be overt in case its referent is not identifiable from the matrix clause or the preceding context. If so, the IC must be tensed for the subject to be overtly realized; see (25).

It should be noted that the coreference/disjoint reference-criterion is merely an epiphenomenon. The primary criteria are (i) identifiability and/or (ii) contrastive focussing of the subject of the IC.

---

[19]See Szabolcsi (2009) for similar data from Hungarian. The relevance of contrastive focus is also mentioned by McFadden & Sundaresan (2018: 513).

5 Interpretation

A crucial question with respect to ICs relates to interpretation: How does the conditional interpretation of the compound sentence arise? There are two competing theoretical possibilities:

1. The compositional view:[20]
   Sentences with ICs contain a covert operator that induces the conditional interpretation.

2. The pragmatic view:[21]
   The conditional interpretation arises due to specific lexical and grammatical triggers, the context, and/or world knowledge.

The facts speak in favor of the second – pragmatic – view (see also Milotová 2011; 2012). First, ICs cannot be used in isolation to answer a question about the condition on the realization of a situation; see (28a) and (28b) for the question and answer, respectively:

(28) a. Pod   jak-ou   podmínk-ou   by   se   to
      under which-INS.SG condition-INS.SG COND.3SG REFL this.NOM.SG.N
      by-l-o   ne-sta-l-o?
      be-LPT-SG.N NEG-happen-LPT-SG.N
      'Under what condition would this not have happened?'
   b. *Otec   bý-t   doma.
      father.NOM.SG be-INF at.home
      intended: 'If father had been at home.'

The reason for this exclusion is that ICs lack an "unambiguous relation marking" ("eindeutige Bezugskennzeichnung"; Reis 1997: 135), i. e. a conditional subjunction.

Second, it is generally possible to identify specific signals in the matrix clause that trigger the irreal, viz. conditional interpretation: (i) irrealis/conditional periphrasis; (ii) (realis/indicative) modal verbs (e. g. *moci* 'can'); (iii) (inherently future-oriented) perfective aspect; see (5); (iv) lexical signals, e.g. *tak* 'so' in (27a),

---

[20]See, e.g., the account of conditionals proposed by Kratzer (1979; 1986; 1991). Accordingly, even so-called bare conditionals (without explicit modal operators) contain silent modal (necessity) operators.

[21]See the analyses of Stalnaker (1975); Byrne & Johnson-Laird (2010); Elder & Jaszczolt (2016). Elder & Jaszczolt (2016: 41) give an overview of existing pragmatic analyses of conditionals.

ted' 'now' (with past tense) in (29); (v) context; see (30); (vi) world knowledge; see (31).

(29) [Ne-bý-t   věc-í,   s   nimi-ž   jsem zbytečně
       NEG-be-INF thing-GEN.PL with which.INS.PL-REL be.1SG pointlessly
       z-tráce-l   čas],   stá-l   jsem ted' před
       PF-lose-LPT-SG.M time.ACC.SG stand-LPT-SG.M be.1SG now in.front.of
       vládc-em.
       ruler-INS.SG
       'If there had not been things with which I pointlessly spent my time, I
       would now be standing in front of the ruler.'   (Czech; Milotová 2012: 5)

(30) ...kdy jsem tu   přenocova-l,   to   však
       when be.1SG here stay.overnight-LPT-SG.M this.NOM.SG.N however
       by-l-o   opravdu jen ojediněl-é,   protože to
       be-LPT-SG.N really   only sporadical-NOM.SG.N because this.NOM.SG.N
       ne-stá-l-o   za t-u   rozmrzelost
       NEG-be worth-LPT-SG.N for this-ACC.SG.F moroseness.ACC.SG.F
       manželk-y. [Ne-bý-t   jí],   nocova-l   jsem na
       wife-GEN.SG NEG-be-INF she.GEN stay.the.night-LPT-SG.M be.1SG at
       chat-ě   určitě častěji.
       cottage-LOC.SG surely more.often
       '...when I stayed overnight here, this, however, happened really only
       sporadically, since it was not worth my wife's moroseness. If it had not
       been for her, I surely would have spent the night in the cottage more
       frequently.'   (Czech; Milotová 2012: 5)

(31) [Ne-bý-t   Clintonov-y   administrativ-y],   ne-jsme   v
       NEG-be-INF Clinton's-GEN.SG.F administration-GEN.SG.F NEG-be-1PL in
       NATO.
       NATO
       'If it had not been for the Clinton administration, we would not be in the
       NATO.'   (Czech; Milotová 2012: 5)

6 Summary

The present investigation allows the following conclusions concerning Czech ICs:

Semantically,

- ICs express conditions without explicit marking as conditional clauses;

- ICs yield their interpretation pragmatically through specific grammatical and lexical signals in the matrix clause (irrealis/conditional periphrasis, modal verbs, temporal adverbs, etc.), the context, and/or world knowledge.

Regarding their use,

- ICs are highly marked, expressive, and context-dependent expressions in Czech, which is why they refer to specific situations only.

Syntactically,

- ICs are (at least) TPs;

- ICs can be tensed (NOM-subject) or untensed (PRO), depending on whether or not their subject is identifiable from the context and/or contrastively focused;

- ICs are clause-external adjuncts to the matrix CP.

From a typological point of view,

- ICs have cross-linguistic parallels, see (2) from Tamil;

- Czech and Slovak are the only Slavic languages that exhibit such structures; for Slovak facts see Ružička (1956) and Hirschová (2005).

A possible explanation for the latter fact is that the Czech (and Slovak) infinitive is lexically unvalued with respect to its T- and Φ-features, so (i) it is not restricted to untensed clauses, but may also occur in tensed clauses, and (ii) it can "invisibly agree" with a NOM-subject. The infinitive in other Slavic languages is perhaps more restricted.

Taken all together, these insights reveal that ICs are no "construction", but arise through syntactic structure building and are interpreted pragmatically.

Abbreviations

1,2,3	first, second, third person	LPT	*l*-participle
ACC	accusative case	M	masculine
AUX	auxiliary	N	neuter
COND	conditional	NEG	negation
DAT	dative case	NOM	nominative case
F	feminine	PF	perfective aspect
GEN	genitive case	PL	plural
IC	infinitival conditional	PST	past tense
INF	infinitive	REFL	reflexive
INS	instrumental case	REL	relative
LOC	locative case	SG	singular

Acknowledgements

References

Bauer, Jaroslav & Miroslav Grepl. 1981. *Skladba spisovné češtiny*. Praha: SNP.

Bhatt, Rajesh & Roumyana Pancheva. 2006. Conditionals. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, 638–687. Oxford: Blackwell.

Błaszczak, Joanna. 2009. Negation and clause structure. In Tilman Berger, Karl Gutschmidt, Sebastian Kempgen & Peter Kosta (eds.), *The Slavic languages: An international handbook of their history, their structure and their investigation* (HSK 32.1), 431–468. Berlin: de Gruyter.

Byrne, Ruth M. J. & Philip N. Johnson-Laird. 2010. Conditionals and possibilities. In Mike Oaksford & Nick Chater (eds.), *Cognition and conditionals: Probability and logic in human thinking*, 55–68. Oxford: Oxford University Press.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures. Vol. 3*, 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2007. Approaching UG from below. In Hans-Martin Gärtner & Uli Sauerland (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, 1–30. Berlin: de Gruyter.

Cowper, Elizabeth. 2002. Finiteness. Ms., University of Toronto.

Diessel, Holger. 2001. The ordering distribution of main and adverbial clauses: A typological study. *Language* 77(3). 433–455. DOI:10.1353/lan.2001.0152

Dokulil, Miloš. 1956. Opravovat zvratné *se*, *si* po čárce? *Český jazyk* 6(3). 108–113.

Dunn, John A. 1982. The nominative and infinitive construction in the Slavonic languages. *The Slavonic and East European Review* 60(4). 500–527. https://www.jstor.org/stable/4208579.

Elder, Chi-Hé & Kasia M. Jaszczolt. 2016. Towards a pragmatic category of conditionals. *Journal of Pragmatics* 98. 36–53. DOI:10.1016/j.pragma.2016.04.013

Engel, Ulrich & Danuta Rytel-Kuc. 1999. *Deutsch-polnische kontrastive Grammatik*. Heidelberg: Julius Groos.

Fried, Mirjam. 1994. Second-position clitics in Czech: Syntactic or phonological? *Lingua* 94(2–3). 155–175. DOI:10.1016/0024-3841(94)90024-8

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of language*, 73–113. Cambridge, MA: MIT Press.

Grepl, Miroslav & Petr Karlík. 1998. *Skladba češtiny*. Olomouc: Votobia.

Haegeman, Liliane. 2004. The syntax of adverbial clauses and its consequences for topicalization. In Martine Coene, Gretel de Cuyper & Yves d'Hulst (eds.), *Current studies in comparative Romance linguistics* (Antwerp Papers in Linguistics 107), 61–90. Antwerp: University of Antwerp.

Hirschová, Milada. 2005. Sekundární predikace v češtině a ve slovenštině. In Mária Šimková (ed.), *Aktuálne otázky súčasnej syntaxe*, 103–115. Bratislava: Jazykovedný ústav SAV.

Iatridou, Sabine & Anthony Kroch. 1992. The licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. *Working Papers in Scandinavian Syntax* 50. 1–24. DOI:10.1.1.47.9530

Junghanns, Uwe. 2002. Klitische Elemente im Tschechischen: Eine kritische Bestandsaufnahme. In Thomas Daiber (ed.), *Linguistische Beiträge zur Slavistik IX* (Specimina Philologiae Slavicae 135), 117–150. München: Otto Sagner.

Junghanns, Uwe & Hagen Pitsch. 2019. Konditionale Infinitive im Tschechischen. In Genia Böhnisch, Uwe Junghanns & Hagen Pitsch (eds.), *Linguistische Beiträge zur Slavistik. XXV. JungslavistInnen-Treffen in Göttingen, 13.–16. September 2016* (Specimina Philologiae Slavicae 199), 111–132. Frankfurt am Main: Peter Lang.

Junghanns, Uwe & Gerhild Zybatow. 2009. Grammatik und Informationsstruktur. In Tilman Berger, Karl Gutschmidt, Sebastian Kempgen & Peter Kosta (eds.), *The Slavic languages. An international handbook of their history, their structure and their investigation* (HSK 32.1), 684–707. Berlin: de Gruyter.

Karlík, Petr. 2007. Zur Integration des Infinitivs in das Verbalsystem des Tschechischen. *Anzeiger für Slavische Philologie* 35. 87–100.

Kaspar, Jiri. 2016. Czech left periphery: A preliminary analysis. *Linguistica Brunensia* 64(1). 71–88. http://hdl.handle.net/11222.digilib/135451.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kiparsky, Valentin. 1960. Über das Nominativsubjekt der Infinitive. *Zeitschrift für Slavische Philologie* 28. 333–342.

Kiparsky, Valentin. 1967. Nochmals über das Nominativsubjekt der Infinitive. *Zeitschrift für Slavische Philologie* 33. 263–266.

Krapova, Iliana & Vassil Petkov. 1999. Subjunctive complements, null subjects and case checking in Bulgarian. In Katarzyna Dziwirek, Herbert S. Coats & Cynthia M. Vakareliyska (eds.), *Formal Approaches to Slavic Linguistics 7: The Seattle Meeting 1998*, 267–285. Ann Arbor, MI: Michigan Slavic Publications.

Kratzer, Angelika. 1979. Conditional necessity and possibility. In Rainer Bäuerle, Urs Egli & Arnim von Stechow (eds.), *Semantics from different points of view*, 387–394. Berlin: Springer. DOI:10.1007/978-3-642-67458-7_9

Kratzer, Angelika. 1986. Conditionals. *Chicago Linguistics Society* 22/2. 1–15.

Kratzer, Angelika. 1991. Conditionals. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: de Gruyter.

Křížková, Helena. 1972. Kontextové členění a typy tázacích vět v současných slovanských jazycích. *Slavia* 41. 241–262.

Lenertová, Denisa. 2004. Czech pronominal clitics. *Journal of Slavic Linguistics* 12(1–2). 135–171. https://www.jstor.org/stable/24599870.

Macháčková, Eva. 1980. Ústup kondicionálu minulého? *Naše řeč* 63(2). 62–67. http://nase-rec.ujc.cas.cz/archiv.php?art=6175.

Marantz, Alec. 1999. Creating words above and below little *v*. Ms., MIT.

McFadden, Thomas & Sandhya Sundaresan. 2018. Reducing *pro* and PRO to a single source. *The Linguistic Review* 35(3). 463–518. DOI:10.1515/tlr-2018-0003

Meyer, Roland. 2010. Mood in Czech and Slovak. In Björn Rothstein & Rolf Thieroff (eds.), *Mood in the languages of Europe*, 358–375. Amsterdam: John Benjamins. DOI:10.1075/slcs.120.20mey

Milotová, Kateřina. 2011. Pozice subjektu infinitivu: Infinitiv v podmínkových větách. In Vladimír Petkevič & Alexandr Rosen (eds.), *Gramatika a značkování korpusů. Korpusová lingvistika Praha 2011. Vol. 3*, 70–79. Praha: Nakladatelství Lidové noviny.

Milotová, Kateřina. 2012. *K infinitivním podmínkovým větám v češtině*. Presented at the fourth international conference *Gramatika a korpus*, 28–30 November 2012, AV ČR, Praha. http://www.ujc.cas.cz/miranda2/export/sitesavcr/ujc/veda-vyzkum/vyzkum/gramatika-a-korpus/proceedings-2012/konferencni-prispevky/MilotovaKaterina.pdf.

Pesetsky, David & Esther Torrego. 2001. T-to-C movement: Causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 355–426. Cambridge, MA: MIT Press. DOI:10.1.1.232.9565

Pitsch, Hagen. 2018. Bulgarian moods. *Journal of Slavic Linguistics* 26(1). 55–100. DOI:10.1353/jsl.2018.0003

Poldauf, Ivan. 1959. Děj v infinitivu. *Slovo a slovesnost* 20(3). 183–202. http://sas.ujc.cas.cz/archiv.php?art=998.

Porák, Jaroslav. 1959. O vývoji infinitivních vět v češtině. *Slovo a slovesnost* 20(4). 241–247. http://sas.ujc.cas.cz/archiv.php?lang=en&art=1005.

Raposo, Eduardo. 1987. Case theory and infl-to-comp: The inflected infinitive in European Portuguese. *Linguistic Inquiry* 18(1). 85–109. https://www.jstor.org/stable/4178525.

Reis, Marga. 1997. Zum syntaktischen Status unselbständiger Verbzweit-Sätze. In Christa Dürscheid, Karl-Heinz Rahmers & Monika Schwarz (eds.), *Sprache im Fokus. Festschrift für Heinz Vater zum 65. Geburtstag*, 121–144. Tübingen: Niemeyer. http://hdl.handle.net/10900/47033.

Reis, Marga & Angelika Wöllstein. 2010. Zur Grammatik konditionaler V1-Gefüge im Deutschen. *Zeitschrift für Sprachwissenschaft* 29(1). 111–179. DOI:10.1515/ZFSW.2010.004

Richards, Marc D. 2007. On feature inheritance: An argument from the Phase Impenetrability Condition. *Linguistic Inquiry* 38(3). 563–572. https://www.jstor.org/stable/40071403.

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

Ružička, Jozef. 1956. *Skladba neurčitku v slovenskom spisovnom jazyku.* Bratislava: Vydavateľstvo SAV.

Stalnaker, Robert. 1975. Indicative conditionals. *Philosophia* 5(3). 269–286. DOI:10.1007/BF02379021

Stowell, Tim A. 1982. The tense of infinitives. *Linguistic Inquiry* 13(3). 561–570. https://www.jstor.org/stable/4178293.

Stowell, Tim A. 1995. The phrase structure of tense. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 277–291. Dordrecht: Kluwer. DOI:10.1007/978-94-015-8617-7_10

Sundaresan, Sandhya. 2014. Making sense of silence: Finiteness and the (OC) PRO vs. *pro* distinction. *Natural Language & Linguistic Theory* 32(1). 59–85. DOI:10.1007/s11049-013-9216-6

Svoboda, Karel F. 1959. Mluvnická povaha infinitivu v současné spisovné češtině. *Slovo a slovesnost* 20(3). 161–183. http://sas.ujc.cas.cz/archiv.php?art=997.

Svoboda, Karel F. 1960a. Infinitivní věty podmínkové. *Naše řeč* 43(3–4). 65–78. http://nase-rec.ujc.cas.cz/archiv.php?lang=en&art=4738.

Svoboda, Karel F. 1960b. Ještě k mluvnické povaze infinitivních vět podmínkových. *Naše řeč* 43(9–10). 278–280. http://nase-rec.ujc.cas.cz/archiv.php?art=4792.

Svoboda, Karel F. 1962. *Infinitiv v současné češtině.* Praha: Academia.

Szabolcsi, Anna. 2009. Overt nominative subjects in infinitival complements in Hungarian. In Marcel den Dikken & Robert M. Vago (eds.), *Approaches to Hungarian 11: Papers from the 2007 NYU Conference*, 251–276. Amsterdam: John Benjamins. DOI:10.1075/atoh.11.11sza

Toman, Jindřich. 1996. A note on clitics and prosody. In Aaron L. Halpern & Arnold M. Zwicky (eds.), *Approaching second: Second position clitics and related phenomena*, 505–510. Stanford, CA: CSLI Publications.

Trávníček, František. 1951. *Mluvnice spisovné češtiny, část II: Skladba.* Praha: Slovanské nakladatelství.

Trávníček, František. 1959. K postavení stálých příklonek po přestávce uvnitř věty. *Naše řeč* 42(3–4). 65–76. http://nase-rec.ujc.cas.cz/archiv.php?art=4673.

Wurmbrand, Susanne. 2001. *Infinitives: Restructuring and clause structure.* Berlin: Mouton de Gruyter.

Xrakovskij, Viktor S. 2009. The conditional. In Tilman Berger, Karl Gutschmidt, Sebastian Kempgen & Peter Kosta (eds.), *The Slavic languages. An international handbook of their history, their structure and their investigation* (HSK 32.1), 275–281. Berlin: de Gruyter.

Chapter 7

A syntactic re-analysis of the Slovenian impersonal *se*-construction

Jakob Lenardič

University of Ljubljana

The following paper discusses the syntactic derivation of the Slovenian impersonal *se*-construction (e.g. *Gradilo se je hišo.ACC* 'People were building a house'). On the one hand, the paper argues for an analysis of the construction in which all the thematic arguments selected by the predicate normally enter the syntactic derivation (e.g. Rivero & Milojević Sheppard 2003; contra Marelj 2004; Grahek 2008). To this end, the paper discusses the construction in relation to the causative alternation (e.g. Alexiadou et al. 2015), which shows that agentive participation is invariably entailed in contrast to the personal variant (*Gradila se je hiša.NOM*). On the other hand, the paper shows that Rivero & Milojević Sheppard's (2003) syntactic derivation is set up in such a way that it cannot offer an explanation as to why the construction only allows unergative verbs. In its stead, the paper offers a new analysis of the construction within the theory of the Voice domain (e.g. Legate 2014; Schäfer 2017). The new analysis posits that *se* is the head of an argument-introducing thematic VoiceP and is a separate element from the phonologically-null thematic argument that the construction introduces. Ungrammaticality arises in the case of unaccusative verbs because such predicates merge the null thematic argument within VP yet still combine with *se*, which introduces a participant variable into the derivation that cannot be saturated.

Keywords: argument structure, impersonal construction, reflexive clitics, Slovenian, unaccusative verbs

1 Introduction

The paper discusses the IMPERSONAL *SE*-CONSTRUCTION in Slovenian, as exemplified by (1), in terms of the syntactic representation of its thematic arguments. The

main morphosyntactic features of this construction, which will be discussed in the paper, include (i) the clitic *se*, (ii) the accusative case on the internal argument assigned by a transitive predicate in absence of an overt external argument bearing nominative case, and (iii) default (singular neuter) agreement which follows from the superficial omission of the external argument.

(1) Celo leto se je   gradilo   hišo.
    whole year SE AUX.3SG build.PTCP.SG.N house.ACC
    'People were building the house for a whole year.'

In the formal literature, there are generally two opposing approaches to the number of arguments that are present in the syntactic derivation of (1). The 1-argument analysis (Marelj 2004, Grahek 2008) proposes that the derivation mimics the surface structure, so that the external argument is excluded from the syntactic component and is only implicitly present in the semantics. The 2-argument analysis (Rivero & Milojević Sheppard 2003) regards the structure as containing – in addition to the overt internal argument *hišo* 'house' – a phonologically null pronoun that fills the external argument slot in the syntactic representation and semantically corresponds to the implicit agentive participants paraphrased as 'people' in the English translation.

   In this paper, I first present the empirical problems of the syntactic set-ups of such existing accounts. On the one hand, I defend the 2-argument analysis by discussing transitive impersonal *se*-sentences headed by causative verbs, which show that the construction entails the participation of a volitional initiator (i.e., an agent). This fact is surprising on the 1-argument analysis given that causative verbs are in contemporary morphosyntactic frameworks analysed as predicates that are not lexically specified for an external causer or agent and consequently give rise to a non-agentive anticausative interpretation when they do not combine with an external thematic argument during derivation (e.g. Alexiadou et al. 2015). On the other hand, I show that the syntactic derivation that underlies Rivero & Milojević Sheppard's (2003) analysis is set up in such a way that it cannot offer a principled explanation as to why the impersonal construction does not allow unaccusative verbs in the case of intransitive predicates.

   To account for such empirical gaps, I present a new analysis of the impersonal *se*-construction that is cast within the theory of the Voice domain (Schäfer 2017). The core idea is that the clitic *se* is re-analysed as the head of a thematic VoiceP which introduces an initiator variable into the semantic derivation, and that the null argument is a pro$_{IMP}$ without ϕ-features (Fenger 2018, Ackema & Neeleman

2013; 2018) and with unmarked nominative case (Kornfilt & Preminger 2015). In this sense, the analysis follows the 2-argument account in that all thematic arguments of a transitive predicate normally enter the syntactic derivation. In the case of unaccusative verbs, ungrammaticality arises because such predicates combine with pro$_\text{IMP}$ within VP to saturate the theme variable, yet still merge with *se*, which introduces an open initiator variable that cannot be saturated.

The paper is structured as follows. §2 discusses impersonal *se*-sentences with transitive predicates in relation to the causative alternation, which shows that a 2-argument analysis fares better empirically than the 1-argument analysis. §3 argues that the inadmissibility of unaccusative predicates is problematic for the core syntactic assumptions of Rivero & Milojević Sheppard's (2003) account. §4 lays out the new analysis. §5 is the conclusion.

2 The number of arguments in transitive impersonal *se*-sentences – evidence from causative predicates

2.1 The clitic *se* as a case absorber

In Slovenian, the impersonal *se*-construction (2) alternates with an agreeing personal variant, which generally has the same interpretation if it is headed by a transitive predicate such as *spoštovati* 'respect' or *graditi* 'build' (3).

(2) Impersonal *se*-sentences

    a. Starše   se spoštuje.
       parents.ACC SE respect.3SG
       'Parents are respected.' / 'People respect their parents.'

    b. Hišo   se gradi.
       house.ACC SE build.3SG
       'The house is being built.' / 'People were building a house.'

(3) Personal *se*-sentences

    a. Starši   se spoštujejo.
       parents.NOM SE respect.3PL
       'Parents are respected.' / 'People respect their parents.'

    b. Hiša   se gradi.
       house.NOM SE build.3SG
       'The house is being built.' / 'People were building a house.'

Marelj (2004: 268) has proposed that the interpretative equivalence points to a 1-argument analysis of both constructions.[1] Concretely, she claims that *se* indirectly reduces the argument structure of a predicate because it is first and foremost a case absorber (see also Givón 1998; Franks 1995). In this respect, *se* can either absorb nominative case, resulting in the impersonal sentences in (2), or accusative case, resulting in the corresponding personal variants in (3). Crucially, since *se* is a case absorber, it ensures that a DP corresponding to an external thematic argument is excluded from a syntactic argument position in both personal and impersonal variants, so the external argument of the predicates *respect* and *build* is only present in the semantic component and is invariably interpreted as an indefinite group of people in all *se*-sentences.

For Slovenian, this approach to the argument structure of *se*-sentences is also the most recent one in the formal literature, proposed by Grahek (2008). Although Grahek does not offer an explicit formal account, she makes a similar claim that "both personal and impersonal [*se*-sentences] contain the same type of *se* which reduces the human subject role during their derivation" and that "the choice between nominative and accusative/genitive is not associated with any interpretative difference" (Grahek 2008: 44–48).

However, the interpretative equivalence does not hold for all predicates. With causative predicates like *odpreti* 'open' and *potopiti* 'sink', it is only the impersonal variants in (4) that entail that the reported events involved a volitional initiator; i.e., an agent.

(4) a. Vrata   se je   odprlo   (*same od sebe).
         door.ACC.PL SE AUX.3SG open.PTCP.SG.N   all   by itself
         'The door was opened by someone.'
         Impossible: 'The door opened (by itself).'

     b. Ladjo   se je   potopilo   (*samo od sebe).
         ship.ACC.SG.F SE AUX.3SG sink.PTCP.SG.N   all   by itself
         'The ship was sunk by someone.'
         Impossible: 'The ship sank (by itself).'

(5) a. Vrata   so   se odprla   (sama od sebe).
         door.NOM.PL AUX.3PL SE open.PTCP.PL   all   by itself
         'The door opened (by itself.)'

---

[1]To be clear, Marelj (2004) does not discuss Slovenian data specifically, but rather focuses on other Slavic languages, such as Polish and Serbo-Croatian, in which a *se*-sentence can also surface either as a non-agreeing impersonal or as an agreeing personal construction with the same structural and interpretative characteristics as the Slovenian sentences under (2) and (3). For cross-linguistic examples, the reader is referred to Marelj (2004: 267–270).

b. Ladja   se je   potopila   (sama od sebe).
ship.NOM.SG.F SE AUX.3SG sink.PTCP.SG.F all   by itself
'The ship sank (by itself).'

In the remainder of this section, I will argue that this interpretative non-equivalence speaks in favour of a 2-argument analysis of the transitive impersonal *se*-construction.

2.2 The (anti)causative alternation and the impersonal *se*-construction

We begin with a brief review of the recent literature on the cross-linguistic properties of the causative alternation (e.g. Levin & Rappaport Hovav 1995; Alexiadou 2010; Alexiadou et al. 2015; Schäfer 2017, amongst others). These authors take causative verbs like *open* to be unique among predicates because they are the only ones that are not specified for a certain type of initiator at the level of the lexicon. As an apparent consequence in English, they can either appear in the transitive causative construction, as exemplified by sentence (6a), or in the intransitive anticausative construction, as exemplified by sentence (6b).

(6)   a. Someone opened the door.
      b. The door opened.

Sentence (6b) is structurally poorer than (6a) – crucially, it lacks in its syntactic representation a functional projection like *v*P or VoiceP, which in the transitive variant (6a) introduces the external argument DP (Kratzer 1996). This alternation in argument structure is possible because the causative predicate *open* is not lexically specified for an agent/causer (Alexiadou et al. 2015). In this sense, causative predicates contrast with predicates like *read*, which are lexically specified for an external participant and which consequently have to combine with a functional head like little *v* or Voice so as to introduce the external argument. For this reason, a sentence like (7) – contrary to (6b) – is ungrammatical.

(7)   * The book read.

In contrast with English (6b), Slovenian intransitive sentences interpreted as anticausatives are additionally marked with the clitic *se*, as seen in the examples under (5). In spite of the clitic's presence in overt syntax, such *se*-sentences do not differ in interpretation from the unmarked variants in English, as shown by the corresponding translations. Since the clitic does not make an interpretative

contribution to the construction, Schäfer (2017), who discusses *se*-anticausatives in Romance languages, analyses it as a semantically-vacuous specifier of a VoiceP (the equivalent of Chomsky's 2001 *v*P) whose head is also semantically vacuous and takes the VP as its complement.

(8) [$_{VoiceP}$ SE [$_{Voice'}$ Voice [$_{VP}$ open door.NOM]]]

In the semantic representation, the denotation of the VP in (8) is a function of the set of all change-of-state events in which the door becomes open (9). By contrast, both Voice and *se* denote identity functions that map a semantic argument corresponding to a set of events onto itself (10), so neither makes any semantic contribution (Schäfer 2017).

(9) $[\![\text{VP}]\!] = \lambda e[\text{OPEN}(e) \wedge \text{THEME}(e, \text{DOOR})]$

(10) $[\![\text{SE}]\!] = [\![\text{Voice}]\!] = \lambda P_{\langle s,t \rangle}[P]$

Consequently, when VP combines with the semantically vacuous Voice head and *se*, the resultant interpretation only entails the existence of a change-of-state event in which the door becomes open. That this is indeed how the Slovenian *se*-anticausatives under (5) are interpreted is further emphasized by the admissibility of the *sama od sebe* phrase, which is the Slovenian equivalent of the English *by itself* phrase. This phrase is used to highlight the fact that such sentences denote situations that can be conceptualized as though occurring "without outside help" (Levin & Rappaport Hovav 1995: 88); that is, without the presence of an agent or causer.[2]

By contrast, the corresponding impersonal variants in (4) are not interpreted as anticausatives, but rather as events externally caused by agents. Crucially, however, such an interpretation does not follow straightforwardly from the semantics of a causative predicate like *odpreti* 'open', considering the fact that such a predicate is taken to be cause unspecified in its lexical semantics and the VP that it projects only denotes the characteristic function of a set of events in which the door becomes open (9). On top of that, Marelj's (2004) 1-argument analysis for impersonal sentences also predicts that a thematic external argument has to be excluded because *se* is analysed as a case absorber. Note that, because of such

---

[2]Note that, if the *by itself* phrase is not merged, then the event denoted by the anticausative construction is still conceptually compatible with an agent opening the door, as shown by the fact that the transitive sentence in (6a) asymmetrically entails its anticausative variant in (6b) (Schäfer & Vivanco 2016). The point is that the possible agent in an anticausative construction is missing from its syntactic-semantic representation due to the fact that Voice and its specifier make no semantic contribution (10).

case absorption on part of *se*, the syntactic derivation that Marelj (2004: 268) proposes only allows a null expletive in the grammatical subject position of an impersonal *se*-sentence:

(11)   [$_{CP}$ door.ACC [$_{TP}$ EXPL [$_{T'}$ SE [$_{VP}$ open ~~door~~]]]]

However, if the grammatical subject is semantically null in impersonal *se*-sentences (EXPL in (11)), then it stands to reason that combining the denotation of the VP in (11), which should be equivalent to (9), with those of its c-commanding categories (which make no semantic contribution to expanding the argument structure of the VP) would lead to an interpretation like 'There was an event in which the door became open'. This, however, would be the same interpretation as that of the anticausatives in (5), contrary to the actual meaning.

To get the desired meaning, it rather seems that the VP of the impersonal construction must be predicated of a thematic external argument, just like in the case of the ordinary transitive sentence (6a). Here, the VP *opened the door* combines via event identification (Kratzer 1996) with the external argument DP *someone* located in the specifier of *v*P/VoiceP so as to bring about an external volitional initiator into the semantics of the construction.

This is exactly what the 2-argument analysis of the impersonal *se*-construction predicts. In this approach, Rivero & Milojević Sheppard (2003: 120) propose that transitive impersonal *se*-sentences have the following syntactic representation:

(12)   [$_{CIP}$ [$_{Cl'}$ SE [$_{TP}$ [$_{vP}$ NP$_1$ [$_{VP}$ V DP$_2$]]]]]

The representation in (12) involves two syntactically-projected thematic arguments. While DP$_2$ corresponds to the overt accusative DP *ladjo* in sentence (4b), NP$_1$ is taken to be the external argument, which is analysed as a phonologically-null pronoun that semantically corresponds to an existential quantifier over a group of humans (Rivero & Milojević Sheppard 2003: 135).

(13) $\exists x_{[\text{hum}]} [\text{OPEN}(x_{[\text{hum}]}, \text{DOOR})]$

The LF in (13) perfectly corresponds to the intuitive meaning of the impersonal *se*-sentences under (4); that is, 'There is a human (or a group of humans) *x* such that *x* opened the door', which is the desired result. What is crucial is that the syntactic derivation of an impersonal *se*-sentence on such an account fundamentally differs from that of the personal variant, which cannot introduce the external argument into the syntax due to the fact that its VoiceP (8) is already filled with semantically vacuous material (10). Consequently, the interpretative difference

between the causative impersonals, which entail an initiator, and the corresponding anticausative personals, which do not, becomes obvious.

In sum, what we have shown in this section is that examples of the transitive impersonal construction with causative predicates support a 2-argument analysis (Rivero & Milojević Sheppard 2003) over a 1-argument analysis (e.g. Marelj 2004; Grahek 2008), as the latter does not account for the discussed interpretative differences between the impersonal and personal variants; that is, it is unclear from the latter account why only the impersonal construction should entail agentive participation.

3 Intransitive predicates and the impersonal *se*-construction

3.1 The data

It has been cross-linguistically observed (e.g. Cinque 1995 and D'Alessandro 2008 for Italian; Ilc & Marvin 2016 for Slovenian) that the impersonal *se*-construction is licit with unergative predicates, but unaccusative predicates are out. Consider the following contrast between the grammatical sentences with unergatives (14) and ungrammatical sentences with unaccusatives (15) in Slovenian, adapted from Ilc & Marvin (2016: 152–154):

(14) a. Včeraj   se je   plesalo.
          yesterday SE AUX.3SG dance.PTCP.SG.N
          'Some people danced yesterday.'

      b. Včeraj   se je   cel   dan spalo.
          yesterday SE AUX.3SG whole day sleep.PTCP.SG.N
          'Some people were sleeping yesterday.'

(15) a. *Včeraj   se je   umrlo   v tej bolnici.
          yesterday SE AUX.3SG die.PTCP.SG.N in this hospital
          Intended: 'Some people died yesterday in this hospital.'

      b. *Včeraj   se je   padlo   na teh   stopnicah.
          yesterday SE AUX.3SG fall.PTCP.SG.N on these stairs.
          Intended: 'Some people fell on these stairs yesterday.'

In this section, I will show that Rivero & Milojević Sheppard's (2003) analysis for the Slovenian impersonal *se*-construction does not predict this contrast in acceptability.

3.2 The syntax of impersonal *se*-sentences in Rivero & Milojević Sheppard (2003) and the Unaccusative Hypothesis

On Rivero & Milojević Sheppard's (2003) account, there are two key structural assumptions that underlie the impersonal *se*-construction. Both are tied to the feature configuration of the null argument NP in the bracketed representation (16) of sentence (14a), which is unergative and thus acceptable.

(16) $[_{ClP}$ $[_{Cl'}$ SE$_\exists$ $[_{TP}$ T$_{[-\phi]}$ $[_{VoiceP}$ **NP$_{[uCASE][iHUM]}$** $[_{Voice'}$ Voice $[_{VP}$ dance]]]]]]

On the one hand, Rivero & Milojević Sheppard (2003) assume that the external argument NP has an uninterpretable case feature that needs to be checked prior to spell-out. Since T, which is otherwise the canonical case-checker in ordinary finite sentences, lacks ϕ-features because the construction is non-agreeing, Rivero & Milojević Sheppard (2003) furthermore assume that NP bypasses TP and checks its case feature against the clitic by remerging in its specifier, as in (17):

(17) $[_{ClP}$ **NP$_{[uCASE][iHUM]}$** $[_{Cl'}$ SE$_\exists$ $[_{TP}$ T$_{[-\phi]}$ $[_{VoiceP}$ ~~NP~~ $[_{Voice'}$ Voice $[_{VP}$ dance]]]]]]

On the other hand, Rivero & Milojević Sheppard (2003) propose that NP is a pronoun with a human feature that is otherwise referentially deficient (i.e., NP is a "simplex-expression anaphor" in the sense of Reinhart & Reuland 1993). Because of this assumption, they claim that NP moves to SpecClP, where it repairs its deficiency by forming a chain with the clitic, which is taken to correspond semantically to an existential operator (Rivero & Milojević Sheppard 2003: 129), labelled here with the subscripted ∃. Thus in the LF in (13), the quantifier represents *se*, while the bound variable corresponds to the copy/trace in SpecVoiceP.

On such a syntactic configuration, the derivation of the unaccusative sentences in (15) should be like (17), with the only difference being the locus of the initial merge of the phonologically null NP. In this respect, the standard assumption of the Unaccusative Hypothesis (beginning with Perlmutter 1978) is that an unaccusative predicate differs from an unergative one only in that it merges its sole argument within the VP (for instance, Burzio 1986, Alexiadou et al. 2004), and thus dispenses with a projection external to VP (like *v*P in Chomsky's 2001 system or VoiceP in Kratzer's 1996 system) that would introduce the argument in the case of an unergative predicate. Translating this idea into the configuration of Rivero & Milojević Sheppard (2003), the impersonal construction headed by an unaccusative predicate should be as follows in syntax before movement:

(18) $[_{ClP}$ $[_{Cl'}$ SE$_\exists$ $[_{TP}$ T$_{[-\phi]}$ $[_{VP}$ die **NP$_{[uCASE][iHUM]}$**]]]]

However, since Rivero & Milojević Sheppard's analysis assumes the only relevant syntactic operation to be the remerger of the null NP in the specifier of ClP driven by feature checking/repairing referential deficiency, there is no obvious reason as to why (18) should not result in a grammatical sentence. Note that using this approach the derivation of the argument structure of the impersonal construction with an unaccusative predicate should run in parallel to that of the perfectly licit English structure in (19), where the DP *some people* also starts off VP internally and is then remerged in a higher position.

(19)   [$_{TP}$ some people [$_{T'}$ T [$_{VP}$ died ~~some people~~]]]

Crucially, the null NP in (18) should also trivially remerge in SpecClP to check its case feature and to form a chain with the clitic, as in the unergative derivation in (17). It seems, then, that such a configuration predicts that the unaccusative sentences in (15) should actually be just as grammatical as the unergative ones in (14).

4 The proposal – *se* is the head of VoiceP

4.1 Schäfer's (2017) Voice typology

To account for the inadmissibility of unaccusative verbs, I will now propose a reanalysis of the Slovenian impersonal *se*-construction that is rooted in the typology of Voice heads presented in Schäfer (2017). Following Kratzer (1996), VoiceP is the functional projection atop VP which introduces the external argument DP into the syntactic derivation. The key idea behind Schäfer's proposal, which is an update of Kratzer's, is that the syntactic and semantic requirements of a Voice head vary from one type of construction to another, as follows: (i) a Voice head can introduce the external argument variable into the semantic representation in several different ways and (ii) a Voice head may or may not have a syntactic need for a DP in its specifier slot. For Slovenian, this approach seems especially fruitful because it provides a straightforward explanation of the licensing properties of the so-called *se*-passives (personal *se*-structures in our terms), which differ from canonical passives in that they do not allow their entailed volitional initiator (i.e., agent) to be expressed via the Slovenian equivalent of the *by*-phrase. To see this approach in action, consider the differences between the feature sets of the Voice heads for the following three sentences, where the first feature in the brackets

corresponds to the specifier requirement of Voice and the second to its semantics (Schäfer 2017: 14).[3]

(20) a. John read a book.

    b. Voice = {D, $\lambda x \lambda e[\text{INITIATOR}(x, e)]$}

(21) a. The book was read (by John).

    b. Voice = {/, $\lambda x \lambda e[\text{INITIATOR}(x, e)]$}

(22) a. Knjiga   se je   brala   (*s   strani Janeza).
       book.NOM.FEM SE AUX.3SG read.PTCP.SG.F on part Janez
       'The book was being read.'

    b. Voice = {D, $\lambda e \exists x[\text{INITIATOR}(x, e)]$}

On this approach, a canonical passive sentence (21) differs from its active variant (20) only in that its Voice head lacks the D-feature, which accounts for the fact that the external argument DP cannot directly enter the derivation. In semantics, however, this self-same Voice head introduces an initiator variable. Since this variable is open, it can be saturated by a DP introduced via the *by*-phrase, which enters the derivation as a VoiceP adjunct. By contrast, the Voice head in the personal *se*-structure in (22a), which does entail agentive participation much like (21a), introduces an initiator variable that is inherently closed off by the existential quantifier. Consequently, merging a *s-strani* 'on-part-of' phrase is correctly predicted to be illicit since the external argument DP that the phrase introduces does not have an open variable to saturate. Finally, since (22b) has the D-feature, the *se* clitic is assumed to be a semantically-null DP that is merged in SpecVoiceP to satisfy this c-selectional requirement of Voice (Schäfer 2017).

As to how the impersonal *se*-construction is positioned within this kind of Voice theory, only a brief speculative account is provided in Schäfer (2017: Footnote 14). Following D'Alessandro (2008), he claims that it is *se* itself that is a thematic syntactic argument that expresses the initiator role in a transitive sentence, in contrast to *se* being semantically-vacuous in the personal variant in (22). However, it seems that this kind of analysis brings us back to the same problems

---

[3]Note that Schäfer (2017) does not discuss Slovenian examples, but rather uses the following *se*-sentence from French to exemplify the Voice head in (22b) (Schäfer 2017: 16):

    (i)   Trois maisons se sont louées (*par des   touristes) hier.
        three houses SE aux rented by some tourists yesterday
        'Three houses were rented (by some tourists) yesterday.'

that were discussed in the previous section – that is, unaccusatives should not be disallowed, since if *se* behaves like an ordinary referring DP, there is no principled reason as to why it cannot also correspond to the internal argument of an unaccusative predicate, especially given that VoiceP is not built in this case.

4.2 The syntactic and semantic features of the new proposal

To solve this issue, I will now propose a partial re-analysis of the construction, where the core idea is that *se* itself does not correspond to a syntactic argument, but is rather the overtly realized head of a VoiceP that introduces an open participant variable into the semantic representation.[4] Since I maintain that the impersonal *se*-construction projects all of its arguments into the syntax, I follow Rivero & Milojević Sheppard's (2003) analysis in assuming that there is a phonologically-null pronoun (henceforth, pro$_{IMP}$) corresponding to the external argument within the structure, which is distinct from *se*.

The main grammatical characteristic of pro$_{IMP}$ is that it is structurally impoverished – crucially, it lacks in its morphosyntactic make-up a layer which encodes ϕ-features. Following recent work on the morphosyntactic properties of person features (Ackema & Neeleman 2013; 2018), I assume that it is the lack of ϕ-features that gives pro$_{IMP}$ its idiosyncratically ambiguous interpretation with respect to its person reference. Concretely, this means that pro$_{IMP}$ allows for at least three readings, disambiguated by other sentential constituents or by the context: in existential sentences, (i) it can refer to a group of people that includes the speaker and/or addressee (23); (ii) it can refer to a group of people that excludes the speaker and/or addressee (24); and (iii) in the case of a generic context, it can refer to an arbitrary participant similar in interpretation to the English generic pronoun *one* (25) (see also the discussion in Fenger 2018, where Germanic *man*-type pronouns are also analysed as completely lacking ϕ-features and thereby allowing for the same kinds of interpretation).

(23) Včeraj   se je   na zabavi ves čas plesalo.   Imeli
      yesterday SE AUX.3SG on party   all time dance.PTCP.SG.N have.PTCP.PL.M
      smo   se prav super.
      AUX.1PL SE just great
      'Yesterday, we were dancing all the time at the party and we had such
      fun.'

---

[4]This assumption falls in line with the idea that clitics like *se* are categorically ambiguous in that they can either function as full-fledged XP's or X^0's (Bošković 2001: 31). In this sense, *se* can have the characteristics of an XP in that it can be merged as the specifier of VoiceP in personal *se*-sentences (22a) or the characteristics of an X^0 in that it is the head of a VoiceP in our proposed re-analysis for impersonal *se*-sentences. For a similar proposal for the Slovenian negative marker (though with different syntactic repercussions), see Ilc (2011).

(24) Včeraj   se nas   je   obvestilo   o   novi knjigi o
     yesterday SE us.ACC AUX.3SG inform.PTCP.SG.N about new book on
     skladenjski teoriji.
     syntactic   theory
     'Yesterday, someone/people told us about a new book on syntactic theory.'

(25) Ko   se gre   v Italijo, se navadno je   pašto.
     when SE go.3SG in Italy,   SE usually   eat.3SG pasta.ACC
     'When one goes to Italy, one usually eats pasta.'

The claim that pro$_{IMP}$ is morphosyntactically impoverished is further motivated by the fact that the verbal element which shows ϕ-feature contrasts is invariantly spelled out with third person singular features even though pro$_{IMP}$ typically has a plural interpretation glossed as 'people' (with a contextually determined person reference) in the translations.[5] Compare, for instance, the first sentence in (23), which is impersonal and thus spells out the auxiliary with non-agreeing third person singular features (i.e., *je*), with the second one, which is personal and thus spells out the auxiliary with first person plural features (i.e., *smo*).[6] According to Ackema & Neeleman (2013), such an apparent mismatch between number marking and meaning in impersonal *se*-structures is possible because third person singular represents the default spell-out of ϕ-features, and is therefore compatible with the absence of matching interpretable features in the pronoun.

In addition, the lack of ϕ-features in the pronoun explains a crucial distributional fact of the Slovenian impersonal *se*-construction that has gone – to the best of my knowledge – unnoticed in the formal literature. Namely, Slovenian impersonal *se*-sentences do not allow pro$_{IMP}$ to be used in the structural object position (26). In this respect, Slovenian pro$_{IMP}$ patterns with the overt impersonal

---

[5]More precisely, the number interpretation of pro$_{IMP}$ actually seems to be ambiguous between a plural and a singular reading; e.g. it is unclear how many people are denoted by pro$_{IMP}$ in (24). However, this ambiguity seems to be expected. In the contemporary semantic/pragmatic approach to grammatical number (e.g. Sauerland et al. 2005), a plural NP makes the same semantic contribution as the alternative singular NP (i.e., both logically mean 'one or more referents'), while the actual 'more than one meaning' of the former obtains through a process of pragmatic strengthening, by means of which the intended pragmatic meaning of the singular NP is negated. Crucially, such strengthening occurs only because a plural NP and the alternative singular NP display a ϕ-feature contrast in number and are seen as competitors for the intended number interpretation. By contrast, the null pronoun in Slovenian impersonal *se*-sentences is invariant (i.e., there is no other null impersonal pronoun in Slovenian with which it would contrast in number), so it does not have a competitor for the number interpretation. As a consequence, the process of pragmatic strengthening does not take place in this case.

[6]Consequently, the second sentence in (23) contains a typical agreeing pro with matching first person plural features as the grammatical subject.

pronoun *men* in Dutch (27a), and contrasts with generic *one* in English, which is allowed in such a position (27b).

(26) * Ta ideja   se spominja   na vojno.
        this idea.NOM SE reminds.3SG on war
        Intended: 'This idea reminds one of the war.'

(27)   a. * Dit herinnert men aan de oorlong.
            this reminds IMP of the war
            Intended: 'This reminds one of the war.'

       b.   This idea reminds one of the war.

                          (Examples under (27) taken from Fenger 2018: 298)

According to Fenger (2018), the Dutch sentence in (27a) is ungrammatical because *men* cannot be assigned accusative case, as the latter is only compatible with those nominal constituents that are able to project a K(ase) Phrase (Bayer et al. 2001). Crucially, only nominals that contain a rich enough morphosyntactic structure can project a KP. In this respect, Fenger (2018) claims that it is the lack of ϕ-features in the Dutch pronoun *men* that precludes the projection of the KP, whereas English *one* is richer in structure in the sense that it does contain a ϕ-layer encoding person features, and can thus occupy the object position where it gets accusative case.[7]

Note, now, that the proposed lack of ϕ-features on pro$_{IMP}$ has repercussions for how case is assigned in impersonal *se*-sentences. Minimalist theory has generally relied on what Kornfilt & Preminger (2015) call "a positively-specified account of case assignment", which means that nominative case is assigned to the grammatical subject only under ϕ-feature agreement with a specific functional head, which is canonically taken to be T (Chomsky 2001). However, the thing is that if pro$_{IMP}$ completely lacks ϕ-features, then nominative case cannot be assigned to it under the standard agreement relation, in which T, bearing uninterpretable ϕ-features, would find its goal with matching interpretable ϕ-features in the null pronoun.

Recall from §3.2, example (17), that Rivero & Milojević Sheppard (2003) circumvented this problem by assuming that nominative case is assigned to the null pronoun under a special variant of A-movement, whereby the null pronoun bypasses TP and checks its case feature against a higher functional projection

---

[7]According to Fenger (2018), English *one* has a set of ϕ-features that are underspecified for person (see also Ackema & Neeleman 2018). Such underspecification means that the pronoun "must be compatible with any arbitrary choice of person, including the speaker/hearer" (Fenger 2018: 307), which limits its distribution to generic contexts.

headed by the clitic. However, I would like to propose a solution that does not require recourse to such a stipulated modification of A-movement: that is, case assignment in impersonal *se*-sentences occurs configurationally (i.e., not under agreement), in the sense of recent approaches to dependent case (e.g. Levin & Preminger 2015; Kornfilt & Preminger 2015).

For transitive *se*-structures such as sentence (24) and the second clause in (25), this entails that the accusative case of the internal argument DP (e.g. *pašto* 'pasta' in (25)) is the dependent case, which means that it is assigned "in opposition to another argument position" (Marantz 2000: 24); that is, the internal argument DP gets dependent accusative case because it is in a position in which it is asymmetrically c-commanded by another DP (i.e., pro$_{\text{IMP}}$).[8] By contrast, nominative case is not assigned to the c-commanding DP, but "represent[s] the absence of any otherwise assigned case" (Kornfilt & Preminger 2015: 298) and is thus possible when no agreement in ϕ-features takes place.[9] On this account, the syntactic structure of a transitive impersonal *se*-sentence is as in Figure 1, with the dotted line representing dependent case assignment between pro$_{\text{IMP}}$ and the overt internal argument DP.
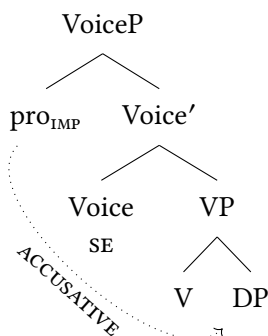
Figure 1: Case assignment in impersonal *se*-sentences

Having dependent case assignment in mind, let's now return to the Slovenian example in (26), where pro$_{\text{IMP}}$ functions as the grammatical object in the same

---

[8]There is an additional requirement: in order to obtain dependent accusative case, the DP must also not be independently assigned idiosyncratic case by a head which imposes its own lexical case requirements, as is typical of prepositions (Marantz's 2000 Disjunctive Case Hierarchy).

[9]For English, Fenger (2018: 313), following previous work (e.g. Schütze 1997), claims that unmarked case should be treated as different from default case, as the latter is spelled out only when the DP is part of a clauseless sentence, as in the exclamation *Me, intelligent?!*

manner as *men* in the Dutch example (27a). In this case, and in contrast with Figure 1, it is pro$_{IMP}$ that is assigned dependent accusative case because it is c-commanded by the grammatical subject *ta ideja* 'this idea', which occupies the SpecVoiceP position in a structure like Figure 1. This kind of structural configuration, however, crashes the derivation, since the morphosyntactically impoverished pro$_{IMP}$ is incompatible with dependent accusative case, which in contrast with nominative imposes structural requirements that the null pronoun, lacking ϕ-features, simply cannot satisfy (Fenger 2018). In other words, it is the lack of ϕ-features that restricts pro$_{IMP}$ to a position where it can only get unmarked nominative case.[10]

Finally, an account of impersonal *se*-sentences has to explain why the external argument is restricted to pro$_{IMP}$ (28a), and why the derivation crashes if a full-fledged referring DP is merged in its stead (28b).

(28) a. Svoje starše   se spoštuje.
          self   parents.ACC SE respect.3SG
          'People respect their parents.'

      b. * Janez   se spoštuje   svoje starše.
          Janez.NOM SE respect.3SG self   parents.ACC
          Intended: 'Janez respects his parents.'

At the beginning of this section, I have posited that *se* is the overt realisation of a special type of impersonal Voice head. This assumption is not without precedent, as there are in fact languages in which a Voice head is overtly realized. According to Legate (2014), one such language is Balinese, in which the Voice head in a low-register passive construction is realized as the morpheme *-a* (p. 44). Crucially, Legate (2014) claims that this morpheme does not display agreement with

---

[10]An anonymous reviewer suggests that the assignment of nominative case to pro$_{IMP}$ may be compatible with an approach which does not assume dependent case, but simply disassociates case assignment, which is still feature-driven, from the checking of ϕ-features (e.g. Bošković 2007). On such an approach, pro$_{IMP}$ would get nominative case by checking its [uCASE] feature against T after moving to the SpecTP position, but in contrast with the traditional minimalist approach (e.g. Chomsky 2001), such case-checking would not hinge on the parallel checking of ϕ-features. However, I am not sure how this approach explains the fact that pro$_{IMP}$ cannot occur in a position where it gets accusative case, given that the valuation of its case feature as either nominative or accusative would simply depend on the type of case-checking head that is available or closest to it (v or Voice for accusative and T for nominative). By contrast, pro$_{IMP}$'s limited distribution follows on the dependent case approach from the fact that nominative is the unmarked case, and is thus not really assigned to a DP (Kornfilt & Preminger 2015) in contrast with accusative.

the grammatical subject of the passive construction, but rather restricts the ϕ-features of the demoted external argument, which is introduced via the preposition *teken* 'by', to 3rd person. Consequently, a 1st person external argument PP is inadmissible in the construction (29b).

(29)   a.   Bli   Man nyidaang masih tepuk-a   teken Made Arini.
            brother man can   still   see-PASS.3 by   Made Arini
            'Brother Man can still be seen by Made Arini.'

       b. *Bli   Man nyidaang masih tepuk-a   teken tiang.
            brother man can   still   see-PASS.3 by   me
            Intended: 'Brother Man can still be seen by me.'

                                          (Balinese; Legate 2014: 44)

Formally, Legate (2014: 39) captures these facts by assuming that the denotation of Voice, which is a function that relates an external participant to an eventuality (Kratzer 1996), combines via a predicate modification rule called Restrict with the denotation of the morpheme, which is semantically a function that imposes a restriction on the kinds of ϕ-features that an initiator DP can have. In this sense, the VoiceP of sentence (29a) is as in Figure 2, where Voice not only introduces an open initiator variable, but specifies that the DP realising the initiator has to have the 3rd person feature.

$$
\begin{array}{c}
\text{VoiceP} \\
\diagup \qquad \diagdown \\
\text{[Spec]} \qquad\qquad \text{Voice}' \\
\lambda x \lambda e[\textsc{initiator}(x, e) \wedge \textsc{3person}(x) \wedge \\
\textsc{see}(e) \wedge \textsc{patient}(\textsc{brother man}, e)] \\
\end{array}
$$

Voice
$\lambda x \lambda e[\textsc{initiator}(x, e) \wedge \textsc{3person}(x)]$

VP
$\lambda e[\textsc{see}(e) \wedge$
$\textsc{patient}(\textsc{brother man}, e)]$

Voice
$\lambda x \lambda e[\textsc{initiator}(x, e)]$

-PASS.3
$\lambda x[\textsc{3person}(x)]$
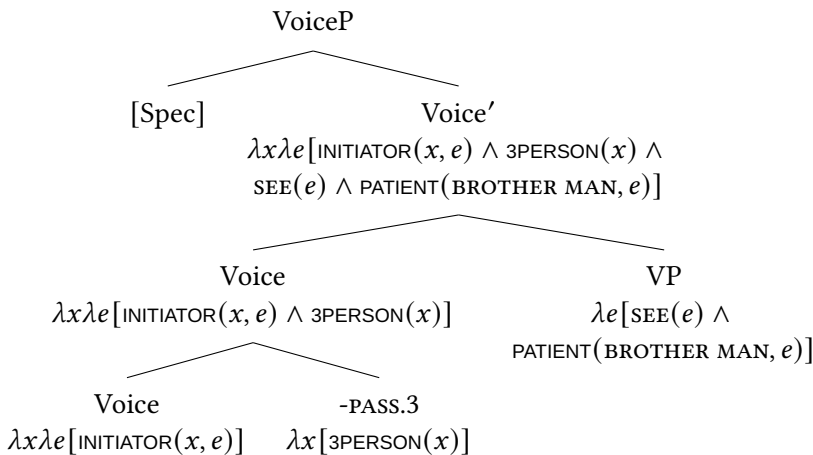
Figure 2: The VoiceP of sentence (29a)

For Slovenian impersonal *se*-sentences, I would tentatively like to propose that *se* operates in a similar manner to the Balinese morpheme, in that it is a Voice

head that restricts the external argument position to an impersonal pronoun, possibly banning all nominals that are specified for ϕ-features. The only difference is that an impersonal *se*-construction is not a passive construction; on the present account, this means that a Voice head realized as *se* has a syntactic D-feature in the sense of Schäfer's Voice typology discussed in §4.1, and thus needs a nominal (i.e., pro$_{\text{IMP}}$) to be merged in the specifier position. By contrast, the Balinese construction lacks this feature, which means that its external argument must be realized as a 'demoted argument'; that is, via the *teken* 'by' phrase, as is standardly assumed to be the case of passives.

4.3 The solution to the unaccusative problem

On this syntactic proposal, the ungrammaticality of impersonal *se*-sentences with unaccusative verbs follows from the fact that – as posited at the beginning of the previous subsection – *se* is the overt head of a VoiceP, so it is an element separate from the syntactically-projected thematic arguments. With this in mind, recall from the Unaccusative Hypothesis discussed in §3.2 that an unaccusative predicate like *die* initially merges its sole argument within VP (in simplified terms, the syntactic derivation of a sentence like *John died* is [$_{\text{TP}}$ John [$_{\text{VP}}$ died ~~John~~]], cf. (19)). It stands to reason, then, that in the Slovenian impersonal *se*-sentence headed by the unaccusative predicate *umreti* 'die' (30a), our proposed pro$_{\text{IMP}}$ is also merged VP-internally, as in (30b).

(30)   a.   * Včeraj   se je   umrlo.
             yesterday SE AUX.3SG die.PTCP.SG.N
             Intended: 'Some people died yesterday.'

         b.   [$_{\text{VP}}$ umrlo pro$_{\text{IMP}}$]

However, the VP in (30b) then combines with *se* heading a VoiceP, and such a *se* has semantic import in that it introduces an open initiator variable into the semantic representation. Crucially, this initiator variable introduced by *se* is distinct from the theme variable of *umreti* 'die', so the key prediction is that the derivation crashes because there is now an additional open variable that a one-place predicate like *umreti* 'die' cannot saturate. To better see the problem, compare Figure 3, which provides the semantic-syntactic representation of the VoiceP of the licit transitive sentence (1), repeated here as (31), with Figure 4, which illustrates the VoiceP of the illicit unaccusative sentence in (30a).

(31)   Celo leto se je   gradilo   hišo.
        whole year SE AUX.3SG build.PTCP.SG.N house.ACC
        'People were building the house for a whole year.'

VoiceP
$\lambda e[\text{INITIATOR}(\text{pro}_{\text{IMP}}, e) \wedge \text{BUILD}(e) \wedge \text{PATIENT}(\text{HOUSE}, e)]$

pro$_{\text{IMP}}$ Voice′
$\lambda x \lambda e[\text{INITIATOR}(x, e) \wedge$
$\text{BUILD}(e) \wedge \text{PATIENT}(\text{HOUSE}, e)]$

Voice   VP
$\lambda x \lambda e[\text{INITIATOR}(x, e)]$ $\lambda e[\text{BUILD}(e) \wedge$
SE $\text{PATIENT}(\text{HOUSE}, e)]$

V   DP
build   house
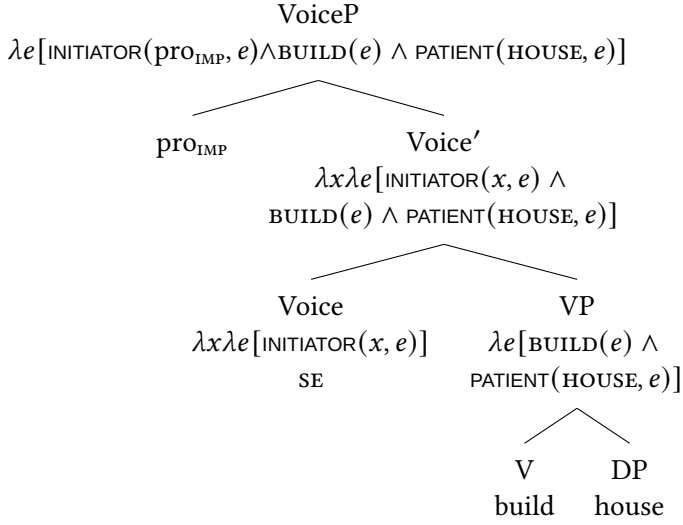
Figure 3: The VoiceP of sentence (31)

VoiceP
$\lambda e[\text{INITIATOR}(???, e) \wedge \text{DIE}(e) \wedge \text{PATIENT}(\text{pro}_{\text{IMP}}, e)]$

???   Voice′
$\lambda x \lambda e[\text{INITIATOR}(x, e) \wedge$
$\text{DIE}(e) \wedge \text{PATIENT}(\text{pro}_{\text{IMP}}, e)]$

Voice   VP
$\lambda x \lambda e[\text{INITIATOR}(x, e)]$ $\lambda e[\text{DIE}(e) \wedge \text{PATIENT}(\text{pro}_{\text{IMP}}, e)]$
SE

V   pro$_{\text{IMP}}$
die
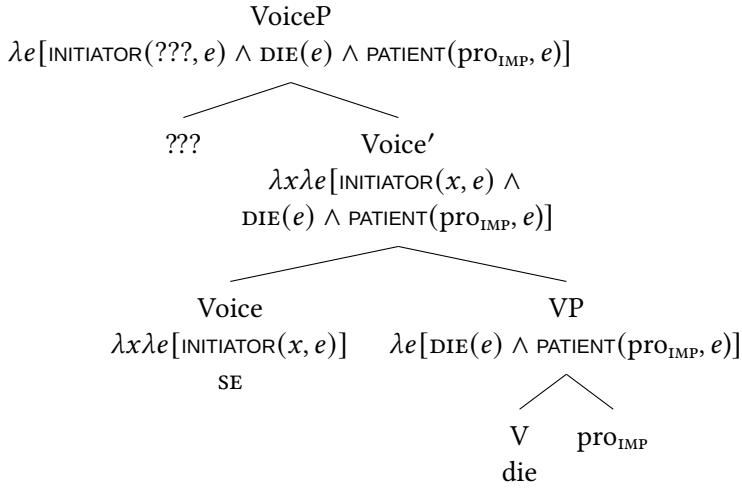
Figure 4: The VoiceP of sentence (30a)

In Figure 3, *se* introduces an initiator variable into the semantics of VP, which by itself denotes the characteristic function of the set of all events in which the house is built. The variable is then saturated by pro$_{\text{IMP}}$, which enters the deriva-

tion as the external argument in SpecVoiceP.[11] As a consequence, the propositional meaning that Figure 3 results in after its higher structure is built is 'People, who acted as volitional initiators (i.e., agents), were building the house', which is indeed how we interpret the Slovenian transitive impersonal *se*-construction. In the illicit structure in Figure 4, however, pro$_{\text{IMP}}$ enters the derivation as the internal argument of *umreti* 'die' so as to saturate its patient variable. The resultant interpretation of the unaccusative VP headed by *umreti* 'die' is one that denotes the characteristic function of the set of all events in which someone (i.e., the denotation of pro$_{\text{IMP}}$) dies. The problem, then, is that the θ-requirements of *umreti* 'die' – being a one-argument predicate – are already satisfied at the VP level, so *se*, which on our account invariably introduces an external participant variable, incorrectly turns *umreti* 'die' into a two-place predicate, violating the θ-Criterion. Lastly, one may wonder, as did an anonymous reviewer, why only an external argument is disallowed in an unaccusative sentence, while an applicative argument, such as the dative clitic *mu* in sentence (32), is allowed, observing that a general θ-Criterion-based constraint on argument structure augmentation would disallow the inclusion of both.

(32) Umrl   mu   je   pes.
      die.PTCP.SG.M him.DAT AUX.3SG dog.NOM
      'His dog died on him.'

Apart from claiming that ethical datives aren't 'real arguments' and thus not subjected to the θ-Criterion, I acknowledge that it needs to be stipulated that unaccusative predicates are somehow able to lexically specify how their default arity can be modified (i.e., by datives, but not by external arguments). In any case, note that, in languages like English and Slovenian, the VP of an unaccusative predicate also cannot combine with just about any Voice head. For instance, if an unaccusative VP merges with an ordinary active Voice head (20b), the result is an ungrammatical sentence:

(33) * Janez   je   umrl   Marijo.
        Janez.NOM AUX.3SG die.PTCP.SG.M Marija.ACC
        Intended: 'Janez caused Mary to die.'

It seems to me that a system of argument structure that introduces the external argument outside of the VP always requires recourse to a stipulation preventing

---

[11]The semantic translations of the nodes in Figure 3 and Figure 4 are – for ease of exposition – somewhat simplified with respect to pro$_{\text{IMP}}$, which remains untranslated. Following Ackema & Neeleman's (2018) account of impersonal pronouns with existential readings, I take it that pro$_{\text{IMP}}$ semantically contributes an existential quantifier that makes an arbitrary selection from the domain of individuals present in the discourse, which explains the ambiguous person reference discussed in §4.2.

an unaccusative predicate to merge its VP with an active Voice head, especially since the only formal requirement of active Voice is that its semantic argument (i.e., the denotation of VP) is of type ⟨s,t⟩, given Kratzer's (1996) rule of Event Identification.

4.4 Additional evidence – the generic impersonal *se*-construction

What the proposal outlined in the previous subsections entails is that, in simplified terms, an unaccusative impersonal *se*-sentence like (30a) is ungrammatical because an argument is missing in the syntactic structure. That this is so is, I believe, further corroborated precisely by the fact that unaccusatives can also be licit in the impersonal *se*-construction, but only in case their interpretation is generic. Compare the ungrammatical non-generic sentences (34) with the grammatical generic variants in (35), taken from Ilc & Marvin (2016).

(34) a. * Včeraj   se je   umrlo.
         yesterday SE AUX.3SG die.PTCP.SG.N
         Intended: 'Someone/People died yesterday.'

     b. * Včeraj   se je   padlo.
         yesterday SE AUX.3SG fall.PTCP.SG.N
         Intended: 'Someone/People fell yesterday.'

(35) a. Še dandanes se umre   od   kuge.
        still nowadays SE die.3SG from plague
        'People still die from the plague nowadays.'

     b. V tem hodniku se zlahka pade.
        in this corridor SE easily fall.3SG
        'Anyone can easily fall in this corridor.'

According to Härtl (2012: 95–97), a sentence that is interpreted generically often allows a lexically transitive predicate, such as *kill*, to be used intransitively (36a). If the interpretation were existential/episodic, such use would generally be ungrammatical (36b).

(36) a. The tiger kills to survive.

     b. * Yesterday, the tiger killed.
         Intended: 'Yesterday, the tiger killed something/someone.'

A similar improvement in acceptability can be observed in the so-called middle construction, where the predicate *read* in both (37a) and (37b) is used intransitively, in the sense that the external argument denoting the reader of the book is missing from the overt syntactic representation.

(37) a. The book reads easily.
          'Anyone can read this book easily.'
      b. * Yesterday, the book read.
          Intended: 'Someone was reading the book yesterday.'

The general idea behind this improvement is tied to the fact that the generic middle construction (37a) interpretatively involves a dispositional 'property' interpretation, which is ascribed to the grammatical subject (Lekakou 2004). In the case of sentence (37a), such a property interpretation can be observed from the fact that its interpretation corresponds to the paraphrase 'This book has properties such that it is easy to read' (Fagan 1992). According to Härtl, it is this property interpretation tied to genericity that licenses the intransitive use of *read* in (37a), in the sense that "the interpretation of a generic property can assist the [inference] of [the implicit] evaluative entity" (Härtl 2012: 114).[12] This means that (37a) is licit because the generic interpretation allows the hearer to easily infer the overtly missing initiator argument, who is an evaluative entity insofar as he/she can read the book easily due to some properties that it possesses.[13]

---

[12]To be clear, I do not wish to assume – and, if I understand him correctly, neither does Härtl – that these generic sentences are licit in contrast with their eventive variants because some kind of post-syntactic operation is at work, accessing the syntactic component after derivation and circumventing the θ-Criterion by suspending the predicate's need for a syntactically-projected external argument. Härtl (2012) himself remains quite neutral with respect to how and at which stage of derivation the argument can be dropped. At the end of this section, I propose a tentative solution, following Bruening (2012), in the sense that such sentences contain a null operator that blocks the merger of an overt DP by closing off the open initiator variable introduced by Voice at the stage when the syntactic derivation is still taking place.

[13]Note that the generic context is not the only environment that can license the intransitive use of a lexically transitive predicate. An anonymous reviewer cites the following two examples with the predicate *kill*, which omit an overt internal argument even though the interpretation is existential:

(i)   Jack the Ripper killed (someone) again last night.
(ii) Donald Glover killed (the audience) at SNL last night.

It seems to me that sentence (i), albeit non-generic, is licit because it also encourages a property-reading of the grammatical subject, in the sense that its interpretation is like that of a characterizing sentence; i.e., *Jack the Ripper killed again last night, as he is wont to do*. Note, also, that the sentence becomes degraded if the adverb is omitted – ?*Jack the Ripper killed last night*. In sentence (ii), *kill* is used in the unrelated sense 'to make a favourable impression on (someone)'. Such a sense of *kill* seems to be very context-specific (i.e., some kind of performance in front of an audience), so a property-meaning can be more easily inferred from the grammatical subject, i.e. *Donald Glover*, which might explain its generic-like behaviour in that it allows the omission of an overt internal argument.

The generically interpreted impersonal *se*-sentences in (35), which are licit in contrast with their eventive variants (34), also involve a property interpretation. However, in contrast with the middle construction in (37a), the property interpretation is not associated with the logical object but rather with an external source, which causes the logical object (i.e., people denoted by pro$_{\text{IMP}}$) to either fall or die. In this sense, (35a) is intuitively interpreted with respect to the properties of the plague, which is such that it still causes people to die nowadays. Along the same lines, (35b) has the interpretation that the corridor has certain properties such that it causes people to fall. In other words, the interpretation of generic impersonal *se*-sentences roughly corresponds to 'Something causes people to fall/die', and what is crucial is that the syntax we have proposed in the previous section does in fact predict this kind of quasi-two-argument interpretation. This is so because we have posited *se* to be the overt head of a Voice head which invariably introduces its own argument corresponding to an external participant/initiator.

The question, then, is why the generic sentences in (35) are licit if – on the present account – *se*'s formal requirement for an additional argument renders their eventive variants in (34) ungrammatical. To try and answer this, I would tentatively like to adopt a proposal by Bruening (2012), which pertains to the argument structure of the English middle construction. Bruening claims that an English middle sentence such as (37a), lacking a syntactic external argument, is derived via an operator that is defined as follows (Bruening 2012: 30):

(38) $\llbracket \text{MiddleOp} \rrbracket = \lambda f_{\langle e, \langle s, t \rangle \rangle} \lambda g_{\langle s, t \rangle} \mathsf{GN} e \exists x [f(x, e) \Rightarrow g(e)]$

This operator, which is merged during the syntactic derivation of VoiceP (Bruening 2012), targets the denotation of an unsaturated Voice projection and maps it into a semantic scope where the initiator variable gets existentially bound. This, in turn, also means that a referring DP cannot be merged as an external argument in the English construction, as its merger would saturate the open initiator variable and thus violate the semantic-type requirement of the first argument of the middle operator in (38), which has to be a function of type $\langle e, \langle s, t \rangle \rangle$.

I suggest that the generic *se*-sentences headed by unaccusative predicates undergo a similar operation tied to their valency. That is, an operator such as (38) targets an unsaturated Voice projection headed by *se* and thereby closes off the open initiator variable. As a result, a sentence like (35b) is interpreted as follows.

(39) $\mathsf{GN} e \exists x [\textsc{initiator}(x, e) \land \textsc{fall}(e) \land \textsc{patient}(\text{pro}_{\text{IMP}}, e) \Rightarrow \textsc{easy}(e)]$

The LF in (39) corresponds to the above-mentioned paraphrase – 'generally, there is an x such that it causes people (denoted by pro$_{\text{IMP}}$) to fall easily'. What is key

here is that the merger of the operator in (38) precludes the merger of an argument in the SpecVoiceP position, just as an external argument DP is blocked in the English middle construction by the operator. This semantic operation, however, is not available with the eventive versions. This means that the only way the unaccusatives in (34) can satisfy *se*'s requirement for an external participant is by introducing a referential argument in the SpecVoiceP position, which is – as discussed in the previous section – inadmissible as it flies in the face of the lexical selectional properties of unaccusatives.[14]

5 Conclusion

On the one hand, the paper has argued that all of the thematic arguments selected by a predicate enter the syntactic derivation of the impersonal *se*-construction. Evidence was given in the form of the causative alternation, since the fact that the impersonal *se*-construction with a causative predicate like *odpreti* always entails agentive participation does not trivially follow from an analysis in which the external argument does not enter the derivation. On the other hand, the paper has shown that the syntactic set-up of Rivero & Milojević Sheppard's (2003) account does not provide an explanation as to why the construction is incompatible with unaccusative verbs. In its stead, the paper has proposed an analysis of the construction in which the clitic heads a VoiceP and introduces an agent variable to the semantic representation of the VP. This variable is saturated by transitive and unergative predicates because they merge a null pronoun in the SpecVoiceP position. Unaccusative predicates, however, merge the pronoun within the VP but do not lexically select for another argument, so the participant variable introduced by *se* remains unsaturated. Finally, I have argued that the licit generic interpretation of the impersonal *se*-construction with unaccusative verbs diagnoses the fact that an unaccusative *se*-sentence is ruled out because of a missing argument.

---

[14]There is a conceptual issue here: even if there is a generic operator that saturates the open variable in the generic impersonal *se*-construction, then the construction must still somehow circumvent the syntactic requirement that an overt DP be merged in SpecVoiceP, given that it is an active construction in terms of Voice and thus has a D-feature that needs to be checked by a nominal argument landing in the SpecVoiceP position (cf. §4.1). For English middles, Schäfer (2007) claims that the internal argument satisfies this feature by stopping off in the SpecVoiceP position on its way to its final position in SpecTP. It may be the case that the pro_{IMP} in the internal argument position of generic impersonal *se*-sentences moves in a similar manner; however, this aspect of the derivation merits further research.

Abbreviations

1	1st person	M	masculine
3	3rd person	N	neuter
ACC	accusative	NOM	nominative
AUX	auxiliary	PASS	passive
Cl	clitic	PL	plural
EXPL	expletive	PTCP	participle
F	feminine	SG	singular
HUM	human	KP	kase phrase
IMP	impersonal	LF	logical form

Acknowledgements

References

Ackema, Peter & Ad Neeleman. 2013. Person features and syncreticism. *Natural Language & Linguistic Theory* 31(4). 901–950.   DOI:10.1007/s11049-013-9202-z

Ackema, Peter & Ad Neeleman. 2018. *Features of person: From the inventory of persons to their morphological realization.* Cambridge, MA: MIT Press.

Alexiadou, Artemis. 2010. On the morphosyntax of (anti-)causative verbs. In Malka Rappaport Hovav, Edit Doron & Ivy Sichel (eds.), *Lexical semantics, syntax, and event structure.* Oxford Scholarship Online. DOI:10.1093/acprof:oso/9780199544325.003.0009

Alexiadou, Artemis, Elena Anagnostopoulou & Martin Everaert. 2004. *The unaccusativity puzzle: Exploration of the syntax-lexicon interface.* Oxford: Oxford University Press.

Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2015. *External arguments in transitivity alternations: A layering approach.* Oxford: Oxford University Press.

Bayer, Josef, Markus Bader & Michael Meng. 2001. Morphological underspecification meets oblique case: Syntactic and processing effects in German. *Lingua* 111(4–7). 465–514.   DOI:10.1016/S0024-3841(00)00041-3

Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Amsterdam: Elsevier.

Bošković, Željko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38(4). 589–644. DOI:10.1162/ling.2007.38.4.589

Bruening, Benjamin. 2012. *By* phrases in passives and nominals. *Syntax* 16(1). 1–41. DOI:10.1111/j.1467-9612.2012.00171.x

Burzio, Luigi. 1986. *Italian syntax: A government-binding approach.* Dodrecht: Reidel.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1995. *Italian syntax and universal grammar.* Cambridge: Cambridge University Press.

D'Alessandro, Roberta. 2008. *Impersonal "si" constructions: Agreement and interpretation.* Berlin: De Gruyter.

Fagan, Sarah M. B. 1992. *The syntax and semantics of middle constructions: A study with special reference to German.* Cambridge: Cambridge University Press.

Fenger, Paula. 2018. How impersonal does *one* get? A study of *man*-pronouns in germanic. *The Journal of Comparative Germanic Linguistics* 21(3). 291–325. DOI:10.1007/s10828-018-9101-0

Franks, Steven. 1995. *Parameters of Slavic morphosyntax.* Oxford: Oxford University Press.

Givón, Talmy. 1998. Tale of two passives in Ute. In Masayoshi Shibatani (ed.), *Passive and voice*, 417–439. Amsterdam: John Benjamins. DOI:10.1075/tsl.16.14giv

Grahek, Sabina. 2008. Middles in Slovene. *Leeds Working Papers in Linguistics and Phonetics* 13. 44–75.

Härtl, Holden. 2012. Generic rescue: Argument alterations and the monotonicity condition. In Patrick Brandt & Eric Fuß (eds.), *Repairs: The added value of being wrong*, 95–130. Berlin: De Gruyter. DOI:10.1515/9781614510796.95

Ilc, Gašper. 2011. Jespersen's cycle in Slovenian. *Linguistica* 51(1). 349–363. DOI:10.4312/linguistica.51.1.349-363

Ilc, Gašper & Tatjana Marvin. 2016. Unaccusatives in Slovenian from a cross-linguistic perspective. In Franc Marušič & Rok Žaucer (eds.), *Formal studies in Slovenian syntax: In honor of Janez Orešnik*, 145–166. Amsterdam: John Benjamins. DOI:10.1075/la.236.07ilc

Kornfilt, Jaklin & Omer Preminger. 2015. Nominative as no case at all: an argument from raising-to-accusative in Sakha. In Andrew Joseph & Esra Predolac

(eds.), *MIT Working Papers in Linguistics 76: Proceedings of the 9th Workshop on Altaic Formal Linguistics*, 109–120. Cambridge, MA: MITWPL.

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

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese.* Cambridge, MA: MIT Press.

Lekakou, Marika. 2004. Middles as disposition ascriptions. In Cécile Meier & Matthias Weisgerber (eds.), *Proceedings of Sinn und Bedeutung 8*, 181–196. Konstanz: Universität Konstanz.

Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity: At the syntax–lexical semantics interface.* Cambridge, MA: MIT Press.

Levin, Theodore & Omer Preminger. 2015. Case in Sakha: Are two modalities really necessary? *Natural Language & Linguistic Theory* 33(1). 231–250. DOI:10.1007/s11049-014-9250-z

Marantz, Alec. 2000. Case and licensing. In Eric J. Reuland (ed.), *Arguments and case: Explaining Burzio's generalization*, 11–30. Amsterdam: John Benjamins. DOI:10.1075/la.34.04mar

Marelj, Marijana. 2004. *Middles and argument structure across languages.* Utrecht: LOT Publishing. https://www.lotpublications.nl/Documents/88_fulltext.pdf.

Perlmutter, David M. 1978. Impersonal passives and the unaccusative hypothesis. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society 38*, 157–189. Berkeley, CA: University of California.   DOI:10.3765/bls.v4i0.2198

Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4). 657–720. http://www.jstor.org/stable/4178836.

Rivero, María Luisa & Milena Milojević Sheppard. 2003. Indefinite reflexive clitics in Slavic: Polish and Slovenian. *Natural Language & Linguistic Theory* 21(1). 89–155.   DOI:10.1023/A:1021841517604

Sauerland, Uli, Jan Anderssen & Kazuko Yatsushiro. 2005. The plural is semantically unmarked. In Stephan Kepser & Marga Reis (eds.), *Linguistic evidence: Empirical, theoretical and computational perspectives*, 413–434. Berlin: De Gruyter. DOI:10.1515/9783110197549.413

Schäfer, Florian. 2007. Middles as voiced anticausatives. In Emily Elfner & Martin Walkow (eds.), *NELS 37: Proceedings of the 37th Annual Meeting of the North East Linguistic Society, Vol. 1.* Amherst, MA: GLSA Publications.

Schäfer, Florian. 2017. Romance and Greek medio-passives and the typology of voice. In Roberta D'Alessandro, Irene Franco & Ángel J.

Gallego (eds.), *The verbal domain*. Oxford: Oxford Scholarship Online. DOI:10.1093/oso/9780198767886.003.0006

Schäfer, Florian & Margot Vivanco. 2016. Anticausatives are weak scalar expressions, not reflexive expressions. *Glossa: a journal of general linguistics* 1(1). 18. 1–36. DOI:10.5334/gjgl.36

Schütze, Carson. 1997. *Infl in child and adult language: Agreement, case and licensing*. Cambridge MA: MIT. (Doctoral dissertation). http://hdl.handle.net/1721.1/10778.

Chapter 8

How to introduce instrumental agents: Evidence from binding in Russian event nominal phrases

Takuya Miyauchi

Tokyo University of Foreign Studies and
Japan Society for the Promotion of Science

The aim of this paper is to argue that instrumental agents in Russian are introduced by a silent P through binding phenomena by instrumental agents in event nominal phrases. Two assumptions are adopted in this paper: one is the absence of the DP layer in Russian based on binding phenomena and the other is a particular structure of event nominal phrases. I show that the appropriateness of proposing a silent P is supported by Generalized Case Realization Requirement in Russian and that the silent P is a lexical preposition, not a functional one due to its ability to bind objects out of PP.

Keywords: Russian, event nominal phrases, DP–NP, instrumental agents

1 Introduction

In this paper, I claim that instrumental agents in Russian are introduced by a silent P (ø), through binding phenomena by instrumental agents in event nominal phrases. Two assumptions are adopted here: one is the structure of event nominal phrases proposed by Miyauchi & Ito (2016) and Miyauchi (to appear), and the other is that Russian nominal phrases are not DP but NP. I discuss that Generalized Case Realization Requirement (GCRR) in Serbo-Croatian (Horvath 2014) can apply (at least partially) to Russian and show that setting a silent P is appropriate via this discussion. Finally, I point out a possibility to bind objects out of PP and demonstrate that the silent P is a lexical preposition, not a functional one.

The rest of this paper is organized as follows. In §2, I outline discussion on the structure of nominal phrases through Russian binding phenomena. §3 offers a syntactic account on a restriction of θ-roles of genitive nouns in event nominal phrases with some assumptions, based on Miyauchi & Ito (2016) and Miyauchi (to appear). In §4, I propose the syntactic structure of event nominal phrases containing instrumental agents and §5 shows the validity of GCRR in Russian. §6 points out that there are binding phenomena out of PP in Russian and the proposed structure is modified. Finally, §7 concludes the paper.

2 Russian binding phenomena and the structure of nominal phrases

The structure of Slavic nominal phrases without overt articles is controversial in terms of whether they have DP in their structure or not. Some researchers insist on the presence of DP even in Slavic (Progovac 1998; Rappaport 2002; Rutkowski 2002; Bašić 2004; Franks & Pereltsvaig 2004; Pereltsvaig 2007; Rutkowski & Maliszewska 2007; LaTerza 2016, etc.), while others maintain that nominal phrases in Slavic are NPs (Zlatić 1998; Trenkić 2004; Bošković 2005; 2007; 2009; Petrović 2011; Despić 2013, etc.).

In this paper, I investigate instrumental agents in event nominal phrases in Russian from the standpoint of the no-DP analysis.[1] In this section, I outline the discussion of the structure of nominal phrases through Russian binding phenomena, which gives support to the no-DP analysis, based on Despić's (2013) paradigm.[2] Despić (2013) argues that binding phenomena and Kayne's (1994) antisymmetry approach provide a key to examine the existence or absence of the DP projection. He concludes that there is no DP in Serbo-Croatian and that the D-like elements are adjuncts.

Following Despić (2013), I adopt Kayne's (1994) definition of c-command given in (1).[3]

(1)   X c-commands Y iff X and Y are categories, X excludes Y and every category that dominates X dominates Y.   (Kayne 1994: 16)

The definition of exclusion is as follows (2):

---

[1]Note that the NP/DP debate itself is beyond the scope of this paper. Please see the references cited above for arguments for and against the DP projection in Slavic.

[2]The content of this section is based on Miyauchi (2016). Please see Miyauchi (2016) for more details.

[3]I use this definition of c-command henceforth.

(2)   α excludes β if no segment of α dominates β.   (Chomsky 1986: 9)

The Russian sentences (3a), (4a), and (5a) are ungrammatical with co-reference between possessors and pronouns, while (3b), (4b), and (5b) are grammatical. There is a clear contrast between (3a), (4a), (5a) and (3b), (4b), (5b).

(3)   a.   * Kolin$_i$ poslednij fil'm sil'no ego$_i$ razočaroval.
           Kolya's latest   film really him disappointed
           Intended: 'Kolya$_i$'s latest film really disappointed him$_i$.'

      b.   Poslednij fil'm Koli$_i$ sil'no ego$_i$ razočaroval.
           latest   film Kolya.GEN really him disappointed
           'The latest film of Kolya$_i$ really disappointed him$_i$.'

(4)   a.   * Vanin$_i$ papugaj ukusil ego$_i$ včera.
           Vanya's parrot   bit   him yesterday
           Intended: 'Vanya$_i$'s parrot bit him$_i$ yesterday.'

      b.   Papugaj Vanin$_i$ ukusil ego$_i$ včera.
           parrot Vanya.GEN bit   him yesterday
           'The parrot of Vanya$_i$ bit him$_i$ yesterday.'

(5)   a.   * Sašin$_i$ mjač včera   udaril ego$_i$ po golove.
           Sasha's ball yesterday hit   him in head
           Intended: 'Sasha$_i$'s ball hit him$_i$ in the head yesterday.'

      b.   Mjač Saši$_i$ včera   udaril ego$_i$ po golove.
           ball Sasha.GEN yesterday hit   him in head
           'The ball of Sasha$_i$ hit him$_i$ in the head yesterday.'

Let us see (3) as a representative case. Following Despić (2013), I argue that (3a) is ungrammatical because the possessor *Kolin* 'Kolya's' binds the co-indexed pronoun *ego* 'him,' which results in Condition B violation. According to the reasoning in Despić (2013), this suggests that Russian nominal phrases lack the DP layer.[4] Figure 1 shows the structure of (3a).[5] Note that under Kayne's (1994) the-

---

[4]An anonymous reviewer argues that the Serbo-Croatian data in Despić (2013) corresponding to the examples (3)–(5) are not ungrammatical, which suggest Despić's (2013) conclusion about the presence/absence of the DP layer in Serbo-Croatian is questionable, but as I am only concerned with Russian here, I do not have much to add. I am only employing the reasoning and the structure of the argument developed in Despić (2013). My argument for the lack of DP in Russian is thus valid regardless of the quality of Despić's (2013) Serbo-Croatian data.

[5]The object *ego* 'him' in (3a) is scrambled and the word order of this sentence becomes SOV. For the sake of simplicity, however, scrambling is ignored in Figure 1. I take the basic word order in Russian to be SVO, following Isačenko (1966).

ory, specifiers are taken to be adjoined phrases. Consequently, specifiers are not distinguished from adjuncts and the bar-level notation does not make sense.[6]
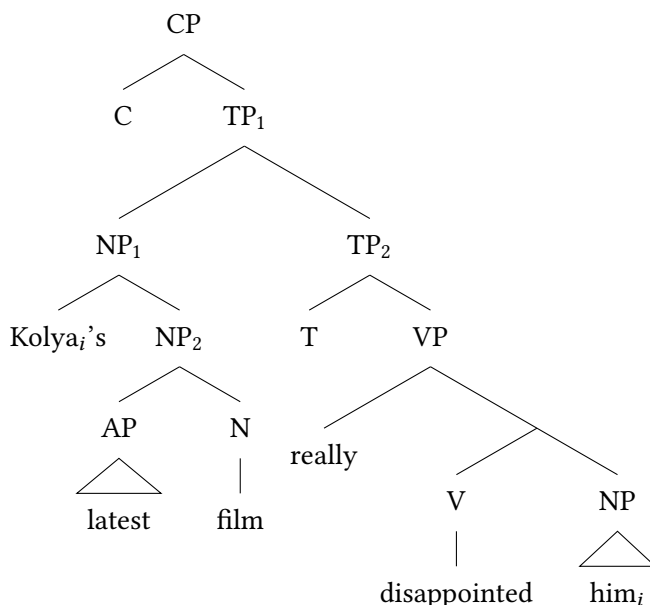
Figure 1: The structure of (3a)

In Figure 1, the first category dominating 'Kolya's' is CP, which also dominates the object NP 'him.'[7] Therefore, the possessor 'Kolya's' c-commands the object 'him', violating Condition B. Accordingly, (3a) is ungrammatical. If there was an additional DP layer in the nominal phrase, as illustrated in Figure 2, the possessor 'Kolya's' would not c-command the pronoun and there would thus be no Condition B violation. (3a) should be grammatical, contrary to fact.

In Figure 2, 'Kolya's' is dominated by the category DP, which does not dominate any other node outside of the nonimal phrase. That is to say, the possessor 'Kolya's', does not c-command the object NP. For this reason, with the DP layer, there would be no violation of Condition B in sentences like (3a) and these sentences should be grammatical. Thus, it is concluded that the ungrammaticality of (3a) shows that there is no DP layer in Russian.

---

[6]I do not use the bar-level notation. The conventional X′ (X-bar/X-prime) is written as XP or is omitted in the trees in this paper.

[7]Note that NP_1 and TP_1 are segments not categories.

```
                         DP
                    ┌─────┴─────┐
                    D          NP₁
                          ┌─────┴─────┐
                      Kolyaᵢ's        NP₂
                                 ┌─────┴─────┐
                                AP           N
                               ╱  ╲          │
                             latest        film
```
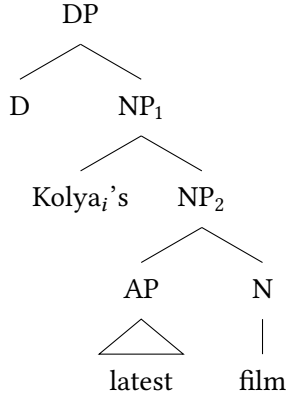
Figure 2: The structure of the subject of (3a) with a DP layer

How can we capture the grammaticality of (3b)? Generally, the genitive possessor NP is supposed to be located in the complement of the (head) NP (Franks 1995: 38; Bailyn 2012: 214; Mitrenina et al. 2012: 84).[8] Figure 3 represents the structure of (3b).

In this case, the categories dominating NP_{GEN} are NP_1 and CP. NP_{GEN} does not c-command the object NP because the subject NP_1 does not dominate the object NP. Thus, there is no Condition B violation and sentences like (3b) are grammatical.

The contrast in grammaticality between prenominal possessive constructions (3a), (4a), (5a) and postnominal ones (3b), (4b), (5b) supports the argument that the DP is not projected in Russian. For the rest of the paper, I adopt the no-DP analysis of Russian nominal phrases.

3 Russian event nominal phrases and their structures

3.1 Russian event nominals

An "event nominal" denotes an event or process and inherits argument structure of its base verb (Grimshaw 1990 in general, Schoorlemmer 1998, Pazelskaya 2007

---

[8]To be precise, Bailyn (2012) does not propose this structure. According to Bailyn (2012), adnominal genitives occupy the complement of QP in (i):

(i) $[_{NP}$ N $[_{QP}$ Q $[_{NP_{GEN}}$]]]   (Bailyn 2012: 214; slightly modified by the author)

Bailyn (2012: 214) proposes that Q assigns the genitive case to the sister NP (there is a case where Q is covert). However, these differences in the positions of genitive NP have no effect on the argument of this paper since genitive possessor NP is located lower than possessee NP.
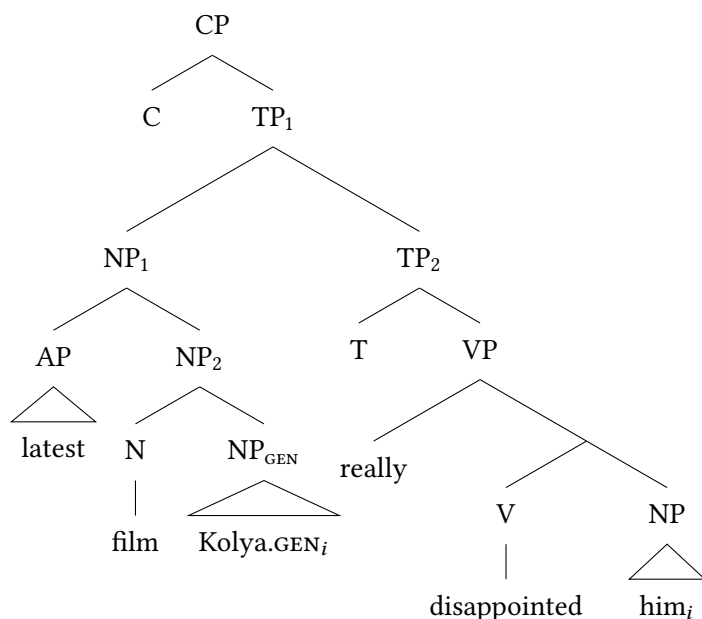
Figure 3: The structure of (3b)

for Russian).[9]

It can be followed by a genitive complement in Russian.

(6) a. Type 1: ✓ external argument / ✗ internal argument
udar {mužčiny / *stola}
hit man.GEN table.GEN
'the hit of {the man / the table}'

b. Type 2: ✓ external argument / ✓ internal argument
ispolnenie {Šaljapina / arii}
performance Chaliapin.GEN aria.GEN
'the performanc of {Chaliapin / the aria}'

c. Type 3: ✗ external argument / ✓ internal argument
razrušenie {*vraga / goroda}
destruction enemy.GEN city.GEN
'the destruction of {the enemy / the city}'

---

[9]The content of this section is mostly based on Miyauchi & Ito (2016) and Miyauchi (to appear).
Please see Miyauchi & Ito (2016) and Miyauchi (to appear) for more details.

This kind of restriction of genitive complements' θ-roles is thought to result from argument structures of event nominals (Pazelskaya 2007). Therefore these phenomena have been dealt with as lexical problems. Miyauchi & Ito (2016) and Miyauchi (to appear) tried to provide more principled explanation for these phenomena as syntactic problems based on phase theory.

I adopt the framework of Distributed Morphology (Halle & Marantz 1993), in which √ (root) moves to a categorizer (*n, v, a*, etc.) to determine its category. The contention of Miyauchi & Ito (2016) and Miyauchi (to appear) is that type 1 nominals and type 2/3 nominals differ structurally. I adopt the structure in Figure 4 for all event nominals and explain below how the two types differ.
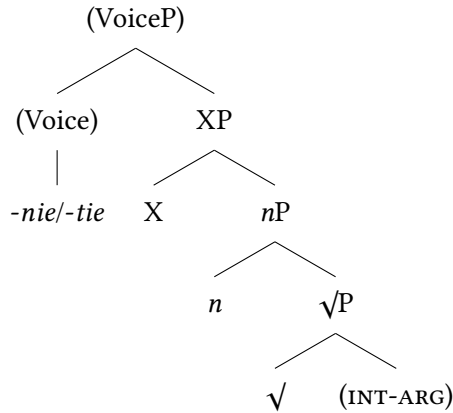
Figure 4: The structure of the event nominal phrases

Unlike in type 2/3, which project the entire structure, VoiceP is not projected in type 1. This structural difference is supported by the absence of a verbal nominalizer *-nie/-tie* in type 1 nominals.[10] I suppose that √ directly takes an internal argument following Harley (2009). Moreover, a functional head, X, licenses genitive Case through Agree.[11]

I assume Chomsky's (2000) phase theory and the proposal that *n*P is a phase (Carstens 2001; Arad 2003; Hicks 2009, etc.). It then follows from Phase Impenetrability Condition, shown in (7), that internal arguments cannot be genitive in type 1 nominals.

---

[10]This argument assumes that the nominalizer occupies the head of VoiceP. Support for this assumption comes from the fact that the nominalizer is morphologically complex and seems to include the passive participle morpheme *-n-/-t-* (Babby 1997).

[11]This X is a counterpart of Num(ber) in Carstens (2001), which is claimed to be a licenser of Case. Bailyn (2012), on the other hand, argues that the genitive case assigner in Russian is Q. The true identity of X lies outside the scope of this paper.

(7)   Phase Impenetrability Condition (PIC)
        In phase α with head H, the domain of H is not accessible to operations
        outside α, only H and its edge are accessible to such operations.

                                                                (Chomsky 2000: 108)

In type 2/3 nominals, the head of √P moves to the nominalizer *-nie/-tie*, the head
of VoiceP, in order to derive its form. Therefore, the *n*P phase slides up to VoiceP
due to PHASE-SLIDING (Gallego 2010).[12] The size of the new phase is shown with
a box in Figure 5.
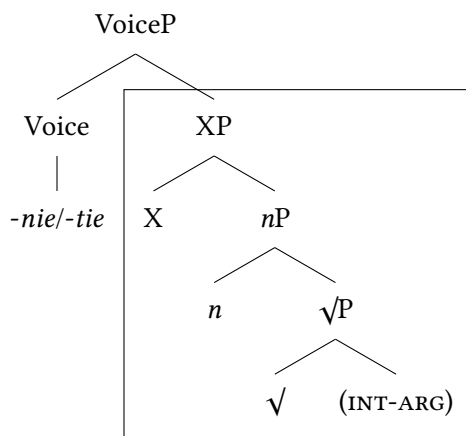
Figure 5: The structure of event nominal phrases with phase-sliding

Thus, phase-sliding makes it possible that X Agrees with an internal argument,
not violating PIC. Consequently, internal arguments are allowed to have genitive
Case.

3.2 Genitive external arguments

To avoid wrong prediction, I suppose two external θ-roles: possessor (POSS) and
agent (AG). In type 1/2, Poss is merged in the specifier of *n*P (Carstens 2000;
2001; Adger 2003, etc.). By contrast, AG in the type 3 is adjoined to VoiceP like

---

[12]According to phase-sliding, when a verb head-moves from v^* to T, the phasehood of v^* also
moves to T. I can generalize this proposal: when X, a phase, head-moves to Y, the phasehood
of X also moves to Y. Thus, phase-sliding can be applied to event nominals. In this example,
√ moves to VoiceP stopping at *n* and X, picking them up along the way because of the head
movement constraint (Travis 1984; Matushansky 2006). Since *n* is a phase head, phase-sliding
also occurs.

by phrases in English (Bruening 2013).[13] Poss is c-commanded by the probe, X and thus X can Agree with it as illustrated in Figure 6. Consequently, external arguments in type 1/2 can be genitive at the postnominal position. On the other hand, Ag is not c-commanded by the probe X and hence X cannot Agree with it as schematized in Figure 6. This is why external arguments in type 3 cannot appear in genitive case postnominally.[14]
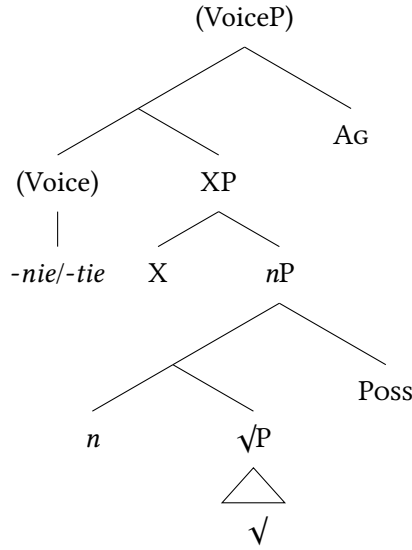
Figure 6: The structure of event nominal phrases with external θ-roles

4 Instrumental agents introduced by silent P

4.1 Russian event nominals with instrumental agents

Double genitives are basically banned in Russian as shown in examples (8a), (9a), where type 2 and type 3 nominals are used. When both internal and external

---

[13]In fact, a distinction between specifiers and adjuncts cannot be made under Kayne's (1994) theory. What is significant here, however, is the structural difference between Poss and Ag. That is, Poss is located in the *n*P domain, while Ag is in the VoiceP domain.

[14]In addition to these structures of event nominal phrases, Miyauchi (to appear) suggests that there are two types on *n*P through semantic analyses. Please see Miyauchi (to appear: section 5) for more details.

arguments are expressed in the same phrase, the former is assigned the genitive case and the latter is assigned the instrumental case as indicated in (8b), (9b).

(8) Type 2

    a.   *ispolnenie   arii   Šaljapina
        performance aria.GEN Chaliapin.GEN
        Intended: 'the performance of aria by Chaliapin'

<div align="right">(Ljutikova 2016: 162)</div>

    b.   ispolnenie   arii   Šaljapinym
        performance aria.GEN Chaliapin.INS
        'the performance of aria by Chaliapin'   (Gerasimova 2016: 64)

(9) Type 3

    a.   *razrušenie   goroda   vraga
        destruction city.GEN enemy.GEN
        Intended: 'the destruction of the city by the enemy'

    b.   razrušenie   goroda   vragom
        destruction city.GEN enemy.INS
        'the destruction of the city by the enemy'

In Serbo-Croatian, if the agent nominal is a complex phrase, the double genitive construction is perfectly acceptable as shown in (10).[15]

(10)   snimak požara Emira   Kusturice
       record   fire.GEN Emir.GEN Kusturica.GEN
       'the shot of the fire by Emir Kustiruca'

However, in Russian, even if agents are complex, double genitives are not permitted as indicated in (11).[16]

(11)   a.   *s'emka požara Ivana   Andreeviča
          shot   fire.GEN Ivan.GEN Andreevich.GEN
          Intended: 'the shot of fire by Ivan Andreevich'

      b.   *ispolnenie   arii   izvestnogo pevca   Šaljapina
          performance aria.GEN famous.GEN singer.GEN Chaliapin.GEN
          Intended: 'the performance of aria by the famous singer, Chaliapin'

---

[15]This was kindly pointed out by an anonymous reviewer.

[16]Although the micro-variation between Serbo-Croatian (10) and Russian (11) is significant, this paper cannot address this contrast since it focuses on only Russian.

Of course, as is the case with (8), (9), the phrases are acceptable when the complex agents are instrumental as show in (12).

(12)   a.   s'emka požara   Ivanom   Andreevičem
             shot   fire.ɢᴇɴ Ivan.ɪɴs Andreevich.ɪɴs
             'the shot of fire by Ivan Andreevich'

       b.   ispolnenie   arii   izvestnym pevcom   Šaljapinym
             performance aria.ɢᴇɴ famous.ɪɴs singer.ɪɴs Chaliapin.ɪɴs
             'the performance of aria by the famous singer, Chaliapin'

Instrumental phrases can appear as agents of ᴄᴏᴍᴘʟᴇx ᴇᴠᴇɴᴛ ɴᴏᴍɪɴᴀʟs (CEN), which have argument structures. CEN obligatorily take internal arguments and overtly express them (Revzin 1973; Schoorlemmer 1998, etc.). That is, an instrumental agent cannot appear without an internal argument unless it is elided as illustrated in (13)–(14).[17]

(13)   a.   *ispolnenie   Šaljapinym
              performance Chaliapin.ɪɴs
              Intended: 'the performance by Chaliapin'   (Revzin 1973: 90)

       b.   *razrušenie   vragom
              destruction enemy.ɪɴs
              Intended: 'the destruction by the enemy'

(14)   a.   ispolnenie   Δ Šaljapinym
              performance   Chaliapin.ɪɴs
              'the performance by Chaliapin of …'

       b.   razrušenie   Δ vragom
              destruction   enemy.ɪɴs
              'the destruction by the enemy of …'

The type 2 and 3 nominals have VoiceP as mentioned in §3. If I assume that instrumental agentive phrases are located at a domain related with VoiceP, it is natural that the type 1 nominals cannot have instrumental agentive phrases because of the absence of VoiceP. This is reflected in the ungrammaticality of (15).

---

[17]Δ in (14) expresses a deleted internal argument.

(15)   Type 1

 a. * udar stola mužčinoj
   hit table.GEN man.INS
   Intended: 'the hit the table by the man'

 b. * udar mužčinoj
   hit man.INS
   Intended: 'the hit by the man'

4.2 Binding in Russian event nominals and instrumental agents as VoiceP specifiers/adjuncts

Contrast in binding similar to the one presented in (3)–(5) is observed also with instrumental agents in event nominal phrases, as shown in (16) and (17).[18]

(16) a. * Ivanovo$_i$ narušenie pravil ogorčaet ego$_i$.
   Ivan's violation rules.GEN distresses him
   Intended: 'Ivan$_i$'s violation of the rules distresses him$_i$.'

 b. Narušenie pravil Ivanom$_i$ ogorčaet ego$_i$.
  violation rules.GEN Ivan.INS distresses him
  'The violation of the rules by Ivan$_i$ distresses him$_i$.'

(17) a. * Ivanovo$_i$ ubijstvo Viti gluboko opečalilo ego$_i$.
   Ivan's murder Vitya.GEN deeply saddened him
   Intended: 'Ivan$_i$'s murder of Vitya deeply saddened him$_i$.'

 b. Ubijstvo Viti Ivanom$_i$ gluboko opečalilo ego$_i$.
  murder Vitya.GEN Ivan.INS deeply saddened him
  'The murder of Vitya by Ivan$_i$ deeply saddened him$_i$.'

(16a) and (17a) are ungrammatical, while (16b) and (17b) are grammatical.

---

[18]An anonymous reviewer pointed out that in Serbo-Croatian the postnominal doubled genitive can co-refer a personal pronoun as shown in (i) below.

 (i) Snimak požara [Emira Kusturice]$_i$ napravio je od njega$_i$ reportersku
   record fire.GEN Emir.GEN Kusturica.GEN made AUX from him reporter
   zvezdu.
   star
   'The shot of the fire by [Emir Kusturica]$_i$ made him$_i$ star reporter.'

Unlike Serbo-Croatian, Russian bans doubled genitives as indicated in (11), so I have nothing further to say on this topic.

Applying the structure of event nominals in Figure 6 to (16a) and (16b), the structure of (16a) and (16b) is illustrated in Figure 7 and Figure 8. I assume that the instrumental agents are specifiers/adjuncts to VoiceP as the genitive external arguments in type 3 nominals.[19]
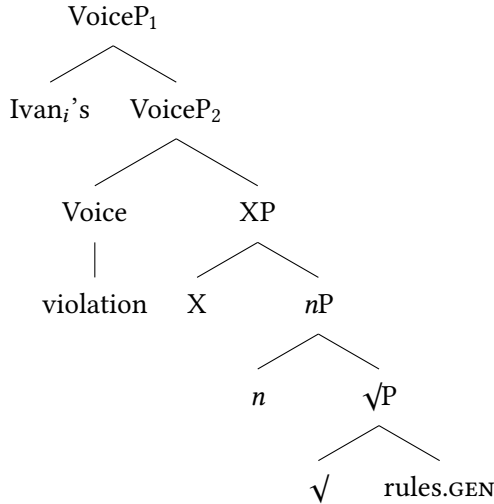
Figure 7: The structure of (16a)

As mentioned above, Schoorlemmer (1998) pointed out that only CEN can have the instrument agentive phrase. That is, X Agrees not with the specifier/adjunct but with the internal argument since X does not c-command the specifier/adjunct but only the internal argument, even under Kayne's (1994) definition of c-command. Thus, the specifier/adjunct cannot be genitive.

In Figure 7 and Figure 8, both the instrumental agents 'Ivan.INS' and the possessive adjective 'Ivan's' do c-command the objects *ego* 'him' since the VoiceP$_1$ and VoiceP$_2$ are segments. In other words, this structure correctly predicts the ungrammaticality of (16a), but it also incorrectly predicts (16b) to be ungrammatical. Thus, the contrast in (16) cannot be captured with Figure 7 and Figure 8. Something needs to be added to address the observed contrast in grammaticality.

I assume that instrumental agents are introduced as PP with a silent P head as in Figure 9. With this extra layer of structure there is no Condition B violation since the PP blocks the instrumental agent's c-commanding the object.

---

[19]This position can explain the fact that type 1 event nominals cannot have instrumental agents as shown in (15).

VoiceP$_1$

VoiceP$_2$ Ivan.INS$_i$

Voice XP

violation X *n*P

n √P

√ rules.GEN
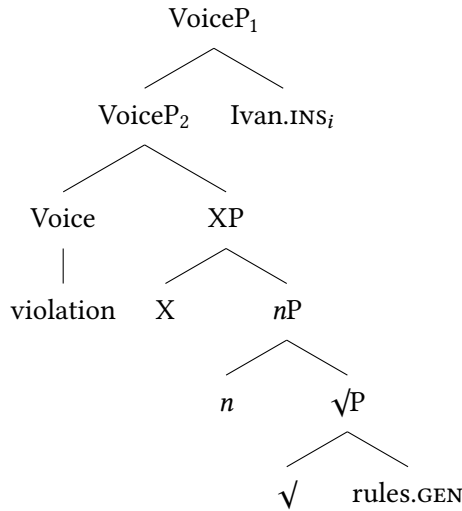
Figure 8: The structure of (16b)

5 Generalized case realization requirement

To capture the contrast in (16)–(17), I assumed that instrumental agents are introduces by the silent P. However, it is undesirable to utilize an abstract element with no evidence. Thus, I need some evidence except the contrasts in (16)–(17). In this section, I demonstrate that assuming the silent P in instrumental agent phrases is a consequence of GCRR, proposed by Horvath (2014).

5.1 GCRR in Serbo-Croatian

Horvath (2014) addressed the distribution of indeclinable nouns in Serbo-Croatian. As shown in (18), the indeclinable name *Miki* is ungrammatical although the declinable name *Larisa* is grammatical in the oblique environment.

(18) a. Divim se {Larisi / *Miki}.
admire.1SG REFL Larisa.DAT Miki
'I admire Larisa/Miki.'

b. Ponosim se {Larisom / *Miki}.
be.proud.1SG REFL Larisa.INS Miki
'I am proud of Larisa/Miki.'
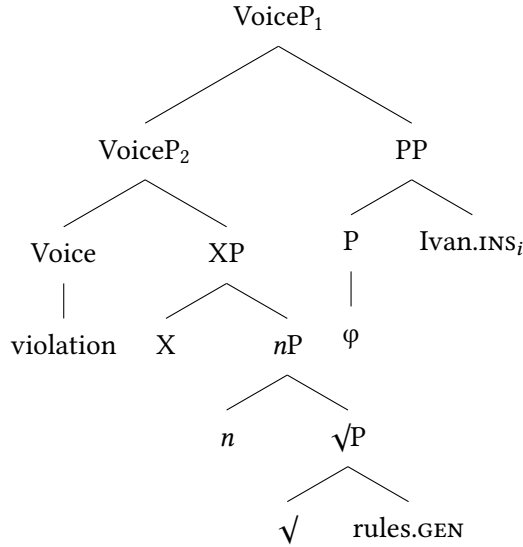
Figure 9: The modified version of the structure of (16b)

c. Oduševjena   sam   {Larisom / *Miki}.
   impressed.ptcp.f aux.1sg Larisa.ins   Miki
   'I am impressed by Larisa/Miki.'   (Horvath 2014: 121)

The indeclinable name is grammatical with the declinable possessive *moj* 'my'
or adjective *lep* 'beautiful', but it is ungrammatical without them or with the
indeclinable adjective *braon* 'brown', as in (19).

(19) a. Divim   se   *(mojoj)   Miki.
         admire.1sg refl   my.dat.sg Miki
         'I admire (my) Miki.'

      b. Oduševjena   sam   *(mojom)   Miki.
         impressed.sg.f aux.1sg   my.ins.sg Miki
         'I am impressed by (my) Miki.'

      c. Divim   se   {*braon / lepoj}   Miki.
         admire.1sg refl   brown   beautiful.dat.sg Miki
         'I admire {brunette/beautiful} Miki.'   (Horvath 2014: 121)

In addition, the indeclinable name is grammatical with P even without the de-
clinable possessive or adjective as illustrated in (20).

(20)   a. On je   trčao   prema (lepoj)   Miki.
          he AUX.3SG run.PTCP.SG towards beautiful.DAT.SG Miki
          'He ran towards (beautiful) Miki.'

       b. Dolazim sa   (mojom) Miki.
          come.1SG with my.INS Miki
          'I am coming with (my) Miki.'

       c. Razgovarali smo   o   (mojoj) Miki.
          talk.PTCP.PL AUX.1PL about my.LOC Miki
          'I talked about (my) Miki.'   (Horvath 2014: 122–123)

Accepting Pesetsky's (2013) theory of Case, Horvath (2014) generalized these complicated phenomena of Serbo-Croatian indeclinable nouns in the form of GCRR shown in (21).

(21)   *Generalized Case Realization Requirement (GCRR)*
       Oblique cases must be overtly realized by some element of the ASSIGNMENT DOMAIN (where assignment domain consists of the assigning head and the assignee – its noun phrase complement).   (Horvath 2014: 125)

According to GCRR, the sentences with the indeclinable name *Miki* in (18) are ungrammatical since no element in the assignment domain overtly realizes oblique cases. As for (19), sentences are grammatical even with the indeclinable name *Miki* if the declinable possessive *moj* 'my' or adjective *lep* 'beautiful' overtly realizes oblique cases. However it is ungrammatical without them or with the indeclinable adjective *braon* 'brown' because of the absence of overt realization of oblique cases. In the case of (20), each P *prema* 'towards,' *sa* 'with,' *o* 'about' manifests oblique cases and thus the sentences are grammatical even if there is no declinable possessive or adjective.

5.2 GCRR in Russian

There are also examples supporting the application of GCRR in Russian.[20]

---

[20]I consulted four Russian native speakers in their twenties for acceptability judgments on (22). All speakers found (22b) unacceptable, but there was variation among speakers on (22a) and (22c); three consider them acceptable (but unnatural) and the other considers them unacceptable. The speakers accepting (22a) also accept (22c) and *vice versa*. What is important here is that some speakers accept (22a), (22c) and that a clear difference in acceptability can be found between (22a)/(22c) and (22b).

(22) a. ? Professor priexal v Moskvu s   okolo pjati   studentov.
          professor arrived to Moscow with about five.GEN students.GEN
          'Professor arrived at Moscow with about five students.'

   b. * Professor rukovodit okolo pjati   studentov.
        professor supervise about five.GEN students.GEN
        Intended: 'Professor supervises about five students.'

   c. ? Professor rukovodit okolo pjat'ju studentami.
        professor supervise about five.INS students.INS
        'Professor supervises about five students.'

Sentence (22a) is grammatical with P requiring the instrumental case *s* 'with'. In this case, there is an overt P and it realizes the instrumental case in the assignment domain. However, (22b), without manifestations of the instrumental case, is ungrammatical. This is because neither P nor its complement overtly realizes the instrumental case. In addition, (22c) is grammatical since the complement *pjat' studentov* 'five students' is declined to bear the instrumental case, although the preposition *okolo* 'about' requires its complement in genitive Case. The (un)grammaticality of these sentences are predicted by GCRR, which means that GCRR is valid not only in Serbo-Croatian but also in Russian.

Pesetsky (2013) points out examples like (22). As shown in (23), without manifestations of the instrumental case, the sentences show low acceptability and with P requiring the instrumental case or with an instrumental adjective, the sentences are grammatical.

(23) a. Ja čital *Po kom zvonit kolokol.*
          I   read *For Whom the Bell Tolls*
          'I read *For Whom the Bell Tolls*.'

   b. ?? Ja vosxiščajus' *Po kom zvonit kolokol.*
        I   admire   *For Whom the Bell Tolls*
        'I admire *For Whom the Bell Tolls*.'

   c. ? Ja vosxiščajus' zamečatel'nym *Po kom zvonit kolokol.*
        I   admire   marvelous.INS *For Whom the Bell Tolls*
        'I admire the marvelous *For Whom the Bell Tolls*.'

   d. Pomnju,   kak diko   rydala   nad *Po kom zvonit kolokol.*
        remember.1 how wildly cried.SG.F over *For Whom the Bell Tolls*
        'I remember how wildly I cried over *For Whom the Bell Tolls*.'

(Pesetsky 2013: 132)

In addition, if GCRR is active, it is predicted that (24a) becomes ungrammatical when the declinable name *Ivan* is replaced with an indeclinable name, as in that case there is no manifestation of the instrumental case. This is confirmed in (24b), where the indeclinable name *Šmidt* is used.

(24) a. Narušenie pravil   Ivanom$_i$ ogorčaet ego$_i$.
        violation   rule.GEN Ivan.INS distresses him
        'The violation of the rules by Ivan$_i$ distresses him$_i$.'

  b. *Narušenie pravil   Šmidt$_i$ ogorčaet ego$_i$.
        violation   rule.GEN Schmidt distresses him
        Intended: 'The violation of the rules by Schmidt$_i$ distresses him$_i$.'

Thus there are reasons to assume the existance of a silent P, which introduces instrumental agent as proposed in Figure 9 since GCRR is (at least roughly) valid in Russian as mentioned above.

6 Binding out of PP

6.1 Functional and lexical prepositions

Yadroff & Franks (2001) proposed a distinction between FUNCTIONAL PREPOSI-TIONS (functinal P) such as *u* 'at', *k* 'toward', *bez* 'without' and LEXICAL PREPOSI-TIONS (lexical P) such as *okolo* 'around', *blagodarja* 'thanks to', *otnositel'no* 'with respect to' in various grammatical respects as illustrated in Table 1.

    To capture the various differences between functional P and lexical P, Yadroff & Franks (2001) assume two different syntactic structures for each type of P. Figure 10 and Figure 11 give their structure of the two PPs in (25).

(25) a. k   ženščinam   b. blagodarja ženščinam
        towards women.DAT   thanks.to   women.DAT
        'towards women'   'thanks to women'

"X" is used to indicate that here, *blagodarja* 'thanks to' is a bleached lexical item, which lacks a functional structure.[21]

---

[21]Please see Yadroff & Franks (2001) for more detail.

Table 1: Functional and lexical prepositions (Yadroff & Franks 2001: 70)

Functional prepositions		Lexical prepositions	
Phonology			
A.	Unstressed	A.	Stressed
B.	Monosyllabic	B.	Polysyllabic
Morphology			
C.	Monomorphemic	C.	Often polymorphemic or compound
D.	Prothetic *n-* before 3rd-person pronouns	D.	No prothetic *n-*
Syntax			
E.	Object is obligatory	E.	Object may be optional
F.	Approximative inversion yields N before P	F.	Approximative inversion yields P before N
G.	Negative particle *ni* does not intervene	G.	*Ni* intervenes
H.	May be doubled in colloquial language	H.	Cannot be doubled
I.	May be lexically selected	I.	Cannot be lexically selected
J.	Allow binding out of PP	J.	Block binding out of PP
K.	No intercalating particles	K.	Intercalating particles
L.	May govern multiple cases	L.	Govern one specific case
Semantics			
M.	Meaning abstract (hence polysemous)	M.	Meaning concrete (therefore fixed)
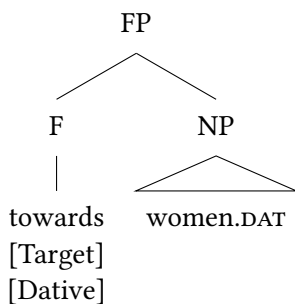
Figure 10: The structure of (25a) (Yadroff & Franks 2001: 77)
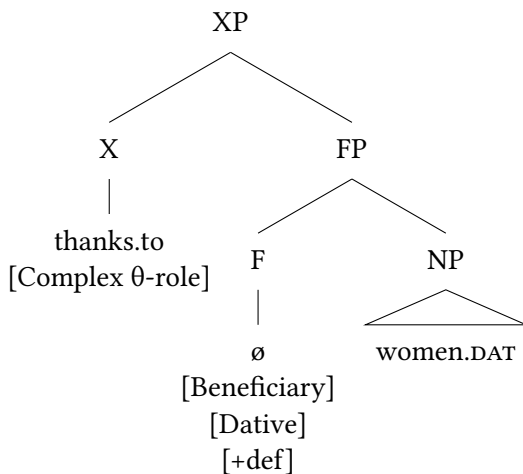
Figure 11: The structure of (25b) (Yadroff & Franks 2001: 78)

6.2 Binding possibility out of lexical PP

I showed that a silent P is required to capture the contrasts in (16)–(17) and that assuming a silent P is valid as shown in the data regarding GCRR in Russian in §5.

However, binding out of PP is not necessarily blocked as, for example, Yadroff & Franks (2001) and Bailyn (2010) point out. With regard to binding possibility

(J in Table 1), a functional P allows binding out of PP as shown in (26a) and a lexical P blocks binding out of PP as in (26b).[22]

(26)   a. U ėtogo čeloveka$_i$ vsegda est' svoi$_i$ original'nye idei.
          at this   person   always be self's original   idea
          'In that person's head there are always his own original ideas.'

      b. * Okolo ėtogo čeloveka$_i$ vsegda est' svoi$_i$ original'nye idei.
           around this   person   always be self's original   idea
           Intended: 'Around that person there are always his own original ideas.'

                                      (Yadroff & Franks 2001: 74)

Given the grammaticality of (26a), the silent P (ø) is a lexical P. Therefore, the structure of (16b), shown in Figure 9, should be modified to Figure 12.[23]

7 Conclusion

In this paper, I have argued that instrumental agents in Russian are introduced by a silent (lexical) P. As I have shown, this PP layer blocks binding of instrumental agents outside the event nominal, which is otherwise possible for agents introduced as possessors. My analysis which assumes Russian event nominals (or noun phrases more generally) lack the DP layer also offers (at least partial) support to the idea that GCRR is active in Russian.

---

[22]Yadroff & Franks (2001) point out that there is also a similar contrast in English.

    (i)   a.   John spoke to [Bill and Mary]$_i$ about each other$_i$'s birthdays.
          b. * John spoke about [Bill and Mary]$_i$ in each other$_i$'s houses.

                             (Yadroff & Franks 2001: 74)

Bailyn (2010) also points out data similar to (26). Binding out of PP *u Petrovyx* 'at the Petrovs' is allowed.

    (ii)   U Petrovyx$_i$ byla svoja$_i$ komnata.
          at the.Petrovs was self's.NOM room.NOM
          'The Petrovs had their own room.'   (Bailyn 2010: 14)

[23]There are two X(P)s in this tree. However, note that it is not guaranteed that they are the same projection. The "X" whose sister is *n*P represents a genitive assigner as described in Miyauchi & Ito (2016) and Miyauchi (to appear). The "X" whose sister is FP is used to indicate that the silent P is a bleached lexical item (Yadroff & Franks 2001).

VoiceP$_1$ tree:

VoiceP$_1$
- VoiceP$_2$
  - Voice
    - violation
  - XP
    - X
    - nP
      - n
      - √P
        - √
        - rules.GEN
- XP
  - X
    - ø
  - FP
    - F
      - ø
    - Ivan.INS$_i$
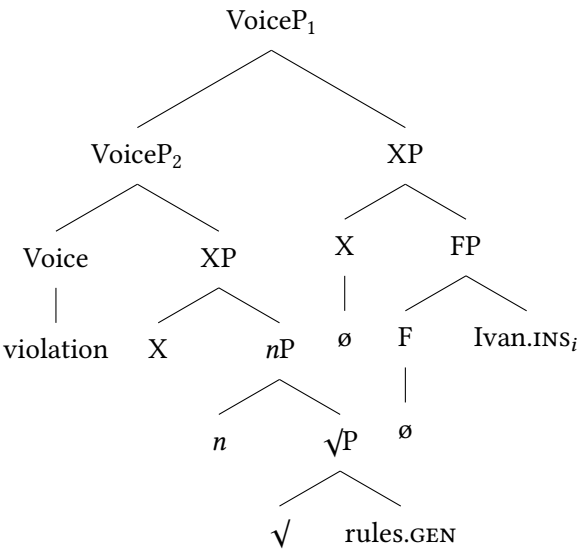
Figure 12: The structure of (16b)

Abbreviations

1	first person		INF	infinitive
3	third person		INT-ARG	internal argument
ACC	accusative		INS	instrumental
AG	agent (θ-role)		LOC	locative
AUX	auxiliary		PIC	phase impenetrability
CEN	complex event nominals			condition
DAT	dative		PL	plural
F	feminine		Poss	possessor (θ-role)
GCRR	generalized case		PTCP	participle
	realization requirement		REFL	reflexive
GEN	genitive		SG	singular

Acknowledgements

References

Adger, David. 2003. *Core syntax: A minimalist approach.* Oxford: Oxford University Press.

Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language & Linguistic Theory* 21(4). 737–778. DOI:10.1023/A:1025533719905

Babby, Leonard H. 1997. Nominalization, passivization, and causativization. *Die Welt der Slaven* 42(2). 201–251.

Bailyn, John F. 2010. Kinds of derivational binding. In Gerhild Zybatow, Philip Dudchuk, Serge Minor & Ekaterina Pshehotskaya (eds.), *Formal studies in Slavic linguistics: Proceedings of Formal Description of Slavic Languages 7.5*, 11–29. Frankfurt am Main, Berlin: Peter Lang.

Bailyn, John F. 2012. *The syntax of Russian.* Cambridge: Cambridge University Press.

Bašić, Monika. 2004. *Nominal subextractions and the structure of NPs in Serbian and English.* Tromsø: University of Tromsø. (MA thesis).

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59(1). 1–45. DOI:10.1111/j.1467-9582.2005.00118.x

Bošković, Željko. 2007. 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*, vol. 1, 101–114. GLSA Publications.

Bošković, Željko. 2009. More on the no-DP analysis of article-less languages. *Studia Linguistica* 63(2). 187–203. DOI:10.1111/j.1467-9582.2009.01158.x

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

Carstens, Vicki. 2000. Concord in minimalist theory. *Linguistic Inquiry* 31(2). 319–355. DOI:10.1162/002438900554370

Carstens, Vicki. 2001. Multiple agreement and case deletion: Against φ-(in)completeness. *Syntax* 4(3). 147–163. DOI:10.1111/1467-9612.00042

Chomsky, Noam. 1986. *Barriers.* Cambridge, MA: MIT press.

Chomsky, Noam. 2000. Minimalist inquiries. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Despić, Miloje. 2013. Binding and the structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44(2). 239–270.   DOI:10.1162/LING_a_00126

Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. Oxford: Oxford University Press.

Franks, Steven & Asya Pereltsvaig. 2004. Functional categories in the nominal domain. In Olga Arnaudova, Wayles Browne, María Luisa Rivero & Danijela Stojanović (eds.), *Formal Approaches to Slavic Linguistics 12: The Ottawa Meeting 2003*, 109–128. Ann Arbor, MI: Michigan Slavic Publications.

Gallego, Ángel J. 2010. *Phase theory*. Amsterdam: John Benjamins.

Gerasimova, Anastasija. 2016. Parametričeskij podxod k vnutrijazykovomu var'irovaniju: Problemy i metody. *Rhema.Rema* 3. 63–74.

Grimshaw, Jane. 1990. *Argument structure*. Cambridge, MA: MIT Press.

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

Harley, Heidi. 2009. The morphology of nominalizations and the syntax of *v*P. In Anastasia Giannakidou & Monika Rathert (eds.), *Quantification, definiteness, and nominalization*, 320–342. Oxford: Oxford University Press.

Hicks, Glyn. 2009. *The derivation of anaphoric relations*. Amsterdam: John Benjamins.

Horvath, Julia. 2014. A note on oblique case: Evidence from Serbian/Croatian. In Anna Bondaruk, Gréte Dalmi & Alexander Grosu (eds.), *Advances in the syntax of DPs: Structure, agreement, and case*, 221–246. Amsterdam: John Benjamins.

Isačenko, Aleksandr. 1966. O grammatičeskom porjadke slov. *Voprosy Jazykoznanija* 6. 27–34.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

LaTerza, Ivana. 2016. Binding in English and south Slavic and the parameterized DP hypothesis. *Linguistic Inquiry* 47(4). 741–753.   DOI:10.1162/LING_a_00230

Ljutikova, Ekaterina. 2016. *Sintaksis imennoj gruppy v bezartikevom jazyke*. Moscow: MGU. (Doctoral dissertation).

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1). 69–109.   DOI:10.1162/002438906775321184

Mitrenina, Ol'ga V., Evgenija Romanova & Natalija Slioussar. 2012. *Vvedenie v generativnuju grammatiku*. Moskva: Knižnyj Dom LIBROKOM.

Miyauchi, Takuya. 2016. Sokubakugensyō to rosiagomeisiku no tōgokōzō [Binding phenomena and the structure of nominal phrases in Russian]. *Suravu Bunka Kenkyu [Slavic Culture Studies]* 13. 69–106.

Miyauchi, Takuya. to appear. A phase-theoretic account of restriction on θ-roles of postnominal genitives in Russian event nominal phrases. In Tania Ionin & Jonathan MacDonald (eds.), *Formal Approaches to Slavic Linguistics 26: The Urbana-Champaign Meeting 2017*. Ann Arbor, MI: Michigan Slavic Publications.

Miyauchi, Takuya & Katsumasa Ito. 2016. Dekigotomeisiku ni okeru zokkaku syōgō no mekanizumu: rosiago, doitugo kara no kensyō [Mechanism of genitive case checking in event nominal phrases: Evidence from Russian and German]. In Linguistic Society of Japan (ed.), *Proceedings of the 153rd meeting of the Linguistic Society of Japan]*, 158–163. Kyoto: Nakanishi Printing Company.

Pazelskaya, Anna. 2007. Argument structure in Russian deverbal nouns in *-nie*. In Franc Marušič & Rok Žaucer (eds.), *Studies in formal Slavic linguistics: Contributions from Formal Description of Slavic Languages 6.5*, 255–272. Frankfurt am Main: Peter Lang.

Pereltsvaig, Asya. 2007. The universality of DP: A view from Russian. *Studia Linguistica* 61(1). 59–94. DOI:10.1111/j.1467-9582.2007.00129.x

Pesetsky, David. 2013. *Russian case morphology and the syntactic categories*. Cambridge, MA: MIT Press.

Petrović, Borivoje. 2011. The DP category in articleless Slavic languages. *Jezikoslovlje* 12(2). 211–228.

Progovac, Ljiljana. 1998. Determiner phrase in a language without determiners (with apologies to Jim Huang 1982). *Journal of Linguistics* 34(1). 165–179. DOI:10.1017/S0022226797006865

Rappaport, Gilbert. 2002. Numeral phrases in Russian: A minimalist approach. *Journal of Slavic Linguistics* 10(1–2). 327–340.

Revzin, Isaak I. 1973. Transformacionnoe issledovanie konstrukcij s sub"ektnym i ob"ektnym priimennym dopolneniem (genitivus subjectivus i genitivus objectivus). In Andrej A. Zaliznjak (ed.), *Problemy grammatičeskogo modelirovanija*, 88–96. Moskva: Nauka.

Rutkowski, Paweł. 2002. Noun/pronoun asymmetries: Evidence in support of the DP hypothesis in Polish. *Jezikoslovlje* 3(1–2). 211–228.

Rutkowski, Paweł & Hanna Maliszewska. 2007. On prepositional phrases inside numeral expressions in Polish. *Lingua* 117(5). 784–813. DOI:10.1016/j.lingua.2006.03.004

Schoorlemmer, Maaike. 1998. Complex event nominals in Russian: Properties and readings. *Journal of Slavic Linguistics* 6(2). 205–255.

Travis, Lisa. 1984. *Parameters and effects of word order variation*. Cambridge, MA: MIT. (Doctoral dissertation).

Trenkić, Danijela. 2004. Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua* 114(11). 1401–1427. DOI:10.1016/j.lingua.2003.09.005

Yadroff, Michael & Steven Franks. 2001. The origin of prepositions. In Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn & Luka Szucsich (eds.), *Current issues in formal Slavic linguistics*, 69–79. Frankfurt am Main: Peter Lang.

Zlatić, Larisa. 1998. *Slavic noun phrase are NPs not DPs*. Paper presented at the Workshop on Comparative Slavic Morphosyntax, Bloomington, Indiana, June 1998.

Chapter 9

Constraining the distribution of the perdurative in Russian

Ilya Naumov

National Research University Higher School of Economics

In this paper, I explore the constraints on the distribution of the perdurative prefix *pro-* in Russian. Applying several diagnostics proposed by Tatevosov (2009; 2013), I show that the perdurative *pro-* is a "selectionally restricted" prefix associated with an additional restriction: it can combine with a predicate built on a secondary imperfective stem only under a pluractional interpretation. I argue that this restriction is an instantiation of a more general semantic requirement imposed by the perdurative: it can be formed from a predicate if there are no subevents making up the activity component of this predicate that are in Landman's (1992) "stage-of relation".

Keywords: aspect, (a)telicity, subeventual structure, verbal prefixes, perdurative, Russian

1 Introduction

This chapter focuses on the distribution of the perdurative in Russian. So-called perdurative verbs in Russian are verbs marked by the so-called perdurative prefix *pro-*. Perdurative-prefixed verbs convey the meaning of temporal duration. The addition of the prefix to a verb stem makes a co-occurring measure expression obligatory, see (1a) and (1b).[1]

---

[1]There is no consensus on the syntactic status of measure expressions occuring with perdurative verbs. A common view is that they are adjuncts (Borik 2006; Ramchand 2008; Gehrke 2008: a.o.). Some researchers argue for a quasi-argument status for these phrases (Fowler & Yadroff 1993). Schoorlemmer (1995) distinguishes different types of accusative phrases: some of them are treated as arguments, while others as adjuncts. The most recent and at the same time the

(1)   a.   Lena govori-la (ves' večer)   po telefonu.
           L.   talk-PST   all   evening on phone
           'Lena talked on the phone for the whole evening.'

      b.   Lena pro-govori-la   *(ves' večer)   po telefonu.
           L.   PRO_PER-talk-PST   all   evening on phone
           'Lena spent the whole evening talking on the phone.'

While the syntax (Babko-Malaya 1999; Ramchand 2004; Svenonius 2004; Romanova 2007; Gehrke 2008; Žaucer 2009; Tolskaya 2015, a.o.) and semantics (Isačenko 1960; Bondarko & Bulanin 1967; Forsyth 1970; Flier 1985; Zaliznjak & Šmelev 2000; Gehrke 2003; Filip 2005, a.o.) of the perdurative in Russian has been extensively discussed, the constraints on its distribution, to the best of my knowledge, have not been the subject of a separate research. Nevertheless, data suggest that these constraints are intricate. For example, the perdurative can be combined with some predicates describing atelic processes, (2) (recall also (1b)), but not with others, (3).

(2)   * Kolja pro-pisa-l   pismo vsjo utro.
        K.   PRO_PER-write-PST letter all   morning
        Intended: 'Kolja spent the whole morning writing the letter.'

(3)   Kolja pro-taska-l   čemodan neskol'ko časov.
      K.   PRO_PER-carry-PST suitcase   several   hours
      'Kolja spent several hours carrying the suitcase.'

As Pazelskaya & Tatevosov (2006) convincingly argue, so-called simplex imperfective stems, like *pisa-* and *taska-*, project an activity event structure. With respect to common diagnostics on subeventual complexity (such as the interpretation of *opjat'* 'again' and negation) both *pisat' pis'mo* 'write a letter' and *taskat' čemodan* 'carry a suitcase' demonstrate identical properties: they are subeventually simplex denoting events that do not lead to the attainment of the result state. If nothing else is added here, it remains a mystery why the perdurative is bad from the former and fine from the latter.

   Moreover, the perdurative cannot be derived from predicates describing some goal-oriented activity that stops before the corresponding result state is attained. An example of such a predicate is *otkryvat' zamok* 'open a lock'. The unavailability of the perdurative from this predicate is demonstrated in (4).

---

most radical analysis is presented in Žaucer (2009; 2012), where all accusative phrases are claimed to be unselected direct objects introduced by the prefix. I will not go into this problem here and refer the interested reader to the cited works.

(4)   *Situation:* The lock in the door is broken. After spending half an hour trying to open it, Vasja gave up.

    *Vasja pro-[[otkr]-yva]-l   zamok polčasa.
    V.   PRO_{PER}-open-YVA-PST lock   half.an.hour
    Intended: 'Vasja spent half an hour opening the lock.'

Unlike in the examples above, in (4) the perdurative attaches to the so-called secondary imperfective verb stem containing the so-called secondary imperfective suffix *-yva*. It has been claimed by Altshuler (2013; 2014) and Tatevosov (2017) that in this type of predicates *-yva* functions as a partitive operator. It extracts (proper) non-final parts (more precisely, stages) from the extension of the original telic event predicate *otrkyt' zamok* 'open a lock'. Why cannot the perdurative successfully operate on this eventuality description?

In principle, the perdurative can be derived from a predicate based on the secondary imperfective verb stem but only if this predicate has a pluractional interpretation. Consider the examples in (5). While the sentence in (5a) has two possible readings, the sentence in (5b), which contains a perdurative-prefixed verb, is only compatible with a scenario in which the agent opened his mouth repeatedly.[2] That is, the restriction on the possible interpretation arises at the stage when the perdurative attaches to the stem.

(5)   a.   Kolja otkr-yva-l   rot   minutu.
         K.   open-YVA-PST mouth minute
         'Kolja was opening the mouth for a minute.'

        i.   ✓ Scenario №1: Kolja was opening the mouth wider and wider for a minute until he got cramps in his cheekbones.

        ii.   ✓ Scenario №2: Kolja was opening the mouth several times for a minute until he realized that he was not producing any sounds.

    b.   Kolja pro-[[otkr]-yva]-l   rot   minutu.
         K.   PRO_{PER}-open-YVA-PST mouth minute
         'Kolja spent a minute opening the mouth.'

        i.   ✗ Scenario №1: Kolja spent a minute opening the mouth wider and wider until he got cramps in his cheekbones.

        ii.   ✓ Scenario №2: Kolja spent a minute opening the mouth several times until he realized that he was not producing any sounds.

---

[2]I would like to thank an anonymous reviewer for suggesting scenario №1 for examples in (5).

Although predicates describing non-culminating events, as in (4), and plural events, as in (5b), are made up from the same morphological pieces that come in the same order (the secondary imperfective suffix *-yva* attached before the perdurative *pro-*), they still differ with respect to the availability of the perdurative. What is the underlying property distinguishing these classes of predicates that the perdurative is sensitive to?

In the remainder of the paper I will explore restrictions on the distribution of the perdurative in greater detail. In §2, I will show that the distribution of the perdurative is subject to (morpho)syntactic constraints. Applying the diagnostics that were proposed by Tatevosov (2009; 2013), I will argue that the perdurative *pro-* is a selectionally restricted prefix. Taking into account the interaction of the perdurative *pro-* with some other affixes, I will claim that it originates in a functional projection between *v*P and AspP. In §3, I will present empirical evidence indicating that the distribution of the perdurative is also subject to semantic constraints. Namely, I will argue that predicates allowing the derivation of the perdurative form a natural class with respect to one semantic property: subevents making up the activity component of these predicates are not in Landman's (1992) "stage-of relation". The main results and several open questions are listed in §4.

2 Determing the position of the prefix

2.1 The perdurative *pro-* in the big picture

One of the most important observations concerning prefixes in Russian and, more broadly, in Slavic languages that has been made so far is that they form a heterogeneous class and fall at least into two types: LEXICAL PREFIXES (LPs) and SUPERLEXICAL PREFIXES (SLPs), or internal and external in other terminology. For the first time this dichotomy was argued for by Babko-Malaya (1999) and subsequently became the subject of intense discussion (see, e. g., Svenonius 2004; Ramchand 2004; Romanova 2004; 2007; Tolskaya 2015). SLPs have several characteristics that distinguish them from lexical ones:

- External prefixes merge outside VP.

- External prefixes are attached over the internal ones.

- External prefixes do not affect the argument structure of the verb stem, or predictably decrease its transitivity.

- External prefixes are semantically compositional.

- External prefixes express temporal or quantifying meanings.

Among the works listed above, there is no agreement on the position of the perdurative *pro-* within this classification. Some authors include it in the list of SLPs, while others do not.

Table 1: The perdurative *pro-* as an SLP

Babko-Malaya (1999)	Ramchand (2004)	Svenonius (2004)	Romanova (2007)	Gehrke (2008)	Tolskaya (2015)
+	+	−	−	+	−

This binary opposition has been recently refined. Focusing on Russian data, Tatevosov (2009; 2013) claims that prefixes commonly subsumed under the label SLPs fall, in fact, into at least two distinct groups with respect to constraints that regulate their distribution. Namely, there are prefixes that demonstrate selectional restrictions, (6a), and prefixes that demonstrate positional restrictions, (6b).[3]

(6) Possible restrictions on the distribution of SLPs:
   a. Selectional restriction. The possibility of attaching a prefix to a stem can be constrained by the stem's formal (im)perfectivity.
   b. Positional restriction. The possibility of attaching a prefix to a stem can be constrained by the positional relationship between this prefix and the secondary imperfective suffix *-yva*.

With respect to these two restrictions there emerge SELECTIONALLY RESTRICTED (SR) and POSITIONALLY RESTRICTED (PR) prefixes. If the distribution of a given prefix can be described through (6a), this prefix is said to be an SR-prefix. If the distribution of a given prefix can be described through (6b), this prefix is said to be a PR-prefix. The list of SR-prefixes includes: delimitative *po-*, cumulative *na-*, distributive *pere-*, inchoative *za-*. The list of PR-prefixes includes: completive *do-*, repetitive *pere-*, attenuative *pod-*, attenuative *pri-*. The perdurative *pro-* is not considered separately by Tatevosov (2009; 2013).

---

[3]The property of being "formally (im)perfective" does not imply carrying any aspectual semantics. This is a morphological notion.

There is one more alternative view on the lexical/superlexical distinction found in Žaucer (2009; 2012). The author of these works consistently argues that at least some SLPs merge within the same resultative projection as LPs. The perdurative *pro-* is claimed to be one of these SLPs. It should be pointed out, however, that Žaucer's (2009; 2012) proposal is based mainly on data from Slovenian. The present work does not set as its goal to revise it. What I aim to do is to try to determine the position of the perdurative *pro-* in Russian. For this I will use diagnostics proposed by Tatevosov (2009; 2013).

2.2 The position of the perdurative *pro-*

In this part of the paper, I will follow Tatevosov (2009; 2013) and assume that SLPs fall into at least two separate classes: SR-prefixes and PR-prefixes. The class membership of a given prefix is determined via the restrictions from (6). Below I will show that the distribution of the perdurative *pro-* is subject to selectional restrictions and argue that the perdurative *pro-* is an SR-prefix.

First, the perdurative *pro-* selects for formally imperfective stems. Its distribution falls under the generalization in (7).

(7)   The perdurative *pro-* merges with formally imperfective stems.

It has already been demonstrated in §1 that the perdurative *pro-* can combine both with simplex imperfective stems and imperfective stems derived through applying the secondary imperfective suffix *-yva*. Here, I repeat the relevant examples.

(8)   Perdurative *pro-* with a non-derived (non-prefixed) imperfective stem:
      Lena pro-govori-la   ves' večer   po telefonu.
      L.   PRO_PER-talk-PST all   evening on phone
      'Lena spent the whole evening talking on the phone.'

(9)   Perdurative *pro-* with a stem imperfectivized by *-yva:*
      Kolja pro-[[otkr]-yva]-l   rot   minutu.
      K.   PRO_PER-open-YVA-PST mouth minute
      'Kolja spent a minute opening the mouth.'
      a.   ✗ Scenario №1: Kolja spent a minute opening the mouth wider and
           wider until he got cramps in his cheekbones.
      b.   ✓ Scenario №2: Kolja spent a minute opening the mouth several
           times until he realized that he was not producing any sounds.

What is the source of the restriction on possible interpretations in (9)?

As was shown in §1 (recall (5a)), when the perdurative *pro-* is not attached above *-yva*, the sentence is also compatible with a progressive interpretation. Tatevosov (2015) claims that when *-yva* induces a progressive interpretation, it functions as a partitive operator and merges in a position between VP and *v*. One can naturally assume that the second (pluractional) interpretation of the sentence in (5a) is also induced by the *-yva* suffix and that in this case it functions as a pluractional operator. It has been proposed that pluractional operators apply very low in the syntactic structure, namely, at the level of V (see, e.g, Lasersohn 1995; van Geenhoven 2004). If this view is correct, the restriction in (9) could be treated as (morpho)syntactic. Namely, despite no surface difference, it could be stated that *-yva* has the possibility to enter the derivation in two distinct hierarchical positions: within VP and above it. When the suffix induces a progressive interpretation, it merges, as Tatevosov (2015) proposes, between VP and *v*. In contrast, when this suffix induces a pluractional interpretation, it adjoins to V. The incompatibility of the perdurative *pro-* and the "partitive" *-yva* could be explained by claiming that they compete for the same position and, thus, block the derivation. Such a configuration would look like in (10).

(10) $[_{vP} \ldots [_{v'} \ldots [_F \textit{pro}_{\text{PER}}\text{- -}\textit{yva}_{\text{PART}} \ldots [_{\text{VP}} \ldots]]]]$.

The explanation for the compatibility of the perdurative *pro-* and the "pluractional" *-yva* would be that the latter occupies a position lower in the tree and does not prevent the former from merging with a stem, (11).

(11) $[_{vP} \ldots [_{v'} \ldots [_F \textit{pro}_{\text{PER}}\text{-} \ldots [_{\text{VP}} \ldots [_{V'} \text{-}\textit{yva}_{\text{ITER}} \ldots]]]]]$.

While this line of reasoning may be true, I cannot come up with any empirical evidence in favor of it. [4] Moreover, as Iordachioaia & Soare (2015) show, there are a number of empirical and theoretical challenges for the view that pluractional operators are V-level operators. They argue for high aspect-level pluractionality. I will pursue a different path and try to show in §3 that the restriction we observe in examples like (5b) and (9) occurs due to semantic reasons and can be explained without any specific assumptions about the syntax of pluractional operators.

On the flip side, the perdurative *pro-* does not select for formally perfective stems. Its distribution falls under the generalization in (12).

---

[4]In fact, below I will argue, relying on the fact that the perdurative *pro-* stacks above the repetitive *pere-*, that the perdurative *pro-* merges above *v*P. This, if true, can be taken as an indirect argument against such a reasoning.

(12)   The perdurative *pro-* does not merge with formally perfective stems.

(13)   Perdurative *pro-* with a non-derived (non-prefixed) perfective stem:

> * Vasja pro-reši-l   zadanije desjat' minut.
> V.   PRO_{PER}-solve-PST task   ten   minutes
> Intended: 'Vasja spent ten minutes solving the task.'

(14)   Perdurative *pro-* with a perfective stem derived by prefixation:

> * Maša pro-[na-[pisa]]-la   pis'mo dva časa.
> M.   PRO_{PER}-NA-write-PST letter   two hours
> Intended: 'Maša spent two hours writing the letter.'

As can be seen from the examples, the perdurative *pro-* cannot combine either with simplex perfective stems, (13), or with perfective stems derived by prefixation, (14).

The intermediate conclusion that can be drawn at this stage is the following. The distribution of the perdurative *pro-* is constrained by the (im)perfectivity of the stem. Specifically, the perdurative *pro-* merges with formally imperfective stems and does not merge with formally perfective ones. Therefore, the distribution of the perdurative *pro-* is subject to selectional restrictions and hence the perdurative *pro-* is an SR-prefix. Next, I will consider constraints that regulate the distribution of PR-prefixes and argue that (i) the perdurative *pro-* is not subject to these constraints; (ii) the syntactic position in which the perdurative *pro-* merges is above *v*P.

PR-prefixes do not impose restrictions on the (im)perfectivity of the stem with which they combine (Tatevosov 2009; 2013). As was shown in (13) and (14), this does not hold for the perdurative *pro-* because the perdurative *pro-* requires the stem to which it attaches not to be formally perfective.

The distribution of PR-prefixes falls under the generalization in (15).

(15)   PR-prefixes do not merge above the secondary imperfective suffix *-yva.*

As was shown in (9), repeated here as (16), this generalization also does not hold for the perdurative *pro-*.

(16)   Perdurative *pro-* above the secondary imperfective suffix *-yva:*

> Kolja pro-[[otkr]-yva]-l   rot   minutu.
> K.   PRO_{PER}-open-YVA-PST mouth minute
> 'Kolja spent a minute opening the mouth.'

Another configuration, when it attaches under the secondary imperfective, is also present, (17).

(17)   The perdurative *pro-* under the secondary imperfective suffix *-yva:*

Prošloj zimoj   ja [pro-[lež]]-iva-l   na divane po desjat' časov v den'.
Last   winter I   PRO_PER-lay-YVA-PST on sofa   by ten   hours in day
'Last winter I spent ten hours a day laying on the sofa.'

Unlike PR-prefixes, SR-prefixes do not have restrictions relative to the position of the secondary imperfective *-yva*. The fact that the perdurative *pro-* can merge both under and above the suffix also unites it with SR-prefixes.

    There is one more observation concerning SR- vs. PR-prefixes dichotomy: SR-prefixes attach above PR-prefixes and they cannot merge as adjacent heads, -*yva* must merge between them (Tatevosov 2009; 2013). As the example in (18) demonstrates this is true for the perdurative *pro-*. In this example, the perdurative *pro-* attaches to the stem that already contains the repetitive suffix *pere-*, which is assumed to be a PR-prefix.[5]

(18)   Vasja {pro-[[pere-[čit]]-yva]-l   / *pro-[pere-[čita]]-l   } etot
V.   PRO_PER-PERE_REP-read-YVA-PST   PRO_PER-PERE_REP-read-PST this
abzac   polčasa,   no ničego ne ponja-l.
paragraph half.an.hour but nothing not understand-PST
'Vasja spent half an hour reading this paragraph over and over again, but did not understand anything.'

PR-prefixes have the possibility to enter the derivation in two distinct syntactic positions: between VP and *v*P, and above *v*P (Tatevosov 2008). The position of the prefix affects the interpretation of the predicate. If the prefix merges before *v*P is projected and takes in its scope only the result state, the restitutive reading obtains. In contrast, if the prefix merges after *v*P is projected and takes in its scope the whole event, the repetitive reading obtains. In (18), we observe the second possibility: *pere-* enters the derivation above *v*P, and, thus, yields the repetitive reading 'Vasja read the paragraph, and that had happened before'. Therefore, as the perdurative *pro-* attaches above *pere-*, it must merge in a position above *v*P. We are open to two different possibilities: (a) the perdurative *pro-* merges in AspP and (b) the perdurative *pro-* merges lower, between *v*P and AspP. Although (a) seems to be more straightforward, below I will speak in favor of (b).[6]

---

[5]Note that (18) is compatible only with a pluractional interpretation, as is reflected in the translation.

[6]Note that in principle both of them are consistent with the semantic proposal made in §3.

In formal Slavic literature, there have been various proposals regarding the position in which SLPs merge. They can be divided into two groups with respect to how they treat the hierarchical relationship between prefixes and perfectivity: (i) perfectivity is introduced as high as prefixes; (ii) perfectivity is introduced higher than prefixes.

The first view is the mainstream and presented in a large number of works on the syntax of prefixation. For example, Babko-Malaya (1999) treats SL-prefixes as left-adjuncts to Asp and Svenonius (2004) treats SL-prefixes as PPs occupying SpecAspP (see also Ramchand 2004 for a similar proposal). These authors assume that perfectivity is directly encoded in SL-prefixes or introduced sufficiently local to them.

The second view is less popular and presented in Piñón (1994) and Filip (2000; 2005; 2008). These authors focus on the delimitative prefix *po-* and claim that it is an event modifier with the semantics of a measure adverbial. The result of its application is an event predicate, not a property of times. According to these works, perfectivity is not part of the meaning of the delimitative *po-*. It is introduced by a phonologically silent operator located higher in the syntactic structure.

The crucial thing is that these approaches have different predictions about the possibility of multiple pieces of aspectual morphology within a word form. Specifically, if we assume that the perdurative *pro-* merges in the projection of aspectual operators, we do not expect any other aspectual morphology after its application. Data, however, suggest just the opposite. It has already been shown in (17), repeated here as (19), that the secondary imperfective suffix *-yva* can attach above the perdurative *pro-*.

(19) Prošloj zimoj   ja [pro-[lež]]-iva-l   na divane po desjat' časov v den'.
      Last   winter I   PRO_PER-lay-YVA-PST on sofa   by ten   hours in day
      'Last winter I spent ten hours a day laying on the sofa.'

If we take aspectual morphology to be merged in the projection of aspectual operators, then in the case when there are multiple pieces of this morphology within a word form, as in (19), we have to postulate several adjacent AspP projections. However, there exists a well-established ban on consecutive identical projections. De Clercq & Wyngaerd (2019) have recently proposed the following formulation of this constraint:

(20) * $\langle X, X \rangle$
      The functional sequence must not contain two immediately consecutive identical projections.

That is, the syntactic structure with several adjacent AspP projections would contradict the constraint on admissible functional sequence and would be undesirable from a theoretical point of view. This is the reason why I opt for (b): the perdurative *pro-* merges in a functional projection between *v*P and AspP.[7]

Summarizing all that has been discussed so far, I make the following conclusions:

- The distribution of the perdurative *pro-* is subject to constraints that regulate the distribution of SR-prefixes.

- The syntactic position in which the perdurative *pro-* merges is between *v*P and AspP.

- The perdurative *pro-* felicitously combines with predicates based on secondary imperfective verb stems only when they receive a pluractional interpretation.

In the next section, I will try to argue for a unified semantic constraint that regulates the distribution of the perdurative in Russian. The proposal will be based on the observation about the specific behaviour of pluractional predicates with respect to Landman's (1992) stage-of relation.

3 A semantic constraint on the distribution of the perdurative

3.1 Theoretical background

I adopt a neo-Davidsonian version of event semantics where verbs are represented as one-place predicates over sets of eventualities. I assume that the denotation of *v*P is a predicate of events which is mapped to predicates of times at Asp (see, e.g., Klein 1994 and much further literature). In addition, adhering to the common view in predicate decomposition, I presuppose that while accomplishment predicates are subeventually complex and consist of, at least, two separate

---

[7]A reviewer poses the following question: "Which projection is positioned there? Are there more projections present? In fact, it seems to me that there could be only different types of aspectual projections like in Cinque (1999)". While I have no basis to make more concrete claims regarding the nature of this projection, I assume that an approach in the spirit of Cinque (1999) can be potentially undertaken. For example, Markova (2011) proposes that some prefixes in Bulgarian are derived in dedicated aspectual projections a la Cinque (1999).

subevents—activity (or process subevent) and result state (or become subevent)—that are connected by a finite set of causal relations (Dowty 1979; Rothstein 2004; a.o.), activity predicates have simple structures consisting of a single activity subevent. I assume, following Pazelskaya & Tatevosov (2006), that predicates based on "simplex imperfective" verb stems are associated with an activity event structure. Following Altshuler (2013; 2014) and Grønn (2003; 2015), I take both simplex imperfectives and secondary imperfectives as denoting (not necessarily proper) parts (or stages) of complete eventualities.

3.2 Previous studies

To the best of my knowledge, the semantics and the distribution of the perdurative in Russian have not been the subject of a separate study so far. Traditional Slavic aspectology is rather consistent in giving a more or less uniform treatment of the meaning of the perdurative. For example, Isačenko (1960: 243–244) proposes that perdurative verbs denote "the completion of a process that lasted for a specific period of time". Bondarko & Bulanin (1967: 16) postulate a "long-term Aktionsart"; its nucleus includes "verbs with the prefix *pro-* denoting an action that covers a specific period of time". Forsyth (1970: 23) notes that verbs marked by the perdurative *pro-*, as opposed to verbs marked by the delimitative *po-*, "suggest a longer period of time" during which an action is performed. A "long-term" or "perdurative" Aktionsart is postulated by Zalizniak & Šmelev (2000: 112–113). The authors claim that this Aktionsart includes "verbs denoting an action that took place within a closed specific period of time". In formal literature, we find similar opinions on the semantics of the perdurative. For example, Gehrke (2003: 26) indicates that the perdurative "refers to an unexpectedly long duration of a situation, where it always has to be made explicit that this duration is specific". Filip (2005: 32) notes that the perdurative *pro-* "indicates a relatively long temporal extent of the event (with connotations of wasted time, boredom, and the like)."

Although the inference that the eventuality from the extension of the verbal predicate marked by the perdurative prefix *pro-* lasts long is indeed frequently present, it is, in fact, optional and can be canceled by some linguistic means. Consider, e.g., the sentence in (21), where the perdurative verb is followed by the adverb *vsego liš'* 'just, only'.

(21) V očeredi Kolja pro-stoja-l   vsego liš' sutki.
      in queue   K.   PRO_PER-stay-PST total only day
      'Kolja stood in the queue only for a day.'

Here, the use of the adverb explicitly indicates that an event of staying described by the perdurative verb *pro-stojat'* 'stay' did not last long relative to expectations. The fact that the inference of long duration arises in the presence of the perdurative *pro-* and can be absent in the appropriate context, suggests that this inference is an implicature. The generation of implicatures is generally taken to be due to the non-use of a non-weaker alternative: the choice of one alternative implies the negation of the other. The only candidate for the position of the alternative item for the perdurative *pro-* is the so-called delimitative prefix *po-*, which also conveys the meaning of temporal duration. Indeed, the perdurative and the delimitative are usually considered together and sometimes assumed to impose similar restrictions on the predicates with which they combine (see, e.g., the references above). Constraints on the distribution of the delimitative have been recently explored by Tatevosov (2017). In the next subsection, I will give an outline of the theory developed in these works and take it as a starting point in identifying constraints on the distribution of the perdurative.

3.3 The delimitative and unique temporal arrangement

Tatevosov (2017) claims that in Russian the derivation of non-culminating accomplishments—predicates that appear in perfective clauses and describe some goal-oriented activity that stops before the corresponding result state is attained—proceeds in two steps: the secondary imperfective suffix *-yva* first merges with a formally perfective stem and then the delimitative prefix *po-* attaches to the resulting complex.[8] Each step is subject to certain semantic restrictions. The ones that we are interested in here are those that emerge when the delimitative comes into play. Tatevosov (2017) proposes that the delimitative can be derived from an event predicate iff the activity component of this predicate does not contain in its extension subevents that are temporally ordered in a unique way. The notion of unique temporal orderedness is formalized through the property of UNIQUE TEMPORAL ARRANGEMENT (UTA). An informal definition of UTA is given in (22).

(22) *Definition:* Unique temporal arrangement

   Whenever an event *e* falls under *P*, there is exactly one way for *e* to start,

---

[8]Tatevosov (2017) proposes that non-culminating accomplishments exist in Russian, and the delimitative *po-* is a means to derive them. There are, however, alternative views on this problem. For example, Martin (2017) claims that Russian does not possess non-culminating accomplishments, and the delimitative-prefixed verbs do not exemplify this class of predicates. I believe that this theoretical debate is orthogonal to the current purposes.

> there is exactly one way for *e* to finish, and for any non-final part of *e* there
> is exactly one follow up.

This property is best understood by looking at concrete examples. Let us compare predicates *zapivat' tabletku* 'wash down a pill' and *zapolnjat' anketu* 'fill in a form', which are both based on secondary imperfective verb stems and within our assumption denote non-final parts of complete eventualities. The predicate *zapivat' tabletku* 'wash down a pill' does not allow the formation of the delimitative, while *zapolnjat' anketu* 'fill in a form' does. These predicates differ with respect to the UTA property. The activity component of the predicate *zapivat' tabletku* 'wash down a pill' consists of very specific subevents that have to come in a very specific order to represent an activity of washing down a pill. In contrast, the predicate *zapolnjat' anketu* 'fill in a form' denotes events in which subevents making up the activity of filling in a form do not have to come in any specific order. Even if some of them are skipped or occur more than once, their sum still represents an activity of filling in a form. According to Tatevosov (2017), the delimitative *po-* is sensitive to this difference. Crucially, the delimitative *po-* can also combine with predicates based on simplex imperfective verb stems, like *govorit' po telefonu* 'talk on a phone', which are associated with an acitivity event structure and also do not possess UTA. Putting these facts together, Tatevosov (2017) concludes that the complement of the delimitative, whether it is derived or non-derived, must be an Activity.

The important fact for the purposes of this work is that, despite superficial similarity, the perdurative and the delimitative impose different requirements on predicates with which they combine. As was shown in §1, the perdurative, unlike the delimitative, can be derived from some activity predicates but cannot be derived from others. Consider the contrast between (23) and (24).[9]

(23) a. * Kolja pro-pisa-l   pismo neskol'ko minut.
          K.   PRO_PER-write-PST letter several   minutes
          Intended: 'Kolja spent several minutes writing the letter.'

      b. Kolja pro-taska-l   čemodan neskol'ko časov.
          K.   PRO_PER-carry-PST suitcase several   hours
          'Kolja spent several hours carrying the suitcase.'

---

[9]It must be pointed out that the verb *taskat'* in (23b) and (24b) is an "indeterminate" motion verb. Indeterminate motion verbs are used to describe undirected motion, motion back and forth, motion that is not associated with any particular path being covered during the unfolding of an event. Neither the delimitative nor the perdurative can be derived from a predicate when it describes motion in a single direction.

(24) a. Kolja po-pisa-l   pismo neskol'ko minut.
          K.   PO$_{DEL}$-write-PST letter several   minutes
          'Kolja spent several minutes writing the letter.'

      b. Kolja po-taska-l   čemodan neskol'ko časov.
          K.   PO$_{DEL}$-carry-PST suitcase several   hours
          'Kolja spent several hours carrying the suitcase.'

Since the boundary separating predicates that allow the perdurative from predicates that do not allow it runs inside the class of activities, more should be said about their internal structure. A more specific question that I would like to address is: are there any properties that differentiate predicates like *pisat' pismo* 'write a letter' from predicates like *taskat' čemodan* 'carry a suitcase'? The next subsection is devoted to seeking an answer to this question.

3.4 Theories of atelicity

In the literature on aspectual composition, it is consistently argued that activity predicates are atelic. What lies under this notion varies across different approaches. In their influential work, carried out within the framework of temporal semantics, Bennett & Partee (1978) propose that atelic predicates must satisfy the SUBINTERVAL PROPERTY.

(25) *Definition:* Subinterval property
      $\text{SUB}(P) \leftrightarrow \forall i[\text{AT}(P, i) \rightarrow \forall j[j \subset i \rightarrow \text{AT}(P, j)]]$
      P possesses the subinterval property iff, if *P* is true at *i*, it is true at every subinterval of *i*.

This definition, however, is too strong and gives rise to the well-known "minimal-parts problem" (Dowty 1979). The problem is that it is not the case that all activity predicates are true at every subinterval during which the events from their extension take place. Dynamic eventualities never hold at points, they take time to establish themselves. Therefore, it is reasonable to judge whether a given predicate possesses the subinterval property only relative to an interval that is "sufficiently large" for the event from its extension to unfold.

Now that we have equipped ourselves with the notion of subinterval property, we can return to the question formulated at the end of the previous subsection and consider the predicates *pisat' pismo* 'write a letter' and *taskat' čemodan* 'carry a suitcase' in detail. These predicates are based on underived imperfective stems and, within the assumption introduced in §3.1, denote non-final parts of complete events and belong to the class of activities. The subinterval property is not

what distinguishes them: both *pisat' pismo* 'write a letter' and *taskat' čemodan* 'carry a suitcase' possess this property. If a nonfinal part of an event of writing a letter holds at an interval i, it also holds at every sufficiently large subinterval of i. Similarly, if a non-final part of an event of carrying a suitcase holds at an interval i, it also holds at every sufficiently large subinterval of i. The perdurative, however, can be derived only from the latter but not from the former.

In the mereological approach to aspectual composition (Krifka 1989; 1992; 1998), atelic predicates are defined as being cumulative. A predicate is CUMU-LATIVE if it satisfies the properties of ADDITIVITY, (26), and DIVISIVITY, (27).

(26)   *Definition:* Additivity

$\forall P[\text{CUM}(P) \leftrightarrow \forall x \forall y [P(x) \wedge P(y) \rightarrow P(x \oplus y)]]$

P is additive iff whenever it applies to the entities x and y, it also applies to the sum $x \oplus y$

(27)   *Definition:* Divisivity

$\forall P[\text{DIV}(P) \leftrightarrow \forall x \forall x' [P(x) \wedge x' \subset x \rightarrow P(x')]]$

P is divisive iff whenever P applies to x, then it must also apply to any x' that is properly included in x.

The notion of cumulativity also cannot help us to distinguish *pisat' pismo* 'write a letter' from *taskat' čemodan* 'carry a suitcase'. Both these predicates are additive and divisive down to minimal parts.[10]

Landman & Rothstein (2009; 2012) claim that neither the subinterval property nor the property of cumulativity is adequate for distinguishing activities from other classes of predicates. They argue that activity predicates are lexically constrained as being incrementally homogeneous.

INCREMENTAL HOMOGENEITY is based on two essential components: CROSS-TEMPORAL IDENTITY and EVENT ONSETS. Cross-temporal identity is a semantic primitive that is used to compare events with different running times. Events with different running times count as the same event if they are in an equivalence relation of cross-temporal identity, (28).

(28)   *Definition:* Equivalence relation of cross-temporal identity

e_1 is cross-temporally identical to e_2, $e_1 \sim e_2$ iff e_1 and e_2 count as one and the same event, i.e. for counting purposes e_1 and e_2 count as one event.

---

[10]Due to space limitations, I leave the verification of this statement to the reader.

The second component is the notion of onset of an event. For e, the onset of e is the smallest initial part of e that is big enough to count both as e and as cross-temporally identical to e, (29).

(29)   *Definition:* Onset of an event
       Let e be an eventuality of verb type V.
       The onset of e, relative to V, $O(e, V)$ is the smallest eventuality of type V such that: $O(e, V) \sim e$ and $\tau(O(e, V) \subseteq_{\text{IN}} e$.

Taking these, the notion of incremental homogeneity is defined, (30).

(30)   *Definition:* Incremental homogeneity
       Let α be a VP with event type α and verbal event type V_α.
       Let $e \in V_\alpha$ and $e \in \alpha$.
       e is incrementally homogeneous w.r.t. α and V_α iff for every interval i: if $\tau(O(e, V)) \subseteq_{\text{IN}} i \subseteq_{\text{IN}} \tau(e)$ then there is an eventuality e' of event type α such that: $e' \sim e$ and $\tau(e') = i$.

For an event $e \in \alpha$ to be incrementally homogeneous, the onset of e must count both as event type V_α and as event type α. If α does not hold at the onset of e, neither e is incrementally homogeneous nor the predicate that is true of e. This is exactly what happens to predicates like *pisat' pismo* 'write a letter'. Let the predicate *pisat' pismo* 'write a letter' be an event type α, with verbal event type V_α, which is *pisat'* 'write'.

(31)   a. $\alpha = \lambda e.\textsc{write}(e) \wedge \textsc{theme}(\textsc{letter})(e)$
            Event type of writing letter events
       b. $V_\alpha = \lambda e.\textsc{write}(e)$
            Event type of writing events

Suppose that e is a non-final part of an event of writing a letter. According to the definition, the onset of e is its most initial proper part that counts as a writing activity. It is not the case, however, that this onset must count as a non-final part of an event of writing a letter. An event of someone diligently tracing out the first letter of the letter undoubtedly counts as writing but very unlikely counts as writing a letter. Thus, the predicate *pisat' pismo* 'write a letter' is not incrementally homogeneous, and the perdurative cannot be derived from it.

In contrast, *taskat' čemodan* 'carry a suitcase' is an incrementally homogeneous predicate. Suppose that e is a non-final part of an event of carrying a suitcase. According to the definition, the onset of e is its most initial proper part that

counts as a carrying activity. What is crucial is that this onset counts also as a non-final part of an event of carrying a suitcase. An event of someone taking a suitcase and making a few steps in different directions counts both as an event of carrying and as an event of carrying a suitcase. Thus, the predicate *taskat' čemodan* 'carry a suitcase' is incrementally homogeneous, and the perdurative can be derived from it.

Can we stop here and say that the restriction on incremental homogeneity is what regulates the distribution of the perdurative? The answer is negative. There are incrementally homogeneous predicates that are marginal with the perdurative. These are the predicates in which, as Tatevosov & Ivanov (2009: 106) claim, "the activity that up to its final point does not contribute to the development of the become subevent at all". An example of such a predicate is *otkryvat' zamok* 'open a lock'. This is an atelic predicate obtained from its telic counterpart *otkryt' zamok* 'open a lock' through applying the secondary imperfective suffix *-yva*. Within our assumption, this predicate denotes non-final parts of a complete event of opening a lock. These non-final parts do not induce any change of state. The result state is brought about by the very final subevent. This subevent, however, is not in the denotation of the secondary imperfective *otkryvat' zamok* 'open a lock'.

Consider the example in (32). In the provided situation, the sentence is judged as ungrammatical.

(32)   *Situation:* The lock in the door is broken. After spending half an hour trying to open it, Vasja gave up.

    * Vasja pro-[[otkr]-yva]-l   zamok polčasa.
    V.   PRO$_{PER}$-open-YVA-PST lock   half.an.hour
    Intended: 'Vasja spent half an hour opening the lock.'

It is easy to show that this predicate is indeed incrementally homogeneous. Let *e* be a non-final part of a complete event of opening a lock. According to the definition of onset given above, the onset of *e* is its most initial proper part that counts as an opening activity. Crucially, this onset counts as a non-final part of an event of opening a lock, too. The very first manipulation with a lock aimed at unlocking it counts both as opening and as opening a lock. Thus, the predicate *otkryvat' zamok* 'open a lock' is incrementally homogeneous. However, the perdurative is blocked.

In the next subsection, I will try to identify the property that distinguishes predicates that allow the perdurative from predicates that do not and try to for-

mulate a unified semantic constraint that regulates the distribution of the perdurative *pro-* in Russian.

3.5 The perdurative and arrangement by stages

I start this subsection from the observation, previously presented in §1, that the perdurative can be derived from predicates based on the secondary imperfective verb stems only if they receive a pluractional interpretation, (33).

(33) Kolja pro-[[otkr]-yva]-l   rot   minutu.
      K.   PRO_PER-open-YVA-PST mouth minute
      'Kolja spent a minute opening the mouth.'

    a. ✗ Scenario №1: Kolja spent a minute opening the mouth wider and wider until he got cramps in his cheekbones.

    b. ✓ Scenario №2: Kolja spent a minute opening the mouth several times until he realized that he was not producing any sounds.

The above contrast suggests that the distinguishing property of pluractional predicates might serve as a clue to our understanding what the constraint on the distribution of the perdurative is. There have been proposed various analyses of event pluractionality in formal semantic literature (Lasersohn 1995; van Geenhoven 2004; Tovena 2010; Henderson 2012, a.o.). Space prevents me from reviewing them in detail here. To put it briefly, the basic idea shared by these works is that a pluractional marker takes an underlying predicate and ensures that there is a multiplicity of atomic events of the same type. Importantly, these atomic events are self-sufficient, they do not stand in any special temporal (like UTA) or causal relation to each other. I capture this property by utilizing Landman's (1992) stage-of relation.

    The notion of a STAGE has been proposed by Landman (1992) to define the semantics of the progressive in English. He claims that the progressive is a function from a set of events denoted by VP to a set of stages of those events. A progressive sentence is true if a VP-event stage develops into an event of the same kind denoted by VP. Stage-of relation is a partial ordering of the set of events. For an event *e* to be a stage of another event *e′*, *e* must share enough characteristics with *e′* and must develop into *e′* in some possible world that is near enough to the world of evaluation.

    Pluractional predicates are organized in a special way with respect to the stage-of relation: although every atomic event from the extension of the pluractional predicate is a stage of the plural event denoted by this predicate, no atomic event

is a stage of any other atomic event. I claim that other predicates allowing the derivation of the perdurative are organized exactly as pluractional predicates with respect to the stage-of relation. I propose that for a predicate to be able to form the perdurative, subevents making up the activity component of this predicate must not be arranged by the stage-of relation. Below I will argue that if a predicate blocks the formation of the perdurative, it possesses the property of ARRANGEMENT BY STAGES (AbS). A formal definition of AbS is given in (34).

(34) Let e be a partial eventuality from the extension of an event predicate P, and

Let e' and e'' be stages of e such that

  a. $e' \subseteq e$, and

  b. $e'' \subseteq e$, and

     P(e) possesses AbS iff

     $\forall e' \forall e'' \, [\, e' \ll_{\scriptscriptstyle T} e'' \rightarrow e'$ is a stage of $e''\,]$

     where $\ll_{\scriptscriptstyle T}$ is a temporal precedence relation on events.

(34) says that an event predicate P is an AbS predicate iff for all contextually salient subevents e' and e'' in its denotation such that both e' and e'' are stages of e, if e' temporally precedes e'', then e' is a stage of e''. AbS is stronger than incremental homogeneity. I claim that this property is what distinguishes predicates that allow the perdurative from predicates that do not.

Let us consider the predicate *otkryvat' zamok* 'open a lock' again. Imagine a scenario in which one has a bunch of numbered keys but does not know which one opens the lock. Suppose that a stage of a partial eventuality e from the extension of the predicate *otkryvat' zamok* 'open a lock' is an event of using key number one. Let it be e'. Suppose that another stage of e is an event of using key number two. Let it be e''. In addition, assume that e' temporally precedes e''. Clearly, both e' and e'' are stages of the bigger event e. The crucial fact is that e' is a stage of e'', too. According to the original definition proposed in Landman (1992), "an event is a stage of another event if the second can be regarded as a more developed version of the first, that is, if we can point at it and say, "It's the same event in a further stage of development." (Landman 1992: 23). Although Landman (1992) does not provide an explanation for what it is to be "a more developed version", intuitively, a more developed event is nothing more than a next step a rational agent takes to achieve the desired goal. Returning to the discussed scenario, if the agent loses the hope of opening the lock with the first two keys, it is very unlikely that she uses key number one again. In case she picks it

and performs the same kind of activity as before (without changing the turning direction of the key, the applied force, etc.), this can hardly count as a more developed version of the previous event, which is a sum of using key number one and using key number two, because this activity does not bring the agent closer to her goal—have the lock opened. For an event to fall under the denotation of *otkryvat' zamok* 'open a lock', this event has to consist of subevents such that each subsequent subevent is a more developed version of the previous one. That is, *otkryvat' zamok* 'open a lock' is an AbS predicate. This is the reason why the perdurative is not allowed from it.

As was shown in the previous sections, predicates like *govorit' po telefonu* 'talk on the phone' and *taskat' čemodan* 'carry a suitcase' allow the formation of the perdurative. My proposal correctly accounts for this fact. Let us consider the predicate *taskat' čemodan* 'carry a suitcase'. Indeterminate verbs of motion in Russian can express a wide range of meanings (Forsyth 1970). In episodic contexts, the perdurative is licensed when a predicate has the so-called multiple directions reading. This reading arises in a scenario with a single event of motion "in various unspecified directions" (Forsyth 1970: 321). It is easy to show that when *taskat' čemodan* 'carry a suitcase' receives the multiple direction reading, it does not possess AbS. Suppose that a stage of a partial eventuality e from the extension of the predicate *taskat' čemodan* 'carry a suitcase' is an event of moving in one particular direction. Let it be e'. Suppose that another stage of e is an event of moving in the other direction. Let it be e''. Assume that e' temporally precedes e''. Both e' and e'' are stages of the bigger event e: an event of moving in multiple directions is a more developed version of an event of moving in a single direction. However, an event of moving in one direction can hardly count as a more developed version of an event of moving in the other direction. Moreover, after the agent tried to go in two different directions, she can return to the starting point and take the first direction again. This action will still be in the extension of the predicate *taskat' čemodan* 'carry a suitcase'. For an event to fall under its denotation, it does not need to consist of subevents such that every subsequent subevent is a more developed version of the previous one. The same is true for the predicate *govorit' po telefonu* 'talk on a phone'.

The perdurative also cannot be derived from predicates associated with an incremental relation between the activity and change of state components. Predicates of this type denote events such that for every activity subevent there is a change of state that it induces (Rothstein 2004). An instance of such a predicate is *pisat' pismo* 'write a letter'. When one writes a letter, every part of the writing activity corresponds to some part of the process of being written. The perdurative cannot be derived from this predicate, (35).

(35) * Kolja pro-pisa-l   pismo vsjo utro.
      K.   PRO~PER~-write-PST letter all   morning
      Intended: 'Kolja spent the whole morning writing the letter.'

My proposal predicts it because this predicate possesses AbS: an activity of writing a letter consists of subevents such that every subsequent subevent is a more developed version of the previous one. Let me show why. Suppose that a stage of a partial eventuality *e* from the extension of the predicate *pisat' pismo* 'write a letter' is an event of writing an address. Let it be *e′*. Suppose that another stage of *e* is an event of writing a salutation. Let it be *e″*. Assume that *e′* temporally precedes *e″*. Since each of these subevents makes a contribution to the state of being written, at the moment when the agent writes the salutation the letter is already finished to the extent of the address. The subevent of writing the salutation extends the degree to which the letter is finished. Therefore, it counts as a more developed version of the subevent of writing the address. After the salutation is done, the agent can return to the address-part of the letter and, e.g., correct the name of the street or re-write it anew if she realizes that she has made a mistake, and this action will count as a more developed subevent and will be in the extension of the predicate *pisat' pismo* 'write a letter'. What she is very unlikely to do is to perform the same kind of activity and write the address again without any changes. The action of writing the address again will not count as a more developed subevent and will not be in the extension of the predicate *pisat' pismo* 'write a letter'.

4 Conclusion

In this study, I have analyzed (morpho)syntactic and semantic constraints on the derivation of the perdurative in Russian. Considering the restrictions that the perdurative prefix *pro-* demonstrates when combined with a number of other affixes, I have argued that this prefix is a selectionally restricted prefix that merges in the functional domain between *v*P and AspP. I have shown that the perdurative *pro-* selects for grammatically imperfective stems and combines with predicates based on secondary imperfective verb stems only when they receive a pluractional interpretation. I have demonstrated that the distribution of the perdurative is also constrained semantically. Having started from the observation that the perdurative can be derived from predicates associated with an activity event structure and having discussed the existing proposals for what it is to be an activity predicate, I concluded that none of them is able to account for its distribution. I have

argued that predicates allowing the formation of the perdurative do not possess the AbS property. In other words, activity subevents from their extension do not stand in Landman's (1992) stage-of relation to each other.

Several questions remain. How does the analysis developed here account for stative predicates like *žit' v Moskve* 'live in Moscow', which allow the perdurative? In principle, the proposed analysis predicts the availability of the perdurative from statives because statives do not have stages (Landman 1992), and hence there is simply nothing that can be arranged by the stage-of relation. This idea, if correct, needs to be spelled out in more detail. If the perdurative is not responsible for introducing perfectivity, what is its semantic contribution? A potential line to follow, as I see it, is to make use of the notion of maximality, recently much discussed in the literature (see Altshuler 2014; Filip 2017, a.o.). The perdurative can be treated as an event modifier extracting a maximal stage (or part) of a partial eventuality. The elaboration of this hypothesis is left for future research.

Abbreviations

AbS	arrangement by stages	PST	past
DEL	delimitative	REP	repetitive
LP	lexical prefixes	SLP	super lexical prefixes
PER	perdurative	SR	selectionally restricted
PR	positionally restricted	UTA	unique temporal arrangement

Acknowledgments

References

Altshuler, Daniel. 2013. There is no neutral aspect. In Todd Snider (ed.), *SALT 23: Proceedings of the 23rd Semantics and Linguistic Theory Conference*, 40–62. Washington, DC: Linguistic Society of America. DOI:10.3765/salt

Altshuler, Daniel. 2014. A typology of partitive aspectual operators. *Natural Language & Linguistic Theory* 32(3). 735–775. DOI:10.1007/s11049-014-9232-1

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

Bennett, Michael & Barbara Hall Partee. 1978. *Toward the logic of tense and aspect in English*. Bloomington, IN: Indiana University Linguistics Club.

Bondarko, Alexander & Lev Bulanin. 1967. *Russkij glagol [Russian verb]*. Moscow: Prosveščenie.

Borik, Olga. 2006. *Aspect and reference time*. Oxford: Oxford University Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.

De Clercq, Karen & Guido Vanden Wyngaerd. 2019. Negation and the functional sequence. *Natural Language & Linguistic Theory* 37(2). 1–36. DOI:10.1007/s11049-018-9415-2

Dowty, David R. 1979. *Word meaning and Montague grammar*. Dordrecht: Reidel.

Filip, Hana. 2000. The quantization puzzle. In Carol Tenny & James Pustejovsky (eds.), *Events as grammatical objects, from the combined perspectives of lexical semantics, logical semantics and syntax*, 39–96. Stanford, CA: CSLI Publications.

Filip, Hana. 2005. Measures and indefinites. In Greg N. Carlson & Francis J. Pelletier (eds.), *Reference and quantification: The Partee effect*, 229–288. Stanford, CA: CSLI Publications.

Filip, Hana. 2008. Events and maximalization: The case of telicity and perfectivity. In Susan Rothstein (ed.), *Theoretical and crosslinguistic approaches to the semantics of aspect*, 217–256. Amsterdam: John Benjamins.

Filip, Hana. 2017. The semantics of perfectivity. *Italian Journal of Linguistics* 29(1). 167–200. DOI:10.26346/1120-2726-107

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.

Forsyth, James. 1970. *A grammar of aspect: Usage and meaning in the Russian verb*. Cambridge: Cambridge University Press.

Fowler, George & Michael Yadroff. 1993. The argument status of accusative measure nominals in Russian. *Journal of Slavic Linguistics* 1(2). 251–279. https://www.jstor.org/stable/24598971.

Gehrke, Berit. 2003. *Aspectual affixes in Russian and Czech*. Manuscript, Utrecht University.

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

Grønn, Atle. 2003. *The semantics and pragmatics of the Russian factual imperfective*. Oslo: University of Oslo. (Doctoral dissertation).

Grønn, Atle. 2015. On (in)definite tense and aspect in Russian. In Zybatow Gerhild, Petr Biskup, Marcel Guhl, Claudia Hurtig, Olav Mueller-Reichau & Maria Yastrebova (eds.), *Slavic grammar from a formal perspective: The 10th Anniversary FDSL Conference*, 175–196. Frankfurt am Main: Peter Lang.

Henderson, Robert Martin. 2012. *Ways of pluralizing events*. Santa Cruz, CA: University of California. (Doctoral dissertation).

Iordachioaia, Gianina & Elena Soare. 2015. Pluractionality with lexically cumulative verbs: The supine nominalization in Romanian. *Natural Language Semantics* 23(4). 307–352. DOI:10.1007/s11050-015-9117-9

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

Klein, Wolfgang. 1994. *Time in language*. London & New York: Routledge.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem & Peter von Emde Boas (eds.), *Semantics and contextual expression*, 75–115. Dordrecht: Foris Publications.

Krifka, Manfred. 1992. Thematic relations as links between nominal reference and temporal constitution. In Ivan Sag & Anna Szabolcsi (eds.), *Lexical matters*, 29–53. Stanford, CA: CSLI Publications.

Krifka, Manfred. 1998. The origins of telicity. In Susan Rothstein (ed.), *Events and grammar*, 197–235. Dordrecht: Kluwer.

Landman, Fred. 1992. The progressive. *Natural Language Semantics* 1(1). 1–32. DOI:10.1007/BF02342615

Landman, Fred & Susan Rothstein. 2009. Incremental homogeneity in the semantics of aspectual for-phrases. In Malka Rappaport Hovav, Ivi Sichel & Edit Doron (eds.), *Syntax, lexical semantics and event structure*, 229–251. Amsterdam: John Benjamins.

Landman, Fred & Susan Rothstein. 2012. The felicity of aspectual for-phrases–part 2: Incremental homogeneity. *Language and Linguistics Compass* 6(2). 97–112. DOI:10.1002/lnc3.323

Lasersohn, Peter. 1995. *Plurality, conjunction, and events*. Dordrecht: Kluwer.

Markova, Angelina. 2011. On the nature of Bulgarian prefixes: Ordering and modification in multiple prefixation. *Word Structure* 4(2). 244–271.

Martin, Fabienne. 2017. *Non-culminating accomplishments*. Manuscript, University of Stuttgart.

Pazelskaya, Anna & Sergei Tatevosov. 2006. Uninflected VPs, deverbal nouns and the aspectual architecture of Russian. In James E. Lavine, Steven Franks, Mila Tasseva-Kurktchieva & Hana Filip (eds.), *Formal Approaches to Slavic Linguistics 14: The Princeton Meeting 2005*, 258–276. Ann Arbor, MI: Michigan Slavic Publications.

Piñón, Christopher. 1994. Aspectual composition and the 'pofective' in Polish. In Sergey Avrutin, Steven Franks & Liljana Progovac (eds.), *Slavic linguistics workshop*, 341–373. Ann Arbor, MI: Michigan Slavic Publications.

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

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

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

Romanova, Eugenia. 2007. *Constructing perfectivity in Russian*. Tromsø: Universitz of Tromsø. (Doctoral dissertation).

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

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

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

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

Tatevosov, Sergei. 2009. Mnozestvennaja prefiksacija i anatomija russkogo glagola [Multiple prefixation and the anatomy of Russian verb]. In Ksenia Kisileva, Vladimir Plungian, Ekaterina Rakhilina & Sergei Tatevosov (eds.), *Korpusnye issledovanija po russkoj grammatike*, 92–156. Moscow: Probel-2000.

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

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

Tatevosov, Sergei. 2017. *On the temporal structure of non-culminating accomplishments*. Manuscript, Moscow State University.

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*, 83–129. Amsterdam: John Benjamins.

Tolskaya, Inna. 2015. Verbal prefixes in Russian: Conceptual structure versus syntax. *Journal of Linguistics* 51(1). 213–243. DOI:10.1017/S0022226714000206

Tovena, Lucia. 2010. Pluractionality and the unity of the event. In Maria Aloni & Karin Schulz (eds.), *Proceedings of the Amsterdam Colloquium 2009*, 465–473. Heidelberg: Springer.

van Geenhoven, Veerle. 2004. For-adverbials, frequentative aspect, and pluractionality. *Natural Language Semantics* 12(2). 135–190. DOI:10.1023/B:NALS.0000031389.36427.af

Zalizniak, Anna A. & Alexei Šmelev. 2000. *Vvedenije v russkuju aspektologiju [Introduction to the Russian aspectology]*. Moscow: Studia Philologica.

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

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. Procedings of FDSL 8*, 339–355. Frankfurt am Main: Peter Lang.

Chapter 10

Complementizer doubling in Slovenian subordinate clauses

Vesna Plesničar

University of Nova Gorica

The focus of the present paper is on complementizer doubling constructions in subordinate clauses in Slovenian. The main goal is to show that complementizer doubling in Slovenian is a syntactic phenomenon comparable to complementizer doubling in other, mainly dialectal variants of Romance languages (e.g. Paoli 2003; Ledgeway 2005; Dagnac 2012; Villa-Garcia 2012; González i Planas 2014; Munaro 2016). The Slovenian complementizer doubling data strongly suggests that the syntactic analysis of such constructions is possible only under the assumption that the complementizer field is split into several functional projections, as was first proposed by Rizzi (1997). Since it seems that the doubling complementizer in Slovenian is always the closing element of the complementizer system, it is reasonable to assume that at least in Slovenian, this element occupies the head of finiteness projection, while the first complementizer in complementizer doubling constructions, which functions as the complement clause introducer, sits in the head of the highest projection of the split CP field, i.e. the force projection. The suitability of force projection as the host of the first complementizer in Slovenian complementizer doubling constructions is justified by the fact that topicalized and focalized phrases necessarily follow it, which is the exact same pattern that was observed also for complementizer doubling constructions in Romance languages (e.g. Ledgeway 2005; Dagnac 2012; Munaro 2016, among others).

Keywords: subordinate complementizer, complementizer phrase, subordinate clause, complementizer doubling, split CP hypothesis

1 Introduction

Complementizer doubling is a phenomenon in which a clause contains two complementizers, of which the first is the top-most element of the subordinate clause,

while the second, doubling complementizer is positioned after the element that occurs right after the first complementizer. The phenomenon was first observed in mainly dialectal variants of Romance languages such as Italian (Paoli 2003; Ledgeway 2005; Munaro 2016), French (Dagnac 2012), Spanish (Villa-Garcia 2012; González i Planas 2014) and Portuguese (Mascarenhas 2007).

Examples of Slovenian complementizer doubling are shown in (1b) and (2).[1, 2]

(1) a. Mislim, da   ker   pošteno dela, mu pripada plačilo.
         think   that because honest   work he belong   payment
         'I think that because he works hard, he should get paid.'

     b. Mislim, da   ker   pošteno dela, **da** mu pripada plačilo.
         think   that because honest   work that he belong   payment
         'I think that because he works hard, he should get paid.'

(2) Rekel je, da Petrovim prijateljem da   ne zaupa.
     said aux that Peter's   friends   that not trust
     'He said that he doesn't trust Peter's friends.'

As is evident from the examples in (1b) and (2), the complementizer can only be doubled if there is some phrase that splits the two complementizers in the left periphery of the embedded clause.[3] If there is no such element, complemen-

---

[1]A subset of our data were collected in a controlled acceptability-judgment task; results of that task are reported in Plesničar (2016). The judgments in Plesničar (2016) and the judgments reported here are by Slovenian speakers from Goriška region. Regional variation with respect to this type of sentences is possible.

[2]Phrases in examples typeset in small capitals are focused. The use of commas in this paper corresponds to Slovenian writing conventions and is not intended to reflect either the syntactic status of constituents that are located between the two complementizers or prosody.

[3]Lenertová (2001) and Veselovská (2008) argue that Czech second position clitics occupy the lowest head in the CP field, i.e. Fin^0. If Slovenian clitics occupied the same position we could say that the second complementizer in Slovenian is also in Fin^0, especially in view of the contrast between (1b) and (i), as the clitic cannot follow the first complementizer.

   (i) * Mislim, da   mu, ker   pošteno dela, da   pripada plačilo.
          think   that he because honest   work that belong payment

But as pointed out in Marušič (2008) and as is evident by comparing (1a) and (ii), Slovenian clitics do not occupy a unique syntactic position, so they cannot be located in Fin^0.

   (ii) Mislim, da   mu, ker   pošteno dela, pripada plačilo.
         think   that he because honest   work belong payement
         'I think that because he works hard, he should get paid.'

Due to the different status of Slovenian and Czech clitics, an argumentation built on the clitic status will most likely not be on the right track.

tizer doubling cannot occur. The intervening element in complementizer doubling constructions must be a constituent which was not base-generated in its left-peripheral position between the two complementizers, but is rather located there as a result of movement (the arguments for this claim will be presented in §2 below, where the properties of complementizer doubling phenomena in Slovenian are discussed in more detail). Possible landing sites for such moved constituents are specifiers of the projections that are positioned between the first and the second complementizer, arguably the specifiers of Topic or Focus phrases. The moved constituent can be an adverbial phrase or adverbial clause, as shown in (1b), but it can also be a PP- or an NP-argument of the embedded clause, as is the case in (2).

Complementizer doubling can thus be understood as an inherent property of complement clauses that can be realized only if some additional element or some additional structure is inserted into the specifier position of one of the available projections in the CP field, i.e., a projection between force and finiteness phrase. The focus of the remaining part of the paper will be on examples such as (1b) and (2), in which a second complementizer is present in the sentence structure.

Before proceeding with a detailed description and analysis, two clarifications are in order. In complementizer doubling constructions, both complementizers – the first one, which introduces the complement clause, and the second, doubling one – occur in the complement clause. At this point, there is no clear picture about what the full structure of the Slovenian CP-domain looks like. In this paper, we are going to adopt one (of course not exhaustive) relevant proposal for the structure of the Slovenian CP domain – that of Mišmaš (2015) – which is based on her investigation of Slovenian multiple wh-fronting. A graphical representation of the relevant structure is presented in Figure 1 below.[4]

Secondly, as becomes clear from (1a) and (1b) above, Slovenian complementizer doubling is optional, and the only obligatory complementizer in potential complementizer doubling constructions is the highest one, as in (1a) above. The optionality of the doubling complementizer is characteristic of complementizer doubling phenomena in general (e.g. Ledgeway 2005; Dagnac 2012; Munaro 2016). At this point we will not try to answer the question about the motivation for the appearance of the doubling complementizer in the syntactic structure. We will, however, assume – in accordance with the cartographic analyses of the left periphery (e.g. Rizzi 2004; though contrary to Mišmaš 2015) – that the position for the doubling complementizer is present in the sentence structure regardless of whether the doubling complementizer is phonetically realized or not.

---

[4] According to Mišmaš (2015), the starred projections are in the CP field only when needed, and their positions are interchangeable.

ForceP
├── SpecForceP
└── Force′
    ├── Force⁰
    └── InterP
        ├── SpecInterP
        └── Inter′
            ├── Inter⁰
            └── TopicP*
                ├── SpecTopicP
                └── Topic′
                    ├── Topic⁰
                    └── FocusP*
                        ├── SpecFocusP
                        └── Focus′
                            ├── Focus⁰
                            └── WhP*
                                ├── SpecWhP
                                └── Wh′
                                    ├── Wh⁰
                                    └── FinP
                                        ├── SpecFinP
                                        └── Fin′
                                            ├── Fin⁰
                                            └── IP
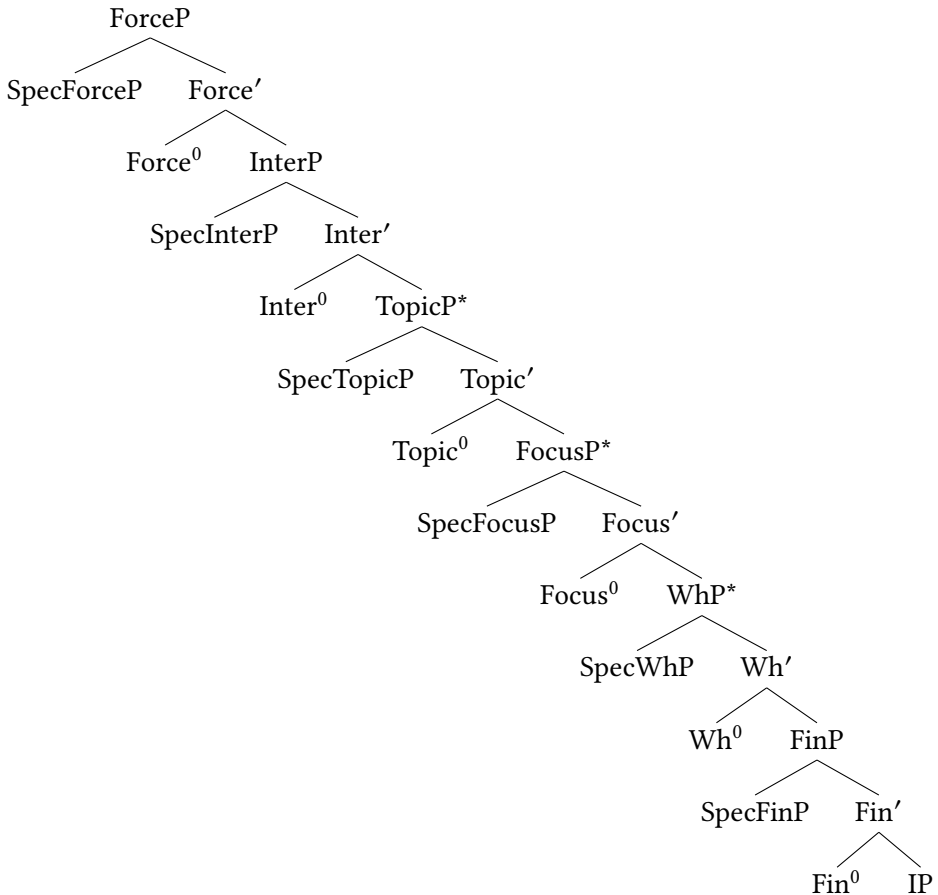
Figure 1: The structure of the CP field in Slovenian (Mišmaš 2015)

In the next section, we will describe the key properties of complementizer doubling constructions: first, the movement of the intervening element; second, the absence of interpretative differences between complementizer doubling constructions and their counterparts without an overt doubling complementizer; and third, the restriction of complementizer doubling to subordinate clauses, and more specifically, to argument clauses/clauses in the syntactic position of argument.

2 Properties of Slovenian complementizer doubling constructions

This section describes and exemplifies three key properties of complementizer doubling constructions in Slovenian linked to their structure and interpretation. We will first show that phrases that occur between the two complementizers must have moved to that position from a lower structural position. Then we will argue that there is no interpretative difference between the complementizer doubling construction and its counterpart without a doubling complementizer. And thirdly, we will claim that Slovenian complementizer doubling is not just an example of speech disfluency, but rather a syntactic phenomenon available only in argument clauses.

2.1 Movement of the intervening element

As shown in (3), the phrase *svojo mamo* 'one's own mother', which sits in the left edge between the two complementizers and contains the bound reflexive adjective *svojo* 'one's own', is bound by *vsak* 'everyone', whose surface position would appear to suggests that it is located lower in the hierarchical structure of the embedded clause. By Principle A of the traditional binding theory, reflexives must have a local antecedent, which essentially means that the phrase that contains the bound reflexive in (3) must have been originally located in the embedded clause (Chomsky 1981).

(3) Rekel je, da svojo$_i$ mamo da ima vsak$_i$ rad.
      said   AUX that one's mother that has everyone like
      'He said that everyone likes their own mother.'

More direct evidence for the claim that the surface position of the phrase *svojo mamo* 'one's own mother' in (3) is a derived position is shown in (4), where movement of the reflexive is illustrated step by step. (4a) is the example with the most salient or preferable word order of the structurally incorporated numeration elements from sentence (3), in which nothing has moved into the CP field, and (4b) is the example with movement of the phrase that contains the bound reflexive from the original structural position into the next available structural position, though not yet as high as the CP field. Another available position for the phrase with the bound reflexive is shown in (4c): since only a copy is left after the movement of this phrase through the CP field of the embedded clause, complementizer doubling is not available in (4c).

(4)   a. Rekel je,   da ima vsak$_i$ rad svojo mamo$_i$.
        said AUX that has everyone like one's mother
        'He said that everyone likes their own mother.'

      b. Rekel je,   da ima svojo mamo$_i$ vsak$_i$ rad.
        said   AUX that has one's mother everyone like
        'He said that everyone likes their own mother.'

      c. Svojo mamo$_i$ je   rekel, da ima vsak$_i$ rad.
        one's mother AUX said   that has everyone like
        'He said that everyone likes their own mother.'

Of course, if the movement explanation from above is accepted for cases like (3), then it is reasonable to try and pursue the approach more generally, among others also for cases like (5), in which the intervening element is not an argument phrase but an adjunct adverbial clause. In other words, we would expect that all intervening constituents found between the two complementizers in complementizer doubling constructions, including adverbial clauses, have moved to the intervening position from their original position, which is lower in the sentence structure. (5) below confirms the correctness of this approach for an intervening adverbial clause adjunct.

(5)   a. Rekel je,   da vsak$_i$ žaluje, če izgubi svojo mamo$_i$.
        said AUX that everyone grieves if loses   one's mother
        'He said that everyone grieves if they lose their mother.'

      b. Rekel je,   da če izgubi svojo mamo$_i$, da vsak$_i$ žaluje.
        said   AUX that if loses   one's mother that everyone grieves
        'He said that everyone grieves if they lose their mother.'

The examples in (5) are direct parallels to the examples in (4). A comparison of (5a) and (5b) shows that the reflexive adjective *svojo* 'one's own', located within an adjunct adverbial clause, must have originated, like the one in (4), lower in the sentence structure or else it could not have satisfied the conditions set by the binding theory's Principle A.

2.2 Interpretation of the complementizer doubling construction and its counterpart without an overt doubling complementizer

Crucially, there is no interpretative difference between cases with the doubling complementizer and their minimal-pair counterparts without the second complementizer, such as (1a) and (1b) above. In the complementizer doubling construction the second, doubling *da* 'that' does not seem to function as anything

other than a doubling complementizer, clearly not having the echo- or discourse-marking role that in certain other cases, such as (6)–(7), *da* 'that' does also have.

(6) A: Petra pride.
        Petra come
        'Petra is coming.'

     B: Kᴅᴏ da pride?
        who that come
        'Who's coming, again?'

(7) Misliš da?
     think that
     'Do you think so?'

As can be seen from the contrasts between examples (1a), (1b), (6) and (7) the Slovenian element *da* 'that' can be used to serve at least four different functions. It can function as a complement clause introducer, as in (1a), as a doubling complementizer in complementizer doubling constructions, as in (1b), as an echo marker, as shown in (6), or as a discourse marker, as (7) shows. The last two functions will be left aside in the remaining part of the paper and only *da* 'that' elements in the function of primary complementizer and *da* 'that' elements in function of doubling complementizer will be the focus of our analysis, since these are the only available candidates that can fill the two edge projections of the complementizer system.

2.3 Restriction of complementizer doubling to argument clauses

In Slovenian, complementizer doubling is possible only in sentences with true complement clauses (clauses in the syntactic position of argument) introduced by the complementizer *da* 'that', as in (8), marginally also in clauses introduced by the complementizer *če* 'if', as illustrated in (9), but not in adverbial or adjectival (relative) subordinate clauses, as shown, respectively, in (10)–(12).[5,6]

---

[5] As will be shown in §3, the complementizer *če* 'if' is possible, but marked, as an introducer of a subject clause in copula sentences or in certain cases as an introducer of a complement clause. In this type of contexts *če* can be used as an alternative to the complementizer *da*. Although the use of this declarative *če*, when compared to *da*, does seem to bring a certain semantic difference, this difference is not directly relevant for the purposes of this paper.

[6] In (10)–(12), the elements *ki* 'who', *ker* 'because' and *ko* 'when' are complementizers that introduce a subordinate clause and are therefore, in this function, more like the complement-clause introducing *da* 'that' from (8) or the subject-clause introducing *če* 'if' from (9) than like a wh-question word (as one might incorrectly conclude especially on the basis of the glosses in examples (10) and (12)).

(8) Sem pa za to, da ker   imamo zadevo na dnevnem redu, da jo
     am   PTCL for this that because have   matter on daily   agenda that it
     čimprej   tudi zaključimo …
     as.soon.as.possible also finish
     'I think that since the matter is already on the agenda, the only
     reasonable thing is to conclude it as soon as possible …' (Gigafida corpus)

(9) Najslabše je, če ker   te ne razume,   (*? če) kričiš.
     worst   is if because you not understand   if   scream
     'The worst is if you scream because he does not understand you.'

(10) Vse preveč   je tistih, ki   ko   naredijo izpit, (*ki) mislijo, da
     all too.many of them, who when pass   exam   who think   that
     znajo vozit.
     know drive
     'There are too many of those people who, when they receive the driver's
     license, think that already know how to drive a car.'

(11) Ne sme se   premaknit, ker   če se premakne, (*ker)   mu lahko
     not may AUX move   because if refl. move   because him may
     počijo kosti.
     crack bones
     'He should not move, because if he moves his bones may crack.'

(12) Ni mu všeč, ko   ker   je jezna, (*ko) mu ne skuha kosila.
     not him like when because AUX angry   when him not cook   lunch
     'He is not happy when she does not make him lunch because she is angry.'

What examples (8) to (12) above show is that the discussed Slovenian comple-
mentizer doubling is a syntactic phenomenon, not just an example of speech
disfluency.[7] If it was not a syntactic phenomenon but just a disfluency-type of
repetition, one would expect it to be available in any type of subordinate clause
introduced by a subordinate complementizer, including adverbial and adjectival
subordinate clauses, rather than being restricted by syntactic context. Clearly,
the way the subordinate clause is incorporated into the sentence structure is one
of the key characteristics of the observed complementizer doubling phenomenon.

---

[7]Note that when asked for a judgment, speakers emphasised that these examples seem a bit odd,
but certainly possible. The speakers judged the examples as clumsy – clumsy with respect to
the standard (normative) Slovenian. However, complementizer doubling can easily be found
in Slovenian corpora, e.g. Gigafida (the largest corpus of contemporary written Slovenian).

What we conclude from the data introduced so far, then, is that the Slovenian complementizer doubling under discussion can occur in clauses in the syntactic position of argument, but not in embedded adjunct clauses.

Summing up §2.1 through §2.3, we showed that phrases positioned between the two complementizers of the embedded clause must have moved to that position from a lower position of a sentence structure. We demonstrated that the presence of a doubling complementizer in complementizer doubling constructions does not result in a difference in the meaning of the sentence, i.e., sentences with a second, doubling complementizer have the same meaning as their counterparts without a doubling complementizer. We also showed that complementizer doubling is only available in argument clauses, from which we concluded that complementizer doubling is a syntactic phenomenon.

3 Doubly-filled COMP filter vs. complementizer doubling

In this section we will briefly introduce another Slovenian construction recently discussed by Bacskai-Atkari (2018), which was argued there to represent a violation of and thus a counterexample to the doubly-filled COMP filter. At first sight, this construction seems very close to our complementizer doubling constructions, so the question arises whether Bacskai-Atkari's analysis can be used to capture our complementizer doubling constructions as well. We will establish that these constructions are not one and the same phenomenon (even though in principle, there could be an indirect relation between these two phenomena). We will argue that there is a structural and interpretative difference between Slovenian sentences that allow complementizer doubling and Bacskai-Atkari's (2018) sentences. While the second complementizer in the doubly-filled COMP constructions always contributes some additional meaning to the sentence interpretation, no such interpretative difference is contributed by the second complementizer in the complementizer doubling constructions.

An example of Bacskai-Atkari's (2018) doubly-filled COMP filter construction (her data is from Hladnik 2010) is in (13).[8]

(13) Vprašal je, če   da pride.
      asked   AUX whether that comes
      'He asked whether it was true that he was coming.'

---

[8]In Bacskai-Atkari (2018), the sentence is marked with a question mark. In Hladnik (2010: (15)), it has no such marking.

According to Bacskai-Atkari, such sentences represent structures with a simultaneously realized specifier and head position of the same CP. Crucially, the presence of the second complementizer (*da*) in (13) above is responsible for the 'it-was-true-that' part of the interpretation: without *da*, the interpretation of the sentence would be just 'He asked whether he was coming'. This means that (13) is felicitous only when used as a response to a statement such as 'He is coming' (Bacskai-Atkari 2018). In our complementizer-doubling Slovenian examples from above, on the other hand, the second complementizer does not seem to contribute any special interpretational difference, and there seems to be no special meaning-related requirement for the use of complementizer doubling construction in Slovenian. Recall from the minimal-pair sentences in (1) and §2.2 that the structure with and without the second complementizer both have the same interpretation, which would suggest that the realization of the doubling complementizer in Slovenian complementizer doubling constructions is optional.

The lack of semantic effects thus clearly separates our complementizer doubling constructions from Bacskai-Atkari's (2018) doubly-filled COMP filter construction. In addition, our complementizer doubling sentences contain not just two complementizers but also a phrase that has moved into the CP layer. For these constructions Bacskai-Atkari's approach therefore cannot work, because there must be at least three positions available in the complementizer system structure of the embedded argument clause to accommodate this CP material. This is sketched in Figure 2, where, as will become clear below, the topmost functional projection (responsible for typing the clause) hosts the first complementizer, the specifier of the intermediate functional projection (TopicP or FocusP) hosts the moved constituent, and the lowest projection hosts the doubling complementizer.

It was shown above that the complementizer doubling and the doubly-filled COMP constructions are not the same phenomenon, contrary to what one might assume on the basis of the superficial similarity between the complementizers involved in these two construction types. Now that we have established the difference between these two phenomena, we can focus on complementizer doubling more closely. This closer examination of complementizer doubling in the next section will then form the basis for the analysis provided in §5, which will focus on determining the syntactic positions that the complementizers occupy in the doubling construction.

```
                           CP
                    ╱            ╲
                  XP              C′
                            ╱          ╲
                          C⁰              CP
                          │          ╱        ╲
                      that/if  ZP          C′
                               │        ╱       ╲
                            moved     C⁰           CP
                          constituent  │        ╱      ╲
                                     YP          C′
                                            ╱         ╲
                                          C⁰            …
                                          │
                                      that/if
```
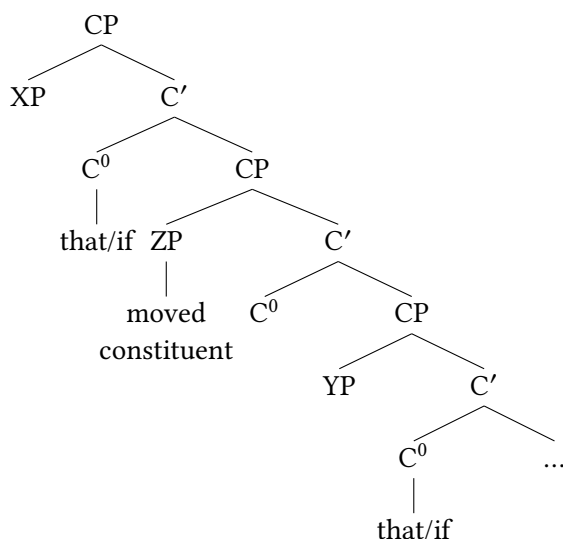
Figure 2: The structure of complementizer doubling construction

4 Further characteristics of Slovenian complementizer doubling

As we already saw in §2.3 above, *če* 'if' is another complementizer that can, in addition to *da* 'that', function as a doubling complementizer (albeit with some degree of degradation). These two complementizers, however, are marked by a clear difference in terms of their doubling positions. The declarative complementizer *če* 'if' can be doubled in the subject clause of copula sentences with predicates such as *pametno je* 'smart is' or *najslabše je* 'the worst is', as shown in examples (14) and (15) below, or in complement clauses with verbs such as *prositi* 'ask/request', as in (16).

(14)   ?? Pametno bi   bilo, če   da se izogneš dezinterpretaciji, če mu
          smart   would be if.DECL that you avoid   misinterpretation if him
          svoje stališče jasno   predstaviš.
          your opinion clearly present
          'It would be good if you state your position clearly, so that you avoid
          misinterpretation.'

(15) ^{??} Najslabše je, če   ker   te ne razume,   če kričiš.
    worst   AUX if.DECL because you not understand if scream
    'The worst is if you scream because he does not understand you.'

(16) ^{??} Prosil je, če   ko   konča z   delom, če pospravi za   sabo.
     request AUX if.DECL when finish with work   if clean.up after yourself
     'He asked if he could clean up after himself when he finishes with his
     work.'

On the other hand, the doubling construction is possible with *da* used in complement clauses introduced by the verbs such as *misliti* 'think' or *reči* 'say', as
shown in (1b) and (2) above, as well as with *da* used in any other environment
where *če* is possible; compare examples (15) and (16) above with examples (17)
and (18) below.

(17) Najslabše je, da ker   te ne razume,   da kričiš.
     worst   AUX that because you not understand that scream
     'The worst is if you scream because he does not understand you.'

(18) Prosil je, da ko   konča z   delom, da pospravi za   sabo.
     request AUX that when finish with work   that clean.up after yourself
     'He asked if he could clean up after himself when he finishes with his
     work.'

More accurately, it is not just *possible* for *da* to be used in positions available for
če: according to our informants, the use of *da* rather than *če* actually improves
the acceptability of such doubling constructions, compare (14), (15) and (16) above
with (19), (20) and (21) below. Moreover, the use of *da* seems to be more natural in
the function of complement subordinator than the use of *če*; compare the contrast
between examples (15) or (20) and (17).

(19) Pametno bi   bilo, če   da se izogneš dezinterpretaciji, da mu
     smart   would be if.DECL that you avoid   misinterpretation that him
     svoje stališče jasno predstaviš.
     your opinion clearly present
     'It would be good if you state your position clearly, so that you avoid
     misinterpretation.'

(20) [?] Najslabše je, če   ker   te ne razume,   da kričiš.
    worst   AUX if.DECL because you not understand that scream
    'The worst is if you scream because he does not understand you.'

(21) $^?$ Prosil je, če ko konča z delom, da pospravi za
      request AUX if.DECL when finish with work that clean.up after
      sabo.
      yourself
      'He asked if he could clean up after himself when he finishes with his
      work.'

The acceptability of a particular complementizer in the doubling construction
thus appears to depend on the matrix predicate, which can also be confirmed
with the availability of all four different combinations in cases where the matrix
verb is such that it accepts either *da* or *če*, as attested through the set of examples
in (15), (17) and (20) above and (22) below.

(22) $^{??}$ Najslabše je, da ker te ne razume, če kričiš.
      worst AUX that because you not understand if scream
      'The worst is if you scream because he does not understand you.'

Given that the acceptability of *da* or *če* in the complementizer doubling construc-
tion depends on the requirements of the matrix predicate, it is not surprising
that the same type of matrix-predicate dependence actually holds for the use of
da/*če* outside the doubling construction. On the one hand, if the matrix predi-
cate allows the use of either one of these complementizers outside the doubling
construction, as in (23) and (24), then both are also acceptable in the doubling
construction.

(23) Najslabše je, če / da kričiš.
      worst AUX if.DECL that scream
      'The worst is if you scream.'

(24) Prosil je, če / da pospravi za sabo.
      request AUX if.DECL that clean.up after yourself
      'He asked if he could clean up after himself.'

On the other hand, in sentences with *misliti* 'think' or *reči* 'say' as the matrix
verb, where the introducer of the complement clause can only be *da*, the use of
če is impossible, as shown by (25)–(26).

(25) * Mislim, če mu pripada plačilo.
      think if.DECL he belong payement
      Intended: 'I think that he should get paid.'

(26) # Rekel je, če   ne zaupa Petrovim prijateljem.
      said AUX if.DECL not trust Peter's   friends
      Intended: 'He said that he doesn't trust Peter's friends.'

In addition to its declarative use from above, *če* 'if' can also function as an introducer of an embedded yes/no question. Like its declarative use, *če*'s interrogative use also allows doubling, as shown in (27) below.

(27) Sprašuje se, če   ker   ga nihče ne posluša, če naj   še   kar
     ask   self if.INT because him nobody not listen   if should more still
     naprej govori.
     on   talk
     'He wonders if he should keep talking, given that no one is listening to him.'

Note that examples very similar to our Slovenian doubling examples from above have also been observed in non-standard varieties of English. Specifically, McCloskey (2006) reports complementizer doubling in declarative and interrogative contexts of the type shown in (28) and (29), which he analyzes with two CPs, one stacked on top of the other.

(28) He thinks that if you are in a bilingual classroom that you will not be
     encouraged to learn English.   (McCloskey 2006: 23, (69b))

(29) John was asking me if, when the house was sold, would they move back
     to Derry.   (McCloskey 2006: 24, (72c))

In perfect parallel to Slovenian, doubling of the English declarative complementizer *that* is realized in complement clauses introduced by verbs like *think*, as shown in (28) above. On the other hand, the parallel between Slovenian interrogative *če* doubling and the English (29) is less straightforward; (29) is a less transparent case of complementizer doubling. McCloskey (2006) argues, however, that (29) nevertheless shows clear evidence for the presence of two CPs through the simoultaneous presence of both *if* and the auxiliary-subject inversion. More specifically, all the CP material in (29) is evidence for the presence of three distinct CP-field positions; one for the yes/no clause introducer *if*, one for the topicalized constituent *when the house was sold* and another one for the inverted auxiliary *would*. Even though (28) and (29) are thus both analyzed as cases of complementizer doubling, they are also marked by a difference, namely, only (28) shows actual doubling of the lexical complementizer.

Similarly, Slovenian also shows a difference between the cases of declarative complementizer doubling and interrogative complementizer doubling. Doubling of *če* in its declarative use is somewhat degraded, and the acceptability of such complementizer doubling examples improves if *da* is used instead, as was shown above. In contexts of embedded yes/no questions, however, only *če* can occupy the position of the doubling complementizer, as can be seen from the comparison between (27) above and (30) below.

(30) *Sprašuje se, če   ker   ga nihče ne posluša, da naj   še
        ask   self if.INT because him nobody not listen   that should more
        kar naprej govori.
        still on   talk
        Intended: 'He wonders if he should keep talking, given that no one is
        listening to him.'

A comparison between doubling in the two contexts, i. e., complement clauses and embedded yes/no questions, implies that although complementizer doubling is possible in both types of constructions, there is an additional restriction in the case of the latter. This points to the fact that the complement clause introducer *da* 'that' and the embedded yes/no clause introducer *če* 'if' are not the same element with respect to their function in the sentence structure. This is further illustrated through the contrast between (31) below and (27) above. (31) shows that the substitution of *če* from (27) above with *da* is not possible in the case of a yes/no embedded question, which is required by the matrix verb *spraševati* 'ask'.

(31) *Sprašuje se, da ker   ga nihče ne posluša, da naj   še
        ask   self that because him nobody not listen   that should more
        kar naprej govori.
        still on   talk
        Intended: 'He wonders if he should keep talking, given that no one is
        listening to him.'

We take this to suggest that, on the one hand, *da* and the declarative *če* occupy the same structural position – the head of the highest CP projection, force projection – since in certain cases they can be used interchangeably, as shown in (15) and (17). On the other hand, despite the fact that it also allows doubling, the embedded yes/no question introducer must be assigned a different position, which makes it impossible for *da* to freely take this position. The fact that it is only in indirect questions that the doubling complementizer must be identical in form to the doubled one does not suggest that we are dealing with two different types of

complementizer doubling constructions; rather, it only confirms the idea that Slovenian complementizer doubling is restricted by requirements of the matrix verb, i.e., both in declarative and in interrogative complement clauses. In fact, Slovenian also seems to allow the doubling of the yes/no-question operator in embedded questions, as shown in (32), further suggesting that we are dealing with a single system of C-doubling.

(32) Vprašal je, a ker   ga ne mara, a naj   kar gre?
      ask   AUX Q because him not like   Q should just leave
      'He asked if he should leave because s(he) don't like him..'

To recapitulate, §3 showed that Bacskai-Atkari's (2018) doubly-filled COMP phenomenon and our complementizer doubling realize two different types of CP configurations. For the realization of the doubly-filled COMP construction, it suffices to have one CP projection, whereas at least three positions are needed to derive our complementizer doubling. The second complementizer of the doubly-filled COMP construction contributes additional meaning to the sentence, while our doubling complementizer does not affect the interpretation of the sentence. Furthermore, as was shown in §4, our complementizer doubling is not exhibited only in declarative complement clauses but also in embedded yes/no questions, and in both of these, the use of the complementizer is governed by the matrix verb.

In the next section we will lay out our analysis of the Slovenian complementizer doubling construction, framing this also in the context of a comparison with complementizer doubling, and its analysis, in Romance.

5 Analysis

We will claim that the Slovenian complementizer doubling construction can best be explained with Rizzi's (1997; 2001) split CP model, and more specifically with the model of Slovenian complementizer field proposed by Mišmaš (2015), which is sketched in Figure 1 above. We will argue that in declarative complement clauses, the first complementizer sits in the highest part of the CP field, namely in the head of the force projection, and the second complementizer in the head of finiteness projection. (For the doubled embedded yes/no clause introducer *če* 'if', we will simply assume that, like in Italian (see Rizzi 2001), it occupies the interrogative projection (InterP) in Figure 1 above.)

As has been observed for the Oïl dialect of French (Dagnac 2012), Slovenian complementizer doubling is closely related to finiteness of the subordinate clause,

with non-finite subordinate clauses not allowing it. The existence of this restriction, however, is not surprising since Slovenian non-finite subordinate clauses are never introduced by a subordinator, so that the use of *da* is ungrammatical regardless of whether the latter is doubled or not, as shown in (33).

(33) *Peter je hotel, da ko   ga ne bo nihče videl, (da) prositi
       Peter AUX want that when he not AUX nobody see   that ask
       Metko za pomoč.
       Metka for help
       Intended: 'Peter wanted to ask Metka for help when no one sees him.'

In other words, whenever *da* can appear in a structure as the subordinator, complementizer doubling can also occur; but when the use of *da* is ungrammatical, doubling cannot occur. But even though Slovenian finite embedded clauses are always introduced by a complementizer and non-finite ones are always without a complementizer, we still expect the information about finiteness/non-finiteness to be present in both types of sentences; as argued by Rizzi (1997), the information carried by FinP expresses a distinction related to tense.

Positing a link between complementizers and the finiteness projection is no novelty either. That complementizers can express distinctions related to tense has been established on the basis of Irish, where sensitivity of the complementizer to the tense of the embedded clause is reflected in the form of the complementizer (van Koppen 2017). As shown by (34), the past-tense form of the complementizer differs from the form used for all other tenses. If the embedded clause shows future tense, as in (34a), the complementizer is go, but when it shows past tense, the complementizer gets a past-tense marker *-r*, surfacing as *gur*, as in (34b).

(34) a. Deir   sé go dtógfaidh sé an peann.
         say.PRES he that take.FUT   he the pen
         'He says that he will take the pen.'

      b. Deir   sé gur   thóg   sé an peann.
         say.PRES he that.PAST take.PAST he the pen
         'He says that he took the pen.'

                                    (Cottell 1995, as cited in van Koppen 2017: (3))

Based especially on evidence from dialectal variants of Romance languages (e.g. Paoli 2003; Mascarenhas 2007; Villa-Garcia 2012; González i Planas 2014; Munaro 2016), it has been proposed that the doubling complementizer in those languages sits in the head of the topic projection. The main evidence for this position is

the fact that the doubling complementizer is never found after a focused phrase, as pointed out for Spanish by Villa-Garcia (2012), consider the contrast between (35a) and (35b) (see also Ledgeway 2005 for comparable data from older varieties of Italian).

(35) a. Me dijeron que a tu primo, que SOLO DOS PORTÁTILES le robaron
       CL said   that your cousin that only two laptops   CL stole
       (no tres).
        not three
       'They told me that it was only two laptops that your cousin got
       stolen, not three.'

     b. * Me dijeron que SOLO DOS PORTÁTILES, que le robaron a tu primo (no
       tres).
       (Villa-Garcia 2012: 30, (24a) and (24b))

This analysis cannot be adopted for our Slovenian data, however, because it wrongly predicts that (36) below will be unacceptable, since the doubling complementizer in it is found after a focused phrase, specifically, after the contrastively focused *dve pivi* 'two beers'.

(36) Rekel je, da DVE PIVI da je včeraj   zvečer spil (in ne
      said   AUX that two beers that AUX yesterday evening drink and not
      treh).
      three
      'He said that he only drank two beers yesterday evening (and not three).'

In fact, (36) could be taken as suggesting the head of finiteness phrase as the site of the doubling complementizer, because the left periphery of the embedded clause is usually seen as not featuring a topic phrase located after the contrastive focus phrase to potentially host the doubling complementizer (Rizzi 1997), and then the next available position (moving downward along the CP field) for the doubling complementizer in this particular case is the head of finiteness phrase. Although on the other hand, given Mišmaš's (2015) structure of Slovenian left periphery in Figure 1 above, Slovenian topic phrases can appear to the right of a focused phrase, in which case (36) above may not be conclusive.

   Furthermore, multiple occurrences of complementizers in the CP field are, contrary to what has been observed for Romance languages (e.g. Ledgeway 2005; Mascarenhas 2007; Villa-Garcia 2012), not acceptable with multiple topicalized phrases in Slovenian: compare (37) from European Portuguese and (38a) from

Slovenian. In Slovenian, only the complementizer that opens and the lower one that closes the complementizer system can be realized, as in (38b), supporting our claim that the position in which doubling complementizers are located in Slovenian is the head of finiteness phrase, as well as again suggesting that the analysis developed for Romance cannot be adopted for Slovenian.

(37) Acho que amanhã   que a   Ana que vai conseguir acabar   o
      think that tomorrow that the Ana that will manage   to finish the
      trabalho.
      assignment
      'I think that tomorrow Ana will manage to finish the assignment.'

                                                          (Mascarenhas 2007: 6, (20))

(38) a. *Ne morem verjet, da   ko   pospravlja stanovanje, da   Andreja
         not can   believe that when clean   apartment   that Andreja
         da   pomete smeti pod   preprogo.
         that sweep   dirt   under rug
         Intended: 'I cannot believe that Andreja sweeps the dirt under the
         rug when she cleans the apartment.'

      b. Ne morem verjet, da   ko   pospravlja stanovanje, da   Andreja
         not can   believe that when clean   apartment   that Andreja
         pomete smeti pod   preprogo.
         sweep   dirt   under rug
         'I cannot believe that Andreja sweeps the dirt under the rug when
         she cleans the apartment.'

In sum, this section showed that in Slovenian, contrary to what has been found for the Romance languages, the doubling complementizer can occur after a focused phrase, and it is only the complementizer that opens and the one that closes the complementizer system that can be realized in sentences with multiple topicalized phrases, both of which separate Slovenian complementizer doubling from the doubling in Romance languages. Furthermore, it was also suggested that the position hosting doubling complementizers in Slovenian is the head of finiteness phrase. The main piece of support for this was the restriction on complementizer doubling set by the finiteness system of the clause.

When a focused phrase is present in the sentence structure, complementizer tripling is also possible, with the focused and topicalized phrases freely ordered, as shown in (39) and (40).

(39) Rekel mu je, da TRI KNJIGE da če hoče naredit izpit, da mora
said him AUX that three books that if want pass exam that need
prebrat (in ne dveh).
read (and not two)
'He told him that that he should read three books (not two) if he wants to
pass the exam.'

(40) Rekel mu je, da če hoče naredit izpit, da TRI KNJIGE da mora
said him AUX that if want pass exam that three books that need
prebrat (in ne dveh).
read (and not two)
'He told him that if he wants to pass the exam he should read three books
(not two).'

As is the case with complementizer doubling, complementizer tripling construc-
tions also seem to show no significant difference in meaning when the second
or the third complementizer is present in the structure. In contrast to comple-
mentizer doubling, complementizer tripling poses an additional question about
the exact position of the linearly second complementizer. Especially given that
the focused and topicalized phrases can surface in either order, it may well be
that the position of the second complementizer of (39) is actually not the same
as the position of the second complementizer of (40). What seems clear is that in
a structure like Figure 3 above, the second complementizers of both (39) and (40)
should sit in one of the available projections between InterP and FinP. Still, for
any real claims to be made regarding the tripling phenomenon, a more in-depth
investigation will be necessary, but this would go well beyond the scope of this
paper and will thus have to wait for future work.

6 Conclusion

In this paper we treated Slovenian complementizer doubling as an inherent prop-
erty of the subordinate clause which can be realized only if there is some addi-
tional element in the specifier position of some projection between force phrase
and finiteness phrase.

We showed that the phrase occurring between the two complementizers of
the embedded clause must have moved there from a lower structural position,
and we claimed that the presence of the doubling complementizer does not re-
sult in any difference in the meaning of the sentence. We suggested that since

complementizer doubling is possible only in complement clauses, this must be a syntactic phenomenon rather than a disfluency-type repetition.

We also argued that the doubly-filled COMP sentences of Bacskai-Atkari (2018) and our complementizer doubling constructions are separate phenomena, or rather, that the complementizer in Bacskai-Atkari's doubly-filled COMP sentences and our doubling complementizer are two different elements. The realization of the doubly-filled COMP construction requires just one CP projection, whereas this does not suffice for our complementizer doubling construction which requires at least three structural positions. Furthermore, we showed that complementizer doubling is also possible in embedded yes/no questions, which confirms the idea that Slovenian complementizer doubling is constrained by the requirements of the matrix predicate.

We suggested that Slovenian complementizer doubling can be nicely accounted for if we assume a split CP model (Rizzi 1997; 2001). A comparison between the characteristics of Slovenian and Romance complementizer doubling revealed that the location occupied by our doubling complementizer cannot be the head of topic projection. We proposed that in Slovenian, the doubling complementizer is hosted by the head of finiteness phrase, while the first complementizer is the highest element of the embedded clause and as such located in force projection.

Abbreviations

AUX	auxiliary	INT	interogative
CL	clitic	PRES	present tense
DECL	declarative	PAST	past tense
FUT	future	PTCL	particle

Acknowledgements

References

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

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrech: Foris.

Cottell, Siobhán. 1995. The representation of tense in Modern Irish. *Geneva generative papers* 3. 105–124.

Dagnac, Anne. 2012. How do you double your C? Evidence from an Oïl dialect. In Christopher Piñón (ed.), *Empirical Issues in Syntax and Semantics 9*, 77–94. Paris: CSSP. http://www.cssp.cnrs.fr/eiss9/eiss9_dagnac.pdf.

González i Planas, Francesc. 2014. On quotative recomplementation: Between pragmatics and morphosyntax. *Lingua* 146. 39–74. DOI:10.1016/j.lingua.2014.04.007

Hladnik, Marko. 2010. *Restrictive relative clauses in Slovene and its spoken varieties*. Utrecht: Utrecht University. (MA thesis). https://dspace.library.uu.nl/handle/1874/40140.

Ledgeway, Adam. 2005. Moving through the left periphery: The dual complementiser system in the dialects of Southern Italy 1. *Transactions of the Philological Society* 103(3). 339–396.

Lenertová, Denisa. 2001. On clitic placement, topicalization and CP-structure in Czech. In Gerhild Zybatow, Uwe Junghanns, Grit Mehlhorn & Luka Szucsich (eds.), *Current issues in formal Slavic linguistics*, 294–305. Frankfurt am Main: Peter Lang.

Marušič, Franc. 2008. Slovenian clitics have no unique syntactic position. In Andrei Antonenko, John F. Bailyn & Christina Bethin (eds.), *Formal Approaches to Slavic Linguistics 16: The Stony Brook Meeting 2007*, 266–281. Ann Arbor, MI: Michigan Slavic Publications.

Mascarenhas, Salvador. 2007. *Complementizer doubling in European Portuguese*. Manuscript, University of Amsterdam / New York University. https://www.researchgate.net/publication/241584392_Complementizer_doubling_in_European_Portuguese.

McCloskey, James. 2006. Questions and questioning in a local English. In Raffaella Zanuttini, Héctor Campos, Elena Herburger & Paul H. Portner (eds.), *Crosslinguistic research in syntax and semantics: Negation, tense, and clausal architecture*, 87–126. Washington, DC: Georgetown University Press.

Mišmaš, Petra. 2015. *On the optionality of wh-fronting in a multiple wh-fronting language.* University of Nova Gorica. (Doctoral dissertation). http://repozitorij.ung.si/Dokument.php?id=3236&lang=eng.

Munaro, Nicola. 2016. Localizing conditional clauses in the left periphery: Evidence from multiple complementizer constructions in Romance. *Linguistica* 56(1). 211–224. DOI:10.4312/linguistica.56.1.211-224

Paoli, Sandra. 2003. *COMP and the left-periphery: Comparative evidence from Romance.* Manchester: University of Manchester. (Doctoral dissertation).

Plesničar, Vesna. 2016. Podvajanje veznikov okoli podrednih stavkov v slovenščini [complementizer doubling around subordinate clauses in slovenian]. In Franc Marušič, Petra Mišmaš & Rok Žaucer (eds.), *Škrabčevi dnevi 9 (Zbornik prispevkov s simpozija 2015)*, 113–126. Nova Gorica: Založba Univerze v Novi Gorici. www.ung.si/media/storage/cms/attachments/2019/09/26/11/21/25/Skrabcevi_dnevi_10_11k18.pdf.

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

Rizzi, Luigi. 2001. On the position "Int(errogative)" in the left periphery of the clause. *Current studies in Italian syntax: Essays offered to Lorenzo Renzi.* 287–296.

Rizzi, Luigi. 2004. On the cartography of syntactic structures. In Luigi Rizzi (ed.), *The structure of CP and IP: The cartography of syntactic structures*, 3–17. Oxford: Oxford University Press.

van Koppen, Marjo. 2017. Complementizer agreement. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley–Blackwell companion to syntax*, 923–962. New York: Wiley–Blackwell. DOI:10.1002/9781118358733.wbsyncom061

Veselovská, Ludmila. 2008. The extended verbal projection in Czech: Three variants of the verb *být*. In Gerhild Zybatow, Luka Szucsich, Uwe Junghanns & Roland Meyer (eds.), *Formal Description of Slavic Languages: The Fifth Conference, Leipzig 2003*, 555–569. Frankfurt am Main: Peter Lang.

Villa-Garcia, Julio. 2012. *The Spanish complementizer system: Consequences for the syntax of dislocations and subjects, locality of movement, and clausal structure.* Storrs, CT: University of Connecticut. (Doctoral dissertation).

Chapter 11

Negation, comparative and alternatives: Experimental evidence from Czech

Iveta Šafratová

Masaryk University in Brno

The semantic interplay of negation with focus and scalar implicatures influences acceptability judgments. This paper describes two readings of sentences with comparatives and negation, namely the equality reading and the interval reading. The experiment provides evidence that sentences with negated comparatives prefer the equality reading in Czech. I argue that Czech negated comparatives result in the preferential equality reading as do English negated comparatives; but I challenge the claim that Czech negation *ne* 'no' activates focus alternatives, unlike in English negated comparatives with *no* where scalar alternatives cause the equality reading. I argue that focus alternatives and scalar alternatives are the same. Both Czech *ne-* 'not' and English *not* in verbal negation comparatives lead to the preferential equality reading if negation has narrow scope over the maximality operator.

Keywords: constituent negation; verbal negation; scalar implicatures; focus alternatives; comparative; Czech; experimental semantics

1 Introduction

In recent years, the topic of scalar implicatures and numerals has attracted considerable attention, and the research gave rise to several influential theories (Larson 1988; Krifka 1999; Sauerland 2004; Fox & Hackl 2006, among others). This article investigates how negation interacts with comparatives involving numerals. I will focus on the comparison of English data with Czech. Though the semantics of many different types of Slavic numerals have recently been explored with considerable success (e.g. Dočekal 2013; Wągiel 2014; 2015), so far little attention has

been dedicated to their behavior in the interaction with negation and comparatives (with the notable exceptions of Dočekal 2017 and Dočekal & Wągiel 2018 respectively). In this paper, I intend to shed new light on this topic by means of experimental investigation.

The article investigates how negation interacts with comparatives. We start with Rick Nouwen's observation about English negated comparatives (Nouwen 2008). He distinguishes two sub-types of comparatives, i.e., strict comparatives (*-er*) in (1a) and non-strict comparatives (*no(t) -er*) in (1b).

(1) a. John found more than 20 mushrooms.
     b. John found no more than 20 mushrooms.

Strict comparatives express either the relation *less* or *more*. We would expect that the only thing non-strict comparatives do is that they simply reverse the relation. But, according to Nouwen (2008), negation changes the relation from > to ≤ and from < to ≥. Non-strict comparatives are ambiguous due to their ability to express either *less/more* or *equality*.

Non-strict comparatives can be negated either by constituent negation (cn), as in (2a), or by verbal negation (vn), as in (2b).

(2) a. John found no more than 20 mushrooms.
     b. John did not find more than 20 mushrooms.

We focus on whether these two types of negation (vn/cn) influence the ambiguity of non-strict comparatives and whether the composition of the meaning of non-strict comparatives happens in the same way in English and Czech.

1.1 English non-strict comparatives

A sentence with a comparative activates scalar alternatives. Scalar alternatives are present in scalar implications developed by Horn (1989; 1996), they are a subtype of generalized conversational implicature associated with scalar values ordered from the weakest value to the strongest value (Horn 2013). Consider the following sample sentences with their scalar implicature (SI).

(3) Some people left.
     SI: ¬ all people left

(4) John has two children.
     SI: ¬ John has three children

(5) John bought a book or a pen.
    SI: ¬ John bought a book and a pen

The strict comparative (6a) and the non-strict comparative (6b) both lead to scalar implications, but with a different result.

(6) a. John found more than 20 mushrooms.
    b. John found no more than 20 mushrooms.

The strict comparative in (6a) means that the minimum number of mushrooms is more than 20. The limit lies in the lower bound, but the upper bound is unbounded, i.e. $(20,\infty)$.

(7) John found more than 20 mushrooms.
    a. truth conditions:
       $\text{MAX}_\text{D}\big(\lambda y \exists x[\#x = y \wedge \text{MUSHROOM}(x) \wedge \text{FIND}(\text{JOHN}, x)]\big) > 20$

The negation reverses the relation from $>$ to $\leq$; therefore the non-strict comparative in (6b) means that the maximum number of mushrooms is 20. The limit lies in the upper bound, whereas the lower bound is not specified, but the natural perception of the world limits the minimal number, i.e., $(0,20)$.

Nouwen (2008) explains the composition of non-strict comparatives. According to him, the non-strict comparative in (6b) has two readings: the interval reading corresponds to the relation *less* – $(0,20)$ and the equality reading corresponds to the *equality* relation – $(0,20)$.

If constituent negation negates a comparative (CN-comparative), as in (8), the equality reading results from the strengthening of the truth condition interpretation via scalar implicature. The sentence has standard truth conditions (8a) and also scalar implicatures (8b), (8c), etc., but scalar implicatures are negated because the proposition stating the number of mushrooms is ≤ 19 is logically stronger than the proposition stating the same number is ≤ 20.[1] The equality reading then arises from the denial of scalar implicatures and strengthening of truth conditions.

(8) John found no more than 20 mushrooms.

---

[1]Logical strength relates to the entailment. The proposition *John didn't find 19 mushrooms* entails the proposition *John didn't find 20 mushrooms*; therefore the first proposition is stronger than the second proposition.

    (i)   John didn't find 19 mushrooms → John didn't find 20 mushrooms → …

a. truth conditions: $\text{MAX}_\text{D}(\lambda d \,.\, \text{the number of mushrooms was } d \leq 20)$
b. SI: $\neg\text{MAX}_\text{D}(\lambda d \,.\, \text{the number of mushrooms was } d \leq 19)$
c. SI: $\neg\text{MAX}_\text{D}(\lambda d \,.\, \text{the number of mushrooms was } d \leq 18)$
d. $\text{MAX}_\text{D}(\lambda d \,.\, \text{the number of mushrooms was } d = 20)$

Nouwen (2008) claims that the most salient interpretation of English cn-comparatives is the equality reading, but he doesn't exclude the interval reading.[2]

If verbal negation negates a comparative (vn-comparative), as in (9), both readings are also possible but due to different reasons than in cn-comparatives. Nouwen (2008) claims that interval and equality readings arise because verbal negation can take two scopes within the proposition: narrow scope, as in (9a), which corresponds to (8a), and wide scope, as in (9b).

(9) John did not find more than 20 mushrooms.

a. $\text{MAX}_\text{D}\big(\lambda y \exists x [\#x = y \wedge \text{MUSHROOM}(x) \wedge \text{FIND}(\text{JOHN}, x)]\big) \leq 20$
b. $\neg\text{MAX}_\text{D}\big(\lambda y \exists x [\#x = y \wedge \text{MUSHROOM}(x) \wedge \text{FIND}(\text{JOHN}, x)]\big) > 20$

The narrow scope leads to scalar alternatives. As in the cn-comparatives, the scalar alternatives are negated, and the strengthening of truth conditions causes the equality reading. The wide scope is interpreted as a denial and truth conditions of the proposition are weaker: 'it is not true that John found more than 20 mushrooms'. Strengthening does not occur in the construction because no scalar alternatives are present; therefore only the interval reading is possible. Nouwen argues that sentences with vn-comparatives show a preference for the interval reading.

Table 1 summarizes Nouwen's observations of English negated comparatives.

Table 1: English negated comparatives

cn-comparatives	equality reading	preferred
cn-comparatives	interval reading	non-preferred
vn-comparatives (wide scope)	interval reading	preferred
vn-comparatives (wide scope)	equality reading	non-preferred
vn-comparatives (narrow scope)	equality reading	preferred
vn-comparatives (narrow scope)	interval reading	non-preferred

---

[2]The interval reading arises because scalar implicatures need not be drawn, and the strengthening of truth conditions doesn't occur.

Table 1 shows a nice pattern of English negated comparatives, i.e., CN-comparatives prefer the equality reading, whereas VN-comparatives prefer the interval reading if verbal negation takes wide scope over the maximality operator. The equality reading is preferred in the case of narrow scope of verbal negation. I now contrast this observation with Czech negated comparatives.

1.2 Czech non-strict comparatives

While English CN *no* and VN *not* differ morphologically, Czech CN and VN share the same morphological form *ne*, but its semantic and syntactic properties vary. The marker of CN is a free morpheme; it stands independently in a sentence, as in (10). The marker of VN is an ordinary prefixal verbal negation; it firmly connects to a lexical verb or an auxiliary, as in (11).

(10) Chci   být   doktorem,   ne učitelem.
      want.PRS.1SG be.INF doctor.INS.SG not teacher.INS.SG
      'I want to be a doctor, not a teacher.'

(11) Nechci   být   učitelem.
      NEG.want.PRS.1SG be.INF teacher.INS.SG
      'I don't want to be a teacher.'

Dočekal (2017) investigates non-strict comparatives in Slavic languages, especially in Czech. Following the previous investigation (Jasinskaja 2016, among others), Dočekal starts with an observation that Slavic focus particles have to c-command their focus marked constituents (12), and they have to be adjacent to the focus marked constituent, unlike English focus particles (13).[3] In this respect, Slavic focus particles resemble German focus particles (see Büring & Hartmann 2001).

(12) a. Tento   slovník   překládá   {pouze / ne} [z
          this.NOM.SG dictionary.NOM.SG translate.PRS.1SG only   not from
          angličtiny]_FOC   do svahilštiny.
          English.GEN.SG to Swahili.GEN.SG
          'This dictionary translates only/not from English to Swahili.'

---

[3]Dočekal (2017) gives examples with the Czech prototypical focus particle *pouze* 'only' and he claims that Czech CN behaves the same way.

    b. Tento   slovník   překládá   z   angličtiny
       this.NOM.SG dictionary.NOM.SG translate.PRS.1SG from English.GEN.SG
       pouze [do svahilštiny]FOC.
       only   to Swahili.GEN.SG
       'This dictionary translates from English to Swahili only.'

    c. Tento   slovník   překládá   z   angličtiny
       this.NOM.SG dictionary.NOM.SG translate.PRS.1SG from English.GEN.SG
       ne [do svahilštiny]FOC.
       not to Swahili.GEN.SG
       'This dictionary translates from English not to Swahili.'

(13)   a. I behave only [seriously]FOC.
      b. I only behave [seriously]FOC.   (Dočekal 2017)

Based on the pattern demonstrated above, Dočekal (2017) concludes that Czech CN is a focus particle, unlike English CN. Negated comparatives activate alternatives in both languages, but the type of alternatives differs: scalar alternatives in English (Nouwen 2008) and focus alternatives in Czech (Dočekal 2017).

We interpret the semantics of focus by the proposal of Rooth (1985; 1992): each sentence with focus has two semantic values: ordinary value $[\![\alpha]\!]^o$ and focus value $[\![\alpha]\!]^f$. Ordinary value is the truth-conditional value of α; focus value is the set of alternatives of α, as in (14b). Focus sensitive operators bear existential presuppositions of focus alternatives. The existential presupposition means that at least one alternative from the set of alternatives is true (see Rooth 1985; 1992).

(14)   a. Charles gave a rose to [Mary]FOC.
      b. {Charles gave a rose to x | x is a person}

CN adds the ¬ operator to a sentence; it negates the assertion and presupposes that at least one alternative is true, as in (15). The assertion is negated (15a), but the constituent negation – being a focus particle – introduces the presupposition in (15b) (which is satisfied by *a hoe*).

(15)   Maxwell killed the judge not with [a hammer]FOC, but with [a hoe]FOC.
      a. ¬ Maxwell killed the judge with a hammer
      b. Presupposition: $\exists x$[Maxwell killed the judge with x]

Dočekal (2017) selected and classified sentences with CN from the largest corpus of contemporary Czech, SYN2010. He formalizes the negated focus marked

constituent cross-categorically because Czech CN may modify various types of constituents, e.g., PP, AP, NumP, AdvP.[4]

The standard consideration about focus is that both the ordinary value and the focus value have to be of the same semantic type. We illustrate this in example (16): both constituents *a doctor* and *a teacher* are of type $\langle e, t \rangle$ and in this case are predicates. They have to share the same property because they belong to the set of alternatives (in this case professions). The ordinary value *a teacher* is negated by CN, and the focus value *a doctor* satisfies the need for at least one alternative to be true.

(16) Chci   být   doktorem,   ne učitelem.
      want.PRS.1SG be.INF doctor.INS.SG not teacher.INS.SG
      'I want to be a doctor, not a teacher.'
      a. ordinary value: a teacher
      b. focus value: a doctor

Dočekal proposes that the Czech focus operator CN targets the comparative morpheme *více než* 'no more' that has the following possible alternatives: ≤, =, >. At least one alternative must be valid. The alternative > is negated; therefore two alternatives remain: the alternative = leads to the equality reading, as in (17a), the alternative ≤ leads to the interval reading, as in (17b).

(17) Peter drank no [more than] 5_FOC beers yesterday.
      a. the equality reading: Peter drank no more, but exactly 5 beers.
      b. the interval reading: Peter drank not more, but less than 5 beers or equal.

Both Czech CN- and VN-comparatives are ambiguous between interval and equality reading, but Dočekal (2017) expresses the intuition that Czech CN-comparatives prefer the interval reading, whereas VN-comparatives prefer the equality reading.[5] Table 2 summarizes Dočekal's observations of Czech CN- comparatives.

---

[4]According to Dočekal (2017), Nouwen's explanation via negated scalar implicatures would not work for Czech because Czech CN modifies various types of constituents, not only numerical phrases. A compromise would be to say that in Czech CN-comparatives activate both scalar alternatives and focus alternatives.

[5]Dočekal (2017) has verified his intuition with a small corpus study; he has checked whether a context of the sentences approves the interval reading or the equality reading. Also, he has consulted the translation of a sample sentence with native speakers of Polish, Bulgarian and Russian. Based on the corpus study and the consultations with several native speakers of Slavic languages, Dočekal argues that CN-comparatives prefer the interval reading in Czech. He describes VN-comparatives only minimally, he mainly investigates CN-comparatives.

Table 2: Czech negated comparatives

cn-comparatives	interval reading	preferred
cn-comparatives	equality reading	non-preferred
vn-comparatives	equality reading	preferred
vn-comparatives	interval reading	non-preferred

Following Dočekal, I hypothesize that Czech cn-comparatives show a preference for the interval reading and vn-comparatives show a preference for the equality reading, although Czech vn-comparatives have been less explored concerning ambiguity than cn-comparatives.

A comparison of Table 1 and Table 2 shows that English cn-comparatives prefer the equality reading, whereas Czech cn-comparatives prefer the interval reading. The verbal negation can take two possible scopes. When English verbal negation is construed with wide scope, there is a preference for the interval reading, but the equality reading is preferred when verbal negation takes narrow scope. An observation of Czech verbal negation is not divided into two possible scopes, the summary is only such that Czech vn-comparatives prefer the equality reading.

The preference for the interval reading or the equality reading of Czech negated comparatives is just Dočekal's hypothesis; therefore it would be appropriate to verify it experimentally. Before introducing an experiment, we will use the remainder of this section to present another approach to alternatives.[6]

We argue so far that the difference between English negated comparatives and Czech negated comparatives comes from different alternatives. According to Nouwen (2008), English constituent negation involves scalar alternatives, but for Dočekal (2017), Czech constituent negation *ne* is a focus particle; therefore Czech constituent negation works with focus alternatives. Focus alternatives differ from scalar alternatives in the fact that scalar alternatives are ordered on a scale from the weakest alternative to the strongest one, whereas focus alternatives are not hierarchically ranked because they are on the same level.

In recent years, the distinction between scalar alternatives and focus alternatives faces doubts (see Katzir 2007; Fox & Katzir 2011; Fox & Spector 2018). Both scalar alternatives and focus alternatives are considered to be determined in the same way, namely as a contextual restriction of the focus value of the sentence. Based on that, alternatives are in a particular logical relationship with the as-

---

[6]Thanks to an anonymous reviewer for raising this point.

sertion or the prejacent. A unified nature of scalar and focus alternatives goes against standard type-theoretic definitions of focus values (Fox & Katzir 2011). For Krifka (1995), scalar items are inherently focused and their alternatives are scales. The revised theory of scalar and focus alternatives challenges Dočekal's assertion that Nouwen's explanation for English cn-comparatives cannot work for Czech cn-comparatives because, according to Dočekal (2017), negated comparatives in Czech activate a different type of alternatives, unlike in English.

As I indicated before (see footnote 4), if we keep the distinction between scalar and focus alternatives, we have to say that Czech cn-comparatives activate both types of alternatives. An approach unifying these two types of alternatives into one group provides a more comprehensive view of the topic of alternatives.

2 Experiment

I experimentally tested the availability of two readings of negated comparatives in Czech. I focus on (i) whether both cn- and vn-comparatives are ambiguous between the equality and interval readings, and (ii) whether as predicted by the initial hypothesis cn-comparatives show a preference for the interval reading, whereas vn-comparatives show a preference for the equality reading.

2.1 Method

2.1.1 Materials

The experiment consisted of a truth value judgment task that tested to what extent Czech native speakers accept sentences with cn-comparatives *ne víc* 'no more' and vn-comparatives *ne + VERB víc* 'not VERB more' with respect to two possible interpretations: the equality reading and the interval reading.[7]

The experiment tested whether a sentence fits a given context. The context preceded the target sentence. Participants were instructed to answer either "yes" (*věta je v daném kontextu pravdivá a vhodná* 'the sentence is true and appropriate in the given context') or "no" (*věta není v daném kontextu pravdivá a vhodná* 'the sentence is not true and appropriate in the given context').[8]

---

[7]The constituent negation *ne* 'no' stands alone in a sentence, whereas the pre-verbal negation *ne-* 'not' must be connected to a verb.

[8]I agree with an anonymous reviewer that a weaker (interval) reading is true in the equality context, but presumably less natural (acceptable). If a participant prioritizes the truth over naturalness, cn-comparatives, which would be judged unnatural in the equality context according to the initial hypothesis, receive an answer "yes" in the equality condition because

The experiment focused on these main issues:

- the interaction between a comparative and the constituent negation: *ne* + more 'no' (cn)

- the interaction between a comparative and the verbal negation: *ne-* + a verb 'not' (vn)

I investigated whether the syntactic position of comparatives could influence the acceptability of sentences; therefore comparative constructions in tested sentences appeared in two types of syntactic environment: (a) in a predicate position and (b) in an object position. Sentences in the experiment did not include comparatives in a subject position. The support for choosing the two positions (predicate and object) follows from Nirit Kadmon's theory (Kadmon 2001) that claims that the readings can differ in these two positions. I omitted the interaction between topic and cn-/vn-comparatives.

I tested the acceptability of each sentence in two contextual environments: equality (e) and interval (n).

Consequently, I tested the acceptability of negated comparatives in eight conditions:

1. ocne – cn-comparative in an object position against the equality context

2. ocnn – cn-comparative in an object position against the interval context

3. ovne – vn-comparative in an object position against the equality context

4. ovnn – vn-comparative in an object position against the interval context

5. pcne – cn-comparative in a predicate position against the equality context

6. pcnn – cn-comparative in a predicate position against the interval context

7. pvne – vn-comparative in a predicate position against the equality context

8. pvnn – vn-comparative in a predicate position against the interval context

---

cn-comparatives are true in the equality context. The weaker reading is also true in the interval context, but more natural. Even if a participant prioritizes the truth or naturalness, the answer should be "yes." But Figure 1 shows that participants answered "no" more often: the interval reading is less acceptable in the interval context than in the equality context.

There were 16 items – each item had 8 conditions (i.e., 128 sentences) – and there were 16 fillers. Each participant judged 16 items and 16 fillers and the order of stimuli was randomized.[9] The experiment was compiled from items and conditions in such a way that each item appeared only once in the experiment, whereas individual conditions cycled (repeated Latin-square design). All items were tested in all eight conditions 1–8 presented above.

Below, two sample items are shown: an item with CN-comparative in a predicate position in (20) and an item with VN-comparative in a predicate position in (21). Items were tested against two contexts: either against the equality context in (18) or against the interval context (19) (not against both contexts simultaneously).

(18) | Petr | nalil | víno | do | všech | 6 | sklenic. | E
Petr.NOM pour.PST.1SG wine.ACC.SG into all.GEN.PL 6 glass.GEN.PL
'Petr poured wine into all 6 glasses.'

(19) | Petr | nalil | víno | do | všech | 5 | sklenic. | N
Petr.NOM pour.PST.1SG wine.ACC.SG into all.GEN.PL 5 glass.GEN.PL
'Petr poured wine into all 5 glasses.'

(20) | Na | stole | je | ne | víc | než | 6 | sklenic. | PCN
on table.LOC.SG be.PRS.1SG no more than 6 glass.GEN.PL
'There are no more than 6 glasses on the table.'

(21) | Na | stole | není | víc | než | 6 | sklenic. | PVN
on table.LOC.SG NEG.be.PRS.1SG more than 6 glass.GEN.PL
'There are not more than 6 glasses on the table.'

The other two sample items are following: an item with CN-comparative in an object position in (24) and an item with VN-comparative in an object position in (25). Again, items were tested against two contexts: either against the equality context in (22) or the interval context (23).

(22) | Všech | 6 | růží | bylo | červených. | E
all.GEN.PL 6 rose.GEN.PL be.PST.1SG red.GEN.PL
'All 6 roses were red.'

(23) | Všech | 5 | růží | bylo | červených. | N
all.GEN.PL 5 rose.GEN.PL be.PST.1SG red.GEN.PL
'All 5 roses were red.'

---

[9]I used functions *shuffle sequence* and *rshuffle* in a statistical analysis. The functions are defined as randomizing and sequencing operations over an array of items in the Ibex manual.

(24) Jan   dal   Marii   k narozeninám   ne víc   než   6
John.NOM give.PST.1SG Mary.DAT to birthday.DAT.PL no more than 6
růží.   OCN
rose.GEN.PL
'John gave Mary no more than 6 roses for her birthday.'

(25) Jan   nedal   Marii   k narozeninám   víc   než   6
John.NOM NEG.give.PST.1SG Mary.DAT to birthday.DAT.PL more than 6
růží.   OVN
rose.GEN.PL
'John didn't give Mary more than 6 roses for her birthday.'

Each sentence is potentially ambiguous between the interval reading and the equality reading, but I assumed a preference for a particular reading. As predicted by the initial hypothesis, I expected that sentences with CN-comparatives like (20) and (24) will be more acceptable in the interval scenario, as in (19) and (23), whereas sentences with VN-comparatives like (21) and (25) will be more acceptable in the equality scenario, as in (18) and (22).

2.1.2 Procedure and participants

The experiment was run on Ibex and participants completed the experiment online. The experiment began with instructions and following that, the practice items and the truth value judgment and appropriateness task were presented.[10], [11] Participants were asked for the acceptability of a sentence against a context. 52 native speakers of Czech participated in the experiment, mainly students from Masaryk University.

2.2 Results

The fillers in the acceptability task were either uncontroversially grammatical or ungrammatical, and I checked whether the average of each participant's responses to ungrammatical fillers was lower than the average of his or her responses to grammatical fillers. Eight participants had to be excluded due to their unsuccessful scores in fillers; therefore I kept 44 subjects for subsequent analysis.

---

[10]The terms *true* and *false* are technical terms of logic and linguistics but the experiment targeted native speakers of Czech irrespective of their academic background; therefore I used *adequate* in the introduction.

[11]Follow the link below for the instructions and practice items, items and fillers used in the experiment: GitHub

I designed the truth value judgment task using a mixed-effects linear model. To model the data I constructed 4 linear models (the standard lm command of R) of the acceptability of the 4 conditions as depending on their E/N sub-conditions. The equality condition was taken as the reference level for each condition (PCN, PVN, OCN, OVN).

For all four conditions the linear model summary resulted in a statistically significant difference between E and N sub-conditions: the mean acceptability was a response vector depending on the E/N as a predictor. The statistical outcome was the following:[12]

1. CN-comparative in a predicate position (condition PCN):
   $t = -3.468, p = 0.000662$

2. VN-comparative in a predicate position (condition PVN):
   $t = -3.149, p = 0.00193$

3. CN-comparative in an object position (condition OCN):
   $t = -4.125, p = 5.73e - 05$

4. VN-comparative in an object position (condition OVN):
   $t = -3.207, p = 0.0016$

Figure 1 charts the boxplots of the acceptability ratings depending on the eight conditions and graphically displays the results using inferential statistics (the dot represents the mean).

Figure 1 shows that I did not find the expected interaction between VN and CN. The reference level for the E sub-condition is PCNE condition and the following statistical output shows that all conditions PCNE, PVNE, OCNE, OVNE are statistically not distinguishable from each other: PVNE: $t = 0, p = 1$; OCNE: $t = -1.226, p = 0.222$; OVNE: $t = -0.509, p = 0.612$. The reference level for the N sub-condition is PCNN condition and all conditions PCNN, PVNN, OCNN, OVNN are statistically not distinguishable from each other again. The statistical output is the following: PVNN: $t = 0.300, p = 0.764$; OCNN: $t = -0.600, p = 0.549$; OVNN: $t = -0.751, p = 0.454$. The formal statistical results report no expected interaction between CN and VN conditions.

The experiment provides the following evidence on the discussed interaction in Czech:

---

[12]The value t states how big a difference there is between the equality reading and the interval reading, the value p expresses how likely it is that the difference between the equality reading and the interval reading is random.
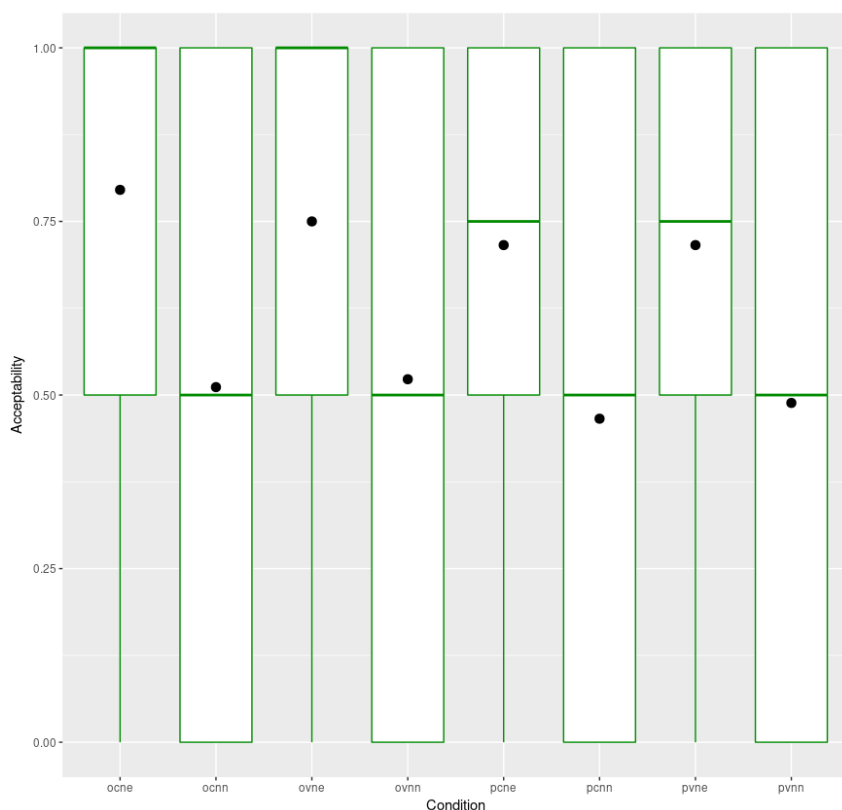
Figure 1: Means of responses average per subject and conditions (20)–(21), (24)–(25) depending on their sub-conditions (18)–(19), (22)–(23) (on the scale 0: inappropriate, 1: appropriate)

(A) Figure 1 shows that the acceptability rates of sentences with a comparative in a predicate position and the acceptability rates of sentences with a comparative in an object position are comparable. It does not support the assumption that the syntactic position of a comparative affects the acceptability of negated comparatives under a particular reading (but see Kadmon 2001).

(B) cn-comparatives and vn-comparatives are comparable.

(C) The acceptability of all four conditions was higher in the equality sub-condition (e) than in the interval sub-condition (n).

The experimental results show that both Czech cn- and vn-comparatives prefer

the equality reading; therefore they do not confirm the initial hypothesis that Czech CN-comparatives lead to the preferential interval reading. Now I turn to the questions: (i) the correlation between focus and two interpretations in Czech, and (ii) reasons leading to two readings of Czech VN-comparatives.

3 Analysis and discussion

3.1 Czech CN-comparatives

The statistical outputs show that sentences with CN-comparatives and VN-comparatives are ambiguous between the equality reading and the interval reading, but both comparative constructions prefer the equality reading. The crucial point is that the results from the experiment go against Dočekal's claim that CN-comparatives show a preference for the interval reading, but I agree with him that Czech CN is a focus operator that has to c-command its focus marked constituent and has to be adjacent to the focus marked constituent.

Previously, Cohen & Krifka (2014) and Geurts & Nouwen (2007) observed that comparatives associate with focus, which can lead to different truth-conditions depending on the focused element. Comparative modifiers like *more than* may focus varying portions of a sentence, as in (26).

(26) a. Ann ate more than [two]$_{FOC}$ apples.
      b. Ann ate more than [two apples]$_{FOC}$.

The sentence with comparative modifiers may have two readings depending on the part of the sentence which is focused. Geurts & Nouwen (2007) observed that (26a) implies that the number of apples that Ann ate exceeds two: 'Ann ate more than two apples, actually she ate four.' (26b) implies that the number of apples that Ann ate was exactly two: 'Ann ate more than two apples, she ate two apples, one pear, and two strawberries.' In the first case, we count how many apples Ann ate, but in the second case, we count how much of everything she ate.

I contend that what we observe in Czech is the same phenomenon as already observed by Dočekal (2017), namely that the comparative marker *než* 'than' can have the equality or interval reading. I add to this an observation that focusing of the numeral would lead to the interval reading only (the presupposed alternatives then are a set of integers $\{0, 1\}$). However, because in Slavic languages the focus operator mainly appears adjacent to the focused element, the focus on the comparative marker *než* signaling the equality reading (and marginally interval

reading) is more salient, which can explain the preference observed in the experiment. Unfortunately, the experiment was a reading task; it did not control for the use of intonation.

3.2 Czech vn-comparatives

The experimental results support Dočekal's assertion that sentences with vn-comparatives indeed lead primarily to the equality reading, but the interval reading is also possible.

At this point, I agree with Nouwen's line of argumentation that a narrow scope of vn (27a) denies a constituent, and a sentence can have the equality reading because: (i) the number is bounded, and (ii) scalar implicatures are present; therefore they can be negated, and the truth condition is strengthened.

(27) John did not find more than 20 mushrooms.

    a. $\text{MAX}_D\big(\lambda y \exists x [\#x = y \wedge \text{MUSHROOM}(x) \wedge \text{FIND}(\text{JOHN}, x)]\big) \leq 20$

    b. truth conditions: $\text{MAX}_D(\lambda d \,.\, \text{the number of mushrooms was } d) \leq 20$

    c. SI: $\neg\text{MAX}_D(\lambda d \,.\, \text{the number of mushrooms was } d) \leq 19$

    d. $\text{MAX}_D(\lambda d \,.\, \text{the number of mushrooms was } d) = 20$

The wide scope of vn (28a) denies the whole proposition and the equality interpretation cannot occur because (i) the number is unbounded, and (ii) truth conditions cannot be strengthened because no scalar implicatures arise. The wide scope of vn leads to the interval interpretation $(20, \infty)$.

(28) John did not find more than 20 mushrooms.

    a. $\neg\text{MAX}_D\big(\lambda y \exists x [\#x = y \wedge \text{MUSHROOM}(x) \wedge \text{FIND}(\text{JOHN}, x)]\big) > 20$

3.3 Discussion

The most salient reading, which Czech speakers associate with both cn- and vn-comparatives, is equality. I analyze it in §3.1 as a result of a particular focus strategy. But I did not control for the focus strategies in the experimental design. This is naturally what I will try to address in the next experiment.

The experiment design would be similar to the experiment presented in this article. Participants would judge whether sentences fit a context. The experiment would test only Czech cn-comparatives because vn-comparatives did not involve focus alternatives. In order to investigate whether focused comparatives indeed

influence the preferred reading, participants would read aloud the tested sentences. The utterances would be recorded and digitized. The appropriate comparison of focused CN-comparatives and unfocused CN-comparatives could clarify the issue of focus and the two possible interpretations of Czech CN-comparatives.[13]

4 Conclusion

I investigated Czech negated comparatives compared with English negated comparatives. I started with the observation that English negated comparatives lead to two interpretations with respect to the type of negation, i.e., the preferred interval reading in the case of VN and the equality reading in the case of CN (Nouwen 2008).

I experimentally tested Czech negated comparatives. Although the experiment to some extent supports Dočekal's observation (Dočekal 2017), it also adds some interesting twists. The experimental results show that both Czech CN-/VN-comparatives are ambiguous between an equality reading and an interval reading, although they strongly prefer the equality interpretation.

Following an approach that unifies scalar alternatives and focus alternatives into one group and claims that both types of alternatives are the same (Katzir 2007; Fox & Katzir 2011; Fox & Spector 2018), I argue that both Czech CN-comparatives and English CN-comparatives result in the preference for equality reading because they activate alternatives.

Czech VN-comparatives behave in the same way as English VN-comparatives: a particular reading depends on whether negation takes wide scope or a narrow scope with respect to a proposition. The narrow scope of VN results in a preference for the equality reading in Czech, whereas the wide scope of VN leads to the interval reading.

---

[13]Thanks to an anonymous reviewer for raising this point.

Abbreviations

1	first person	N	interval reading
ACC	accusative	NEG	negation, negative
CN	constituent negation	NOM	nominative
DAT	dative	PL	plural
E	equality reading	PRS	present
FOC	focus	PST	past
GEN	genitive	SG	singular
INF	infinitive	SI	scalar implicature
INS	instrumental	VN	verbal negation
LOC	locative		

Acknowledgements

References

Büring, Daniel & Katharina Hartmann. 2001. The syntax and semantics of focus-sensitive particles in German. *Natural Language & Linguistic Theory* 19(2). 229–281. DOI:10.1023/A:1010653115493

Cohen, Ariel & Manfred Krifka. 2014. Superlative quantifiers and meta-speech acts. *Linguistics and Philosophy* 37(1). 41–90. DOI:10.1007/s10988-014-9144-x

Dočekal, Mojmír. 2013. What do we count with numerals? Semantic analysis of Czech group-denoting and kind-denoting NPs. In Uwe Junghanns, Dorothee Fehrmann, Denisa Lenertová & Hagen Pitsch (eds.), *Formal Description of Slavic Languages: The ninth conference. Proceedings of FDSL 9, Göttingen 2011*, 87–105. Frankfurt am Main: Peter Lang.

Dočekal, Mojmír. 2017. Upper bounded and un-bounded 'no more'. *Acta Linguistica Academica* 64(2). 213–231. DOI:10.1556/2062.2017.64.2.3

Dočekal, Mojmír & Marcin Wągiel. 2018. Event and degree numerals: Evidence from Czech. In Denisa Lenertová, Roland Meyer, Radek Šimík & Luka Szucsich (eds.), *Advances in formal Slavic linguistics 2016*, 77–108. Berlin: Language Science Press. DOI:10.5281/zenodo.2554021

Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586. DOI:10.1007/s10988-006-9004-4

Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19(1). 87–107. DOI:10.1007/s11050-010-9065-3

Fox, Danny & Benjamin Spector. 2018. Economy and embedded exhaustification. *Natural Language Semantics* 26(1). 1–50. DOI:10.1007/s11050-017-9139-6

Geurts, Bart & Rick Nouwen. 2007. *At least* et al.: The semantics of scalar modifiers. *Language* 83(3). 533–559. https://www.jstor.org/stable/40070901.

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

Horn, Laurence. 1996. Presupposition and implicature. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 299–319. Oxford: Blackwell.

Horn, Laurence. 2013. Implicature. In Gillian Russell & Delia Graff Fara (eds.), *Routledge companion to philosophy of language*, 75–88. New York: Routledge.

Jasinskaja, Katja. 2016. Information structure in Slavic. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 709–732. Oxford: Oxford University Press.

Kadmon, Nirit. 2001. *Formal pragmatics: Semantics, pragmatics, presupposition, and focus.* Oxford: Blackwell.

Katzir, Roni. 2007. Structurally-defined alternatives. *Linguistics and Philosophy* 30(6). 669–690. DOI:10.1007/s10988-008-9029-y

Krifka, Manfred. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* 25(3–4). 209–257. DOI:10.3765/salt.v0i0.2462

Krifka, Manfred. 1999. At least some determiners aren't determiners. In Ken Turner (ed.), *The semantics/pragmatics interface from different points of view*, 257–291. Oxford: Elsevier.

Larson, Richard K. 1988. Scope and comparatives. *Linguistics and Philosophy* 11(1). 1–26. DOI:10.1007/BF00635755

Nouwen, Rick. 2008. Upper-bounded *no more:* the exhaustive interpretation of non-strict comparison. *Natural Language Semantics* 16(4). 271–295. DOI:10.1007/s11050-008-9034-2

Rooth, Mats. 1985. *Association with focus.* Amherst, MA: University of Massachusetts. (Doctoral dissertation). DOI:1813/28568

Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1(1). 75–116. https://www.jstor.org/stable/23748778.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3). 367–391.   DOI:10.1023/B:LING.0000023378.71748.db

Wągiel, Marcin. 2014. Boys, girls, and scissors: A semantic analysis of Polish NPs headed by the numeral *dwoje*. In Ludmila Veselovská & Markéta Janebová (eds.), *Nominal structures: All in complex DPs*, 69–84. Olomouc: Palacký University.

Wągiel, Marcin. 2015. Sums, groups, genders, and Polish numerals. 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*, 495–513. Frankfurt am Main: Peter Lang.

Chapter 12

Syntax predicts prosody: Multi-purpose morphemes in Serbo-Croatian

Marko Simonović

University of Nova Gorica

Boban Arsenijević

University of Graz

We consider four Serbo-Croatian suffixes which appear in various structural positions and display different prosodic behaviour in these positions. Such suffixes allow us to establish the effects of the structural context on prosody by constructing minimal pairs between e.g. derivation and inflection. All four suffixes are shown to fit the generalization that derivational morphology is more accented than inflectional morphology. We propose a formal explanation and discuss the functional benefits of a surface differentiation between the two uses.

Keywords: prosody, derivation, inflection, multi-purpose morphemes, Serbo-Croatian, Optimality Theory

1 Introduction

Morphemes in different structural positions have different phonological properties. This insight has been formalised within various frameworks in both phonology and syntax. In phonology, roots have long been recognised as allowing more phonological contrast than affixes, an observation which has been formalised within Optimality Theory as the constraint family Root Faithfulness (see e.g. McCarthy & Prince 1993 and Beckman 1997 for a discussion of roots as one of the privileged positions in phonology). In a related model, Revithiadou (1999) presents evidence for the prosodic dominance of syntactic heads (stems and

derivational affixes can be heads, but inflectional affixes cannot). Several accounts couched in Distributed Morphology (Halle & Marantz 1993; 1994) deal with prosodic asymmetries of this kind. Doner (2017) argues that Spanish suffixes that express phi-features behave as prosodic adjuncts, which excludes them from the domain of stress assignment under certain circumstances. Marvin (2002) presents evidence for prosodic behaviour that is a function of syntactic phasehood.

To our knowledge, few analyses along these lines have been proposed of cases where the same morpheme appears in different environments, and virtually no analysis targets the same affix in its derivational and inflectional uses. Data involving the same affix are crucial because they constitute minimal pairs in which the only difference is the structural position of the affix. Such minimal pairs are the only type of evidence immune to alternative accounts that explain asymmetries between different structural positions as results of accident or functional pressures on lexicalisation, which do not need to be recorded in the grammar. In other words, while there are formal accounts of phonological asymmetries between inflection and derivation (e.g. McCarthy & Prince 1993; Beckman 1997; Revithiadou 1999) the fact that the derivational suffix X is accented whereas the inflectional suffix Y is not does not immediately strike researchers as a fact in need of a grammatical explanation. However, in cases where the same affix gets different prosodic treatment in different structural positions, we can be sure to see a grammatical mechanism at work.

The first analysis of the same morpheme in different structural environments we are aware of is a cursory discussion of nominalising *-ost* in Slovenian in Marvin (2002). This suffix seems to combine with the adjective *mlad* 'young' in two different nominalisations: in *mlad-óst* 'youth' and in *mlád-ost* 'youngness' (several other pairs of *-ost*-nominalisations are listed). For Marvin, the relevant distinction is that between root nominalisations and deadjectival nominalisations. In the root nominalisation *mlad-óst*, there is no adjectival head between the root and the nominal head *-ost*, so the root and the suffix are in the same syntactic phase. As a consequence, the suffix imposes its prosody (*mlad-óst*). In the deadjectival nominalisation *mlád-ost*, there is a (silent) adjectival head between the root and the nominal head *-ost*, which causes a separate spell-out of *mlád*. The suffix arrives "too late" to affect the stress pattern of the whole, so the resulting stress pattern is that in *mlád-ost*. Arsenijević & Simonović (2013) analyse the Serbo-Croatian cognate of the same suffix using the lexicalist mechanism of Lexical Conservatism (first proposed in Steriade 1997), a constraint which enforces copying the prosody of the base in all paradigm members. Lexical Conservatism has no influence on non-paradigm members. To stay with the same Slovenian

example, *mlád-ost* counts as part of the paradigm of the adjective *mlád* (based on semantic transparency and the fact that the pattern is productive). As such, *mlád-ost* copies the prosody of *mlád* due to the pressure of Lexical Conservatism. *Mlad-óst*, on the other hand, is a separate lexical item and its stress only depends on the general constraints (which in this case seem to enforce the faithfulness to the stress specification of the affix *-ost*). Simonović & Arsenijević (2014) present an analysis along the same lines of the Serbo-Croatian deverbal nominalisations, to which we will turn in §4.1.

Marvin (2002) and Arsenijević & Simonović (2013) use different formal tools, but both account for the targeted data sets. Given this background, our main goal is to expand the data set. We achieve this by discussing the influence of the structural position on prosody in cases of maximal MULTI-FUNCTIONALITY: those cases where one of the structures in which the affix surfaces is clearly inflection whereas the other one is derivation. Furthermore, we observe various cases of such multi-functionality within the same language in order to establish generalisations which hold across morphemes.

The main focus of this paper lies on the prosodic behaviour of affixes which occur both in inflection and derivation and have different prosodic effects in the two domains. To the best of our knowledge, while multi-functional affixes are relatively frequent, so far, this kind of systematic dichotomy at the level of prosody has only been attested in Serbo-Croatian. This is why the empirical focus of this paper will be on data from this language.

The rest of this paper is organised as follows. §2 addresses the issue of identifying affixes which can surface in both inflection and derivation as well as predictions concerning the prosodic effects of such affixes. Based on the existing literature, the prediction is put forward that the derivational uses of the affixes should go hand in hand with more accentedness, whereas the inflectional uses should be characterised by less accentedness. In §3 we present the key features of Serbo-Croatian prosody and its notation. We then list four ways in which the prosodic pattern of the base can be influenced by an affix in Serbo-Croatian. §4 presents a detailed overview of four Serbo-Croatian affixes which appear in both inflection and derivation. We keep track of their SURFACE DISTINGUISHABILITY and ACCENTEDNESS ASYMMETRIES in the two contexts. In §5 we identify the common patterns in the data presented in §4, observing that the prosodic pattern in the derivations seems to be the same at least across the suffixes performing the same function. In §6 we consider the theoretical consequences of the observed asymmetries. §7 concludes this paper.

2 Multi-functional affixes

Cases of the same affix appearing in both derivation and inflection are not hard to come by. Below we quote examples from English and Italian. Both English *-ed*$_A$ and Italian *-uto*$_A$ appear as regular past/passive participle endings when combined with verbs, but also as adjectivisers when combined with nouns.

(1) English *-ed*$_A$
    a. fear fear-ed   (Inflection)
    b. beard beard-ed   (Derivation)

(2) Italian *-uto*$_A$
    a. tem-ere tem-uto   (Inflection)
       fear-INF fear-PASS.PTCP
       'to fear' 'feared'
    b. barba   barb-uto   (Derivation)
       beard   beard-uto

Serbo-Croatian has several suffixes which behave in a similar way. Moreover, Serbo-Croatian pairs of this type are often characterised by surface distinguishability by means of prosody. In (3) and (4) we show how Serbo-Croatian *-at*$_A$ and *-an*$_A$ appear in different constellations.

(3) Serbo-Croatian *-at*$_A$
    a. prìzna-ti   prìzna-at   (Inflection)
       recognise-INF recognise-PASS.PTCP
       'recognise'   'recognised'
    b. pȑs-a pȑs-at   (Derivation)
       bust   bust-at
       'bust' 'busty'

(4) Serbo-Croatian *-an*$_A$
    a. pòsla-ti pòsla-an   (Inflection)
       send-INF send-PASS.PTCP
       'to send' 'sent'
    b. gȉps-a   gȉps-an   (Derivation)
       plaster-GEN plaster-an
       'plaster'   'made of plaster'

We will discuss the details of the prosodic representation of Serbo-Croatian in §3. For the moment, suffice it to say that both suffixes systematically display different prosodic patterns in the two uses and, previewing our findings, that the derivational endings are more accented. This asymmetry is in the same direction as those observed in Marvin (2002) and Arsenijević & Simonović (2013), and it also matches the cross-linguistic tendencies that will be discussed below.

Before moving on, we briefly address one potential objection to this presentation of the data. There is an obvious alternative: accidental homonymy between unrelated affixes. So English *-ed*, Italian *-uto*, as well as Serbo-Croatian *-at* and *-an* may be not a single affix but pairs of unrelated affixes which happen to have the same form or, more precisely, the same segmental content. Our arguments against this view can be summarised as follows:

- Whether inflection or derivation, the category of the word resulting from affixation is the same (adjective in all the cases discussed above). On the accidental homonymy analysis, this would be another accident.

- Most of these suffixes are old in both uses, without a diachronic tendency to phonologically split into two different suffixes – which is what would be expected were they different items.

- In Serbo-Croatian, the two uses of the same affix are systematically distinguished by different prosodic patterns, as discussed in this paper.

- Finally, it would be quite surprising for a Germanic, a Romance and a Slavic language to have an accidental homonymy between the suffixes with exactly the same purposes: the passive participle and an adjectival suffix.

Once we accept that the derivational and inflectional affixes in question are indeed the same affix, prosodic asymmetry is predicted to exist in some language. This follows from the cross-linguistic generalisation that derivational affixes are phonologically more prominent than inflectional affixes (see e.g. Beckman 1997; Revithiadou 1999). Based on facts from several languages from different language families, Revithiadou (1999) proposes two constraints that favour prosodic prominence of derivational affixes. HEADFAITH is a faithfulness constraint which protects lexical prominence of syntactic heads (derivational affixes and stems are argued to be syntactic heads, unlike inflectional affixes). HEADSTRESS is a markedness constraint that militates against stress on non-heads. This constraint is violated whenever inflectional affixes are stressed.

The observed asymmetry predicts that there should exist two types of languages:

- languages such as Italian in which both derivational and inflectional affixes
  can be stressed (or otherwise prosodically strong) and

- languages such as English or Dutch in which derivational affixes can be
  prosodically prominent but inflectional affixes cannot.

To our knowledge, there is no language which is a mirror image of English and
Dutch. In such a language, inflectional suffixes would be more stressed than
derivational affixes.

In Serbo-Croatian, both derivational and inflectional affixes can be either ac-
cented or accentless (but accented inflectional suffixes are becoming rare, see
Simonović & Kager 2020 [this volume]). Prosodic prominence in Serbo-Croatian
involves stress, tone and vowel length and the lexical sponsor of the surface
prosody is not easily determined. This is why we first turn to Serbo-Croatian
prosody and its representation in the following section.

3 Standard Serbo-Croatian prosody

Serbo-Croatian is traditionally classified as a PITCH-ACCENT system in which the
distribution of stress is predictable from that of HIGH TONE (Inkelas & Zec 1988;
Zec 1999). Every prosodic word is characterised by a single TONAL ACCENT headed
by the single stressed syllable of the word. Classical descriptions distinguish be-
tween falling and rising tonal accents. In the FALLING ACCENTS, stress and high
tone co-occur on the same syllable, which is typically the first syllable of the
word. Depending on the length of the stressed syllable, there are long falling and
short falling accents (in *lâđa* 'boat' and *krăđa* 'theft', respectively). The RISING AC-
CENTS are traditionally analysed as spans of two adjacent syllables which both
have a high tone, whereas only the first syllable also carries stress (but see Zsiga
& Zec 2013 for arguments that in some varieties the first syllable of the rising ac-
cents only carries stress but no high tone). The rising accents can also be long or
short, depending on the length of the stressed syllable (as in *báka* 'grandmother'
and *màča* 'sword.GEN', respectively).

Most accounts of Serbo-Croatian prosody share some central assumptions. The
rising accents are generally assumed to have a lexical sponsor in the rightmost
syllable of the span, which automatically spreads onto the preceding syllable: the
rising accent in *màča* 'sword.GEN' then derives from /maʧaH/. This spreading
accounts for the fact that rising accents do not occur in monosyllables, where
only falling accents are possible. Falling accents are the realisation of high tones
that cannot spread to the left: those on the initial syllables. This also accounts for

the fact that falling accents are restricted to initial syllables. The falling accents are assumed to get assigned to the initial syllable as a default in the absence of lexically specified prosody: /novine/ will become *nȍvine* 'newspaper' if no high tone gets assigned in the derivation.

An obvious disadvantage of the traditional diacritics is that they do not mark the second part of the rising accent, so the reader needs to memorise the diacritics and 'imagine' a high tone after every rising-accent diacritic. A disadvantage of using an IPA notation is that it has some overlap with the traditional notation, which confuses those Serbo-Croatian speakers who are used to the traditional notation. This is why we use a different, more transparent notation in this paper: high tone is marked by capitalisation and vowel length by doubling. Stress is not marked, as it predictably falls on the leftmost high-toned (= capitalised) syllable.

Table 1: Traditional diacritics and the notation used here

	monosyllables	polysyllables
long falling	grâd = grAAd 'town'	lâđa = lAAđa 'boat'
short falling	grȁd = grAd 'hail'	krȁđa = krAđa 'theft'
long rising	/	báka = bAAkA 'grandmother'
short rising	/	màča = mAčA 'sword.GEN'

Serbo-Croatian suffixes display varying behaviour with respect to prosody. Table 2 illustrates four suffixes interacting with bases which have a rising accent, i.e. they have an underlying high tone in their representation. The behaviour of these four affixes can be described as:

- ACCENT-BEARING (*-ana* deletes the accent of the base and imposes its own),

- ACCENT-ATTRACTING (*-iji* moves the accent closer to itself),

- ACCENT-NEUTRAL (*-oost* does not change the prosody of the base) and

- ACCENT-ERASING (with *-aaj*, the high tone of the base is deleted, but no other high tone is added by the suffix).

Note that in the fourth example, the accent-erasing suffix leads to default prosody i.e. to the initial short falling accent.

Table 2: Prosodic effect of suffixes in Serbo-Croatian

Base	šEćEr 'sugar'	rAnjIv 'vulnerable'	rAnjIv 'vulnerable'	pOkUš-ati 'try.INF'
Derivation	šećer-AnA 'sugar factory'	ranjIv-Ijii 'more vulnerable'	rAnjIv-oost 'vulnerability'	pOkuš-aaj 'attempt'
Behaviour	accent-bearing	accent-attracting	accent-neutral	accent-erasing

4 Case studies

In this section we present a detailed overview of the prosodic behaviour of four Serbo-Croatian affixes which appear both in inflection and derivation: the nominal -$VVje_N$ and the adjectival -en_A, -an_A, -at_A.

Throughout the discussion of the affixes, we will keep track of two aspects of their behaviour. One is SURFACE DISTINGUISHABILITY: to what extent are the two uses of the suffix distinguishable in the surface form? The other concerns the formal status of the relevant uses of the suffix in relation to protypical derivation or inflection. In each case we consider the PROSODIC DOMINANCE (accentedness), PRODUCTIVITY and SEMANTIC TRANSPARENCY of the suffixes. In line with the typological generalisations and what we found in Serbo-Croatian in Arsenijević & Simonović (2013) and Simonović & Arsenijević (2014), the general expectation is that inflectional uses should be prosodically non-dominant, productive and semantically transparent, whereas the derivational uses should be more prosodically dominant, less productive and less semantically transparent.

In the literature on Serbo-Croatian, the affixes under scrutiny here, especially the adjectival -en_A, -an_A and -at_A, are analysed as different morphemes in their derivational and inflectional uses. Since we are proposing a new unified analysis of the suffixes in question, we limit our attention to those cases where the presence of the suffix is unquestionable. We therefore restrict our corpus to the cases of concatenation of morphemes without any irregular over- or underapplication of phonological or morphological processes e.g. unexpected intervening segments or consonant mutation. We do make an exception regarding one process, as it is fully productive in at least one of the contexts, that of so-called *iotation*. Iotation is a consonant palatalisation, typically before a *j*, e.g. in /sxʋati+en/ which surfaces as *sxʋaȶ͡ɕen* 'understood' via the intermediate /sxʋatjen/, as *tj* is palatialised to *ȶ͡ɕ*.

In the rest of the paper, we distinguish between derivational and inflectional versions using the following notation: *affix*$_{DERIV}$ will be used for the derivational

versions, whereas *affix*_{INFLECT} will be used for the inflectional versions of the same affixes. In §4.1 the nominal suffix *-VVje*_N is discussed, whereas in §4.2 the adjectival suffixes *-en*_A, *-an*_A and *-at*_A are in focus.

4.1 Case study 1: *-VVje*$_N$

This case study summarises some of the findings presented in Simonović & Arsenijević (2014), placing them in the context of this paper. The suffix *-VVje*$_N$ consists of a vowel length that gets realised on the last syllable of the base and the segmental content *-je*. It combines with verbal bases, as well as with phrasal units, mainly VPs, N(um)Ps and PPs. In this paper, we only focus on its application in the verbal domain, where it derives neat minimal pairs depending on the aspectual properties of the base.

When combined with imperfective verbs, the suffix *-VVje*$_N$ derives event-denoting deverbal nominalizations. This pattern is fully productive (the suffix combines with all imperfective verbs), typically semantically transparent (fully compositional), and the suffixation does not affect the prosody of the base. This type of derivation hence shows a number of properties typical for inflection and the suffix can therefore be represented as *-VVje*$_{N.INFLECT}$ (Table 3).

Table 3: *-VVje*$_{N.INFLECT}$

unapređIIvAti	'promote.IPFV.INF'	unapređIIvAAnje	'promoting.IPFV'
sAAdIti	'plant.IPFV.INF'	sAAđEEnje	'planting.IPFV'
čEkati	'wait.IPFV.INF'	čEkaanje	'waiting.IPFV'
psOvAti	'swear.IPFV.INF'	psOvAAnje	'swearing.IPFV'

When combined with perfective verbs, suffix *-VVje*$_N$ derives factitive nominalizations. The derivation is idiosyncratic, barely productive, frequently lexicalized and imposes its own prosodic shape. It is hence much closer to prototypical derivation and the suffix can be represented as *-VVje*$_{N.DERIV}$ (Table 4, see also Simonović & Arsenijević 2018).

Summarising the prosodic bahaviour of *-VVje*$_N$ when combined with vebal participles, we can establish two patterns with a clear divide between them: *-VVje*$_{N.INFLECT}$ behaves as accent-neutral as illustrated in Table 3 above, whereas the nominalisations with *-VVje*$_{N.DERIV}$ predictably have penult stress. In other words, the *-VVje*$_{N.DERIV}$ behaves as accent-attracting.

There is full surface distinguishability between the productive and transparent pattern, which shows properties of inflectional morphology, on the one hand and

Table 4: *-VVje*_{N.DERIV}

unaprEEdIti	'promote.PFV.INF'	unapređEEnjE	'promotion'
zasAAdIti	'plant.PFV.INF'	*zasađEEnjE	
sAčEkati	'wait.PFV.INF'	*sačekAAnjE	
opsOvAti	'swear.PFV.INF'	*opsovAAnje'	

the idiosyncratic pattern, which acts like prototypical derivation, on the other hand. This asymmetry fits the typological generalisation that the properties of inflection are more likely to coincide with prosodic inactivity, and the properties of derivation go hand in hand with prosodically active behavior.

4.2 Case study 2: *-en*_A, *-an*_A, *-at*_A

Serbo-Croatian has a set of three different suffixes which are equivalent in relevant respects to the English *-ed* and the Italian *-uto* discussed in §2, i.e. which are used both for the passive participle and for the derivation of adjectives. These suffixes are *-en*_A, *-an*_A and *-at*_A. We applied the following selection criteria in assembling our data set of forms which contain *-en*_A, *-an*_A and *-at*_A. Regarding the passive participle use, we only included in our corpus verbs in which *-en*_A, *-an*_A and *-at*_A can clearly be reconstructed as the PASS.PTCP morphemes (app. 90% of all the verbs). A detailed overview of verbal paradigms with prosodic information can be found in Klaić (2013).

Our corpus of non-participial adjectives derived by *-en*_A, *-an*_A and *-at*_A was assembled using the Reverse dictionary of the Serbian language (Nikolić 2000) and various available descriptions of Serbo-Croatian (Babić 2002; Stevanović 1979; Barić et al. 1995). As explained at the beginning of this section, we restricted our corpus to the clearest cases of the use of the suffixes in question. Only words which have a clear structure stem+*en*_A/*an*_A/*at*_A were included. Specifically, words with a more complex suffix structure (e.g. *papir*_N+*n*_A+*at*_A 'made of paper'), and words with stems synchronically lacking a semantic relation to the derivation (e.g. *iskr*_N+*en*_A 'honest', synchronically not related to *iskra* 'sparkle') were excluded from the corpus. Additionally, words with stem modifications other than iotation (e.g. *stamben* 'residential' which is clearly related to *stan* 'apartment') were excluded as well.

For the prosodic specification of the bases and results of suffixation, we have consulted the prosodic intuitions of modern Serbo-Croatian speakers. The full set of words with the derivational versions of *-en*_A, *-an*_A and *-at*_A can be found

in the Appendix. The inflectional versions of these suffixes are productive (especially for the first two), which is why we worked with verb classes rather than a corpus. We discuss each of the three suffixes in a separate subsection, including a quantitative overview of their prosodic behaviour.

4.2.1 *-an*

Adjectives in *-an*$_{\text{DERIV}}$ are mostly denominal and have the interpretation of being made of the material denoted by the base noun, or having a property related to its semantics to a large extent. Their distribution in the corpus is shown in Table 5.

Table 5: Adjectives in *-an*$_{\text{DERIV}}$ in the corpus

Base type	N	V	A	Phrase
Example	Ulj-An 'made of oil' (UUlj-e 'oil')	pIj-An 'drunk' (pI-ti 'drink')	mEk-an 'soft' (mEk 'soft')	/ / /
Prosodic behaviour	Attracting	Attracting	Attracting Neutral	/
Number of items	33	1	2	0

-an$_{\text{DERIV}}$ virtually always surfaces as the second part of the rising accent (the only exception being the adjective *mEk-an* illustrated in Table 5). It is therefore prosodically active and overwrites the prosody of the base. *-an*$_{\text{DERIV}}$ is only attested with monosyllabic bases.

-an$_{\text{INFLECT}}$ shows a prosodically inactive behavior. Without exceptions, passive participles in *-an*$_{\text{INFLECT}}$ have a prosodic pattern that exists elsewhere in the verbal paradigm, as illustrated in Table 6, where the other form with the same prosodic pattern is represented in bold. In other words, it does not affect the prosody of the base, and therefore we classify it as unaccented (i.e. neutral).

Even though they are segmentally identical, the two uses of *-an*, the participial *-an*$_{\text{INFLECT}}$ and the denominal adjectivizer *-an*$_{\text{DERIV}}$, are surface-distinguishable:

- *-an*$_{\text{DERIV}}$ is always part of a rising span (*UljAn* 'made of oil').

- *-an*$_{\text{INFLECT}}$ is never part of a rising span (*pIItaan* 'ask.IPFV.PASS.PTCP').

- *-an*$_{\text{DERIV}}$ never surfaces in a long syllable (UljAn 'made of oil').

- *-an*$_{\text{INFLECT}}$ always surfaces in a long syllable (*pIItaan* 'ask.IPFV.PASS.PTCP').

Table 6: Participles in -an_{INFLECT}

INF	PRES.1SG	PST.PTCP	PASS.PTCP	Gloss
pIItAti	**pIItaam**	pIItAo	**pIItaan**	'ask.IPFV'
zapIItiAti	**zApIItaam**	zapIItAo	**zApIItaan**	'ask.PFV'
čItAti	čItAAm	**čItao**	čItaan	'read.IPFV'
pročItAti	pročItAAm	**prOčitao**	**prOčitaan**	'read.PFV'

The prosodic behaviour of the two uses of the suffix *-an* thus fully complies with the generalizations in §2. Suffix *-an*_{DERIV} is prosodically active: it always imposes the same pattern, overwriting the prosody (including the vowel length) of the stem (*UljAn* 'made of oil' vs *UUlje* 'oil'). Suffix *-an*_{INFLECT} is accentless, i.e. neutral: the result of suffixation bears a prosodic pattern which has already been present in the paradigm of the base.

4.2.2 *-at*

*-at*_{DERIV} derives denominal adjectives with the structure base_N-*at*, and the interpretation of having the denotation of the base noun to a large extent. It has the following quantitative distribution in the corpus.

Table 7: Adjectives in -at_{DERIV} in the corpus

Base type	N	V	A	Phrase
Example	zUb-At	/	/	/
	'toothy'	/	/	/
	(zUUb 'tooth')	/	/	/
Prosodic behaviour	Attracting	/	/	/
Number of items	17	0	0	0

The use of *-at*_{DERIV} is additionally constrained by one phonotactic and one semantic restriction on bases: the bases are strictly monosyllabic, and all denote body parts. The participial *-at*_{INFLECT} always has a prosodic pattern that exists elsewhere in the verbal paradigm (typically in the past participle).

This situation leads to the same generalisation as with the suffix *-an*. The two uses of the same suffix, *-at*_{DERIV} and *-at*_{INFLECT} are surface-distinguishable:

Table 8: Participles in *-at*_{INFLECT}

INF	PRES.1SG	PST.PTCP	PASS.PTCP	Gloss
prepOznAti	prepOznAAm	**prEpoznao**	**prEpoznaat**	'recognise'
prOdAti	prOdAAm	**prOdao**	**prOdaat**	'sell'
porAvnAti	porAvnAAm	**pOravnao**	**pOravnaat**	'flatten'

- *-at*_{DERIV} is always part of a rising span (*zUbAt* 'toothy'), while *-at*_{INFLECT} never is (*prOdaat* 'sell.PASS.PTCP');

- *-at*_{INFLECT} is always part of a long syllable (prOdaat 'sell.PASS.PTCP'), whereas *-at*_{DERIV} always surfaces in a short syllable (*zUbAt* 'toothy').

All in all, the prosodic behaviour of *-at* is as expected: *-at*_{INFLECT} is accentless and *-at*_{DERIV} always imposes the same prosodic pattern, deleting the prosody of the stem (e.g. removing the vowel length of *zUUb* 'tooth' in *zUbAt* 'toothy').

4.2.3 *-en*

The two adjectival suffixes that we have considered so far display prosodic behaviour that neatly fits the tendencies outlined in §2. The situation is somewhat less black-and-white with the suffix *-en*, which shows relatively higher productivity in derivation.

*-en*_{DERIV} derives adjectives from bases of different categories and yields four different prosodic patterns. With phrasal bases, the stress falls on the final syllable of the first member of the phrasal base (*jednO-cIfr-en* 'one-digit'), which indicates that the initial syllable of the second member, which heads the construction, bears a high tone. We only found one example with an adjectival base, and it is a rather unique form that is in a suppletion relation with its own base (*mAl-En* 'little' cf. the definite form *mAAl-ii* 'little'). This one example, as well as a vast majority of denominal adjectives derived by the suffix *-en*_{DERIV}, show a stress-attracting behavior of the suffix similar to that of *-an*_{DERIV} and *-at*_{DERIV}. All such cases involve a monosyllabic base. In six cases – all with polysyllabic nouns as bases – *-en*_{DERIV} shows a neutral behavior (the derived adjective has the accent pattern of the base). Tellingly, in all such cases, the stress pattern of the base is not stem-final (e.g. *IzlOžb-a* 'exhibition', *IzlOžb-en* 'related to an exhibition'), so that the accent-attracting property of the suffix would have caused a stress shift (**izlOžb-En*). Finally, in two cases *-en*_{DERIV} erases the lexical specification of the prosody of the base – hence the derived adjective receives the default

prosody (short falling initial accent). In sum, -*en*_{DERIV} displays several patterns, out of which the most frequent one is the same as that of -*an*_{DERIV} and -*at*_{DERIV}.

Table 9: Adjectives in -*en*_{DERIV} in the corpus

Base type	N	V	A	Phrase
Example	rAž-En 'made of rye' (rAAž 'rye')	/ / /	mAl-En 'little' (mAAlii 'little')	dvOsmIsl-en 'ambiguous' (dvAA smIIsla 'two senses')
Prosodic behaviour	Attracting (59) Neutral (6) Erasing (2)	/	Attracting	Pre-stressing
Number of items	67	0	1	10

This behavior suggests an interplay between syntactic and phonological factors in the assignment of prosody. On the syntactic side, there seems to exist a sensitivity to complexity (phrasal vs. simplex bases) and to categorial specifications (nouns vs. adjectives). On the phonological side, the length of the base seems to play a role. Taking a more detailed look reveals another generalisation: -*en*_{DERIV} never shifts the stress of the base to another syllable (but it can delete the H and the vowel length of the base). This is also true of all cases of -*an*_{DERIV} and -*at*_{DERIV}, simply due to the fact that these two always combine with monosyllabic stems.

Passive participles derived using -*en*_{INFLECT} always have a prosodic pattern that exists elsewhere in the verbal paradigm, as illustrated in Table 10. The only exception is formed by four 'rising' classes, where the suffix seems accent-erasing, yet without affecting vowel length. These are illustrated in Table 11.

Table 10: Participles in -*en*_{INFLECT}

INF	PRES.1SG	PST.PTCP	PASS.PTCP	Gloss
vAditi	**vAdiim**	**vAdio**	**vAđen**	'take out'
stvOrIti	**stvOriim**	stvOrIo	**stvOren**	'create'
otvOrIti	**OtvOriim**	otvOrIo	**OtvOren**	'open'
odlUUčIti	**OdlUUčiim**	odlUUčIo	**OdlUUčen**	'decide'

Table 11: Participles in *-en*_{INFLECT} in 'rising' classes

INF	PRES.1SG	PST.PTCP	PASS.PTCP	Gloss
lOmIti	lOmIIm	lOmIo	**lOmljen**	'break'
žElEti	žElIIm	žElEo	**žEljen**	'want'
trUUbIti	trUUbIIm	trUUbIo	**trUUbljen**	'honk'
žIIvEti	žIIvIIm	žIIvEo	**žIIvljen**	'live'

In accounting for this pattern, we should take into account that in most analyses *-en*_{INFLECT} attaches to verbal bases that include the theme vowel *-i-* which becomes consonantal and causes iotation of the stem-final consonant. The standard analysis is that *lomljen* corresponds to the underlying /lomi+en/ which first becomes /lomjen/. As in these verbs the underlying H seems to originate on the theme vowel *-i-*, it seems plausible for the H to disappear together with the syllabicity feature of the vowel. As a result, the form remains without an underlying H and therefore surfaces with a short falling accent: *lOmljen.* A functional gain of such a change is that distinguishability is improved: participles are kept different from denominal forms with *-en*_{DERIV}, such as *rAžEn* 'made of rye'.

Summarising the picture, *-en*_{INFLECT} and *-en*_{DERIV} are not surface-distinguishable either in polysyllables (*OtvOr-en* 'open.PASS.PTCP' vs. *OpOrb-en* 'related to opposition') or in monosyllables (*smIšljen* 'conjecture.PASS.PTCP' vs. *smIslen* 'meaningful'). At the same time, in short *en*-participles there seems to exist an active process that enforces distinguishability between them and the main pattern in denominal derivations. The prosodic behaviour of *-en*_{INFLECT} and *-en*_{DERIV} shows a partial overlap. *-en*_{INFLECT} is neutral or erasing, whereas *-en*_{DERIV} is neutral, erasing or attracting. The observed pattern still exhibits an asymmetry and still in the expected direction since the derivational suffix *-en*_{DERIV} is more accented than the inflectional *-en*_{INFLECT}.[1]

---

[1]Accent-attracting and accent-bearing *-en*_{INFLECT} are attested in some of the inherited verbal classes which lack a theme vowel, which were excluded from our corpus due to the fact that the morphological structure of the participle is opaque. The peculiar pattern which we report without analysing it here is that, at least for some speakers, perfective verbs display accent-bearing *-en*_{INFLECT}, whereas their imperfective counterparts display an accent-attracting version of *-en*_{INFLECT}. In the example below, we show the feminine version of the passive participle in order to illustrate the contrast.

(i) a. 'grind.PFV': sAmlEti sAmEljeem sAmlEo samlev-En-A
          INF   PRES.1SG   PST.PTCP PASS.PTCP

      b. 'grind.IPFV': mlEti mEljeem mlEo mlEv-En-a
              INF   PRES.1SG PST.PTCP PASS.PTCP

5 Common pattern

The prosodic behaviour of the four affixes analysed in §4 fits the generalisation that derivational affixes are more prosodically dominant than inflectional affixes. However, the commonalities seem to go even further. The dominant prosodic behavior is essentially the same for the four observed suffixes: it can be modelled by assuming an underlying representation with a high tone and the capacity to erase (parts of) the prosody of the stem. As mentioned above, this ability is somewhat more limited for the adjectival suffixes -*an*_{DERIV}, -*at*_{DERIV} and -*en*_{DERIV}, which can delete the length and tone of the base, but cannot cause a stress shift. On the other hand, -*VVje*_{DERIV} seems to leave no traces of the base prosody whatsoever, also shifting the stress position of the base. The representation of the four suffixes would then be along the following lines:

- /-VVje$_H$/ + deletion of base tone, vowel length and stress,

- /-an$_H$/ + deletion of base tone and vowel length,

- /-en$_H$/ + deletion of base tone and vowel length,

- /-at$_H$/ + deletion of base tone and vowel length.

Implementing such representations would account for the fact that e.g. the base *unaprEEdIti* /unapreedi$_H$ti/ 'promote.PFV.INF' loses both its vowel length and its H in *unapredEEnjE* /unapređeenje$_H$/ 'promotion'. Such a solution would be similar to what Marvin (2002) proposes as the underlying representation of the Slovenian nominalising suffix -*ost*, which we repeat in (5). Note that this bracket insertion amounts to overwriting the stress of the base.

(5)   -ost
      Delete stress on the stem, insert a bracket at the right edge of the stem:
      ...* * *(

In addition to these very elaborate underlying representations, we would need another mechanism that prevents these lexical prosodic specifications from surfacing in inflection. In Marvin (2002), this is spell-out which proceeds in phases, in Arsenijević & Simonović (2013), this is Lexical Conservatism which enforces the preservation of the base prosody in paradigm members. However, what both approaches seem to leave unaccounted for is the fact that all the derivational uses of different suffixes cause the same pattern: in Slovenian it is stem-final stress, in

Serbo-Croatian it is stem-final H. Given the strikingly similar prosodic behaviour in the four suffixes, a prefered explanation would be that the prosodic behavior of the suffix depends entirely on the structure in which this suffix occurs, not only in the case of inflectional uses, but also in the derivational ones. According to such an explanation, all the suffixes we consider here would underlyingly be without any prosodic prominence and they would behave as accented (Slovenian) and accent-attracting (Serbo-Croatian) due to prominence which they receive when occurring in a particular structure. In the next section, we consider what the implementation of such a solution would entail.

6 Theoretical implementation

Distributed Morphology offers an insightful way to distinguish between morphological structures. As outlined in Marvin (2002), the difference between *nóvost* 'newness' and *novóst* 'novelty' would be in their structural complexity, as shown in Figure 1. The difference in prosody would then naturally follow due to the phasal spell-out.
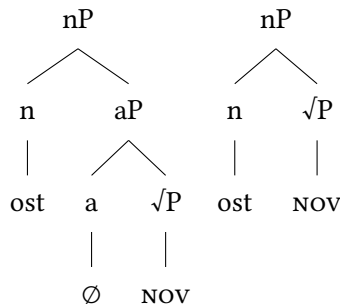
Figure 1: *nóvost* 'newness' and *novóst* 'novelty'

Two important predictions that this approach makes are:

- In root nominalisations (e.g. *novóst*), the suffix can impose (idiosyncratic) selectional requirements on bases with which it combines and the pattern therefore has limited productivity.

- The meaning of the root nominalisations cannot be compositionally derived from the meanings of their parts.

Both of these predictions seem to be born out. However, the same model makes some less desirable predictions:

- The nouns derived by means of the unstressed *-ost* are expected to have a compositional interpretation. There are, however, clear exceptions, e.g. *znán-ost* 'science' is clearly related to *znán* 'known' but its meaning cannot be derived from that of *znán* compositionally.

- The root nominalisation analysis predicts relative freedom of the stressed *-óst* in combining with roots which otherwise surface as verbs, nouns or do not surface independently. This is unfortunately not born out. Out of the very few *-ost/-óst* nouns which have roots which do not surface as independent adjectives, some have the stressed *-óst* (e.g. *krepóst* 'virtue') but others have the unstressed *-ost* (e.g. *kakóvost* 'quality').

The same problems of the root-derivation analysis extend to our Serbo-Croatian data. The three adjectival suffixes whose base category we examined show a clear tendency to select nominal bases. Finally, and most importantly, a phasal-spellout account seems unable to model the assignment of prominence by the structure and leaves us with several suspiciously similar underlying representations of different affixes.

What is necessary, then, is an alternative which would allow for the prosody of the derivational versions of the affixes to be assigned by the structure. We believe that a viable alternative can be offered by Revithiadou (1999). If the distinction between derivational and inflectional affixes is in headhood (only derivational affixes being heads), then HEADFAITH (a faithfulness constraint which protects lexical prominence of syntactic heads) is already sufficient to produce the asymmetry between inflectional and derivational uses. This would still mean that we have to stick to all the affixes having an underlyingly specified H, and the account would be as strong as those presented by Marvin (2002) or Arsenijević & Simonović (2013). However, HEADSTRESS (a markedness constraint that militates against prominence on non-heads) can get us further. In stress systems, this constraint can enforce adding epenthetic stress to a head that has no lexically sponsored prominence (e.g. in the Slovenian *nov-óst* 'novelty'). In Serbo-Croatian, this constraint can enforce the epenthesis of a high tone, in e.g. *unapređEEnjE* 'promotion'.

A final piece of the puzzle is the fact that at least in our Serbo-Croatian data set, the nominal affix *-VVje*$_N$ overrides the prosodic specification of the base more radically than the adjectival affixes *-en*$_A$, *-an*$_A$ and *-at*$_A$: the former is able to cause stress shifts with respect to the surface prosody of the base. We believe that this is a consequence of a cross-linguistic tendency for nominal content to receive more prominence than other categories, which has been discussed in the

literature under the rubric of Noun privilege (see Smith 2011 for a discussion). While constraints enforcing Noun privilege have been proposed for roots which surface as nouns, there is no reason not to extend them to nominalising affixes.

7 Conclusions

We have analysed four Serbo-Croatian affixes which occur both as derivational and as inflectional. We provided an account that mutually relates their most prominent semantic, structural and prosodic properties in a systematic way, thus supporting the view that these cases indeed manifest different uses of the same suffix rather than pairs of homonymous suffixes. In each of the four cases, we compared the inflectional and the derivational uses of the suffix, sharing the same target category, yet with differences in interpretation that can be derived from the different contexts. The prosodic patterns of the derived words confirm the initial generalization that derivational affixes are more prosodically prominent than inflectional affixes. We speculated about both functional and formal mechanisms behind this regularity. Our tentative analysis lends support to the interface model presented in Revithiadou (1999).

Abbreviations

√	root	IPFV	imperfective
1	1st person	N	noun
2	2nd person	*n*	nominal category
3	3rd person	NOM	nominative
A	adjective	PASS	passive
a	adjectival category	PCPT	participle
DERIV	derivational	PFV	perfective
GEN	genitive	PRES	present
H	high tone	PST	past
INF	infinitive	PL	plural
INFLECT	inflectional	SG	singular

Acknowledgements

References

Arsenijević, Boban & Marko Simonović. 2013. The role of syntax in stress assignment in Serbo-Croatian. In Raffaella Folli, Christina Sevdali & Robert Truswell (eds.), *Syntax and its limits*, 192–217. Oxford University Press. DOI:10.1093/acprof:oso/9780199683239.003.0011

Babić, Stjepan. 2002. *Tvorba riječi u hrvatskome književnome jeziku* [word derivation in standard croatian]. Zagreb: HAZU – Nakladni zavod Globus.

Barić, Eugenija, Mijo Lončarić, Dragica Malić, Slavko Pavešić, Mirko Peti, Vesna Zečević & Marija Znika. 1995. *Hrvatska gramatika* [croatian grammar]. Zagreb: Školska knjiga.

Beckman, Jill. 1997. Positional faithfulness, positional neutralisation and Shona vowel harmony. *Phonology* 14(1). 1–46. DOI:10.1017/S0952675797003308

Doner, Julianne. 2017. Spanish stress and lexical accent across syntactic categories. *Probus* 29(1). 233–285. DOI:10.1515/probus-2017-0001

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

Halle, Morris & Alec Marantz. 1994. Some key features of distributed morphology. In Andrew Carnie & Heidi Harley (eds.), *MIT Working Papers in Linguistics 21: Papers on phonology and morphology*, 275–288. Cambridge, MA: MITWPL.

Inkelas, Sharon & Draga Zec. 1988. Serbo-Croatian pitch accent: The interaction of tone, stress, and intonation. *Language* 64(2). 227–248. DOI:10.2307/415433

Klaić, Bratoljub. 2013. *Naglasni sustav standardnoga hrvatskoga jezika* [The standard Croatian pitch-accent system]. Božidar Smiljanić (ed.). Zagreb: Nova knjiga Rast.

Marvin, Tatjana. 2002. *Topics in the stress and syntax of words*. MIT. (Doctoral dissertation). http://mitwpl.mit.edu/catalog/marv01/.

McCarthy, John J. & Alan Prince. 1993. *Prosodic morphology: Constraint interaction and satisfaction*. Vol. 14 (Linguistics Department Faculty Publication Series). https://scholarworks.umass.edu/linguist_faculty_pubs/14.

Nikolić, Miroslav. 2000. *Obratni rečnik srpskoga jezika* [serbian reverse dictionary]. Beograd/Novi Sad: Institut za srpski jezik SANU/Matica srpska.

Revithiadou, Anthi. 1999. *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems*. Leiden: Leiden University. (Doctoral dissertation). https://rucore.libraries.rutgers.edu/rutgers-lib/38604/PDF/1/play/.

Simonović, Marko & Boban Arsenijević. 2014. Regular and honorary membership: On two kinds of deverbal nouns in Serbo-Croatian. *Lingue e Linguaggio* 13(2). 185–210. DOI:10.1418/78407

Simonović, Marko & Boban Arsenijević. 2018. The importance of not belonging: Paradigmaticity and loan nominalizations in Serbo-Croatian. *Open Linguistics* 4(1). 418–437. DOI:10.1515/opli-2018-0021

Simonović, Marko & René Kager. 2020. Serbo-Croatian is developing stem-based prosody. Why so? In Franc Marušič, Petra Mišmaš & Rok Žaucer (eds.), *Advances in formal Slavic linguistics 2017*, 305–322. Berlin: Language Science Press. DOI:10.5281/zenodo.3764867

Smith, Jennifer. 2011. Category-specific effects. In Marc van Oostendorp, Colin Ewen, Beth Hume & Keren Rice (eds.), *The Blackwell companion to phonology*, 2439–2463. Malden, MA: Wiley-Blackwell.

Steriade, Donca. 1997. Lexical conservatism. In *SICOL 1997: Linguistics in the morning calm*, 157–79. Hanshin Publishing House.

Stevanović, Mihailo. 1979. *Savremeni srpskohrvatski jezik: Gramatički sistemi i književnojezička norma* [Modern Serbo-Croatian: The systems of grammar and the standard-language norm]. Beograd: Naučna knjiga.

Zec, Draga. 1999. Footed tones and tonal feet: Rhythmic consistency in a pitch-accent language. *Phonology* 16(2). 225–264. https://www.jstor.org/stable/4420150.

Zsiga, Elizabeth C. & Draga Zec. 2013. Contextual evidence for the representation of pitch accents in Standard Serbian. *Language and Speech* 56(1). 69–104. DOI:10.1177/0023830912440792

Appendix

This appendix contains the annotated corpus material for the observed adjectival suffixes.

Table 12: *an*-adjectives with nominal bases and no segmental change

an-adjective		Base	
brOnz-An	'made of bronze'	brOOnz-a	'bronze'
zEmlj-An	'made of soil'	zEmlj-A	'soil'
grOžđ-An	'made of grape'	grOOžđ-e	'grape'
Ulj-An	'made of oil'	UUlj-e	'oil'
cIgl-An	'made of brick'	cIIgl-A	'brick'
tAft-An	'made of taffeta'	tAft	'taffeta'
plEh-An	'made of tin'	plEh	'tin'
gIps-An	'made of plaster'	gIps	'plaster'
cIc-An	'made of textile'	cIc	'textile'
plIš-An	'made of velvet'	plIš	'velvet'
rAž-An	'made of rye'	rAAž	'rye'
štOf-An	'made of cloth'	štOf	'cloth'
zvjEzd-An	'starry'	zvijEEzd-A	'star'

Table 13: *an*-adjectives with nominal bases and iotation

an-adjective		Base	
sUnč-An	'sunny'	sUUnc-e	'sun'
žIvč-An	'nervous'	žIIvAc	'nerve'
dAšč-An	'made of bars'	dAsk-A	'bar'
kOnč-An	'made of thread'	kOnAc	'thread'
brOnč-An	'made of bronze'	brOOnc-a	'bronze'
lAnč-An	'chain-like'	lAAnAc	'chain'
nOvč-An	'related to money'	nOvAc	'money'
pUpč-An	'umbilical'	pUpak	'belly button'
pUšč-An	'related to rifle'	pUšk-a	'rifle'
tRšč-An	'made of cane'	tRsk-a	'cane'
nEpč-An	'palatal'	nEpc-E	'palate'
brOjč-An	'made of numbers'	brOOjk-a	'number'
sRč-An	'brave'	sRc-e	'heart'
mOžd-An	'brain-related'	m(O)Ozak	'brain'
žIč-An	'made of wire'	žIc-a	'wire'
svEč-An	'celebrative'	svEEtAk	'holiday'
vOšt-An	'made of wax'	vOsak	'wax'
kOšt-An	'related to bones'	kOOst	'bone'
zUpč-An	'geary'	zUUbAc	'gear'
pjEšč-An	'made of sand'	pijEEsAk	'sand'

Table 14: *an*-adjective with a verbal base

an-adjective		Base	
pIj-An	'drunk'	pI-ti	'drink'

Table 15: *an*-adjective with an adjectival base and no segmental modifications

an-adjective		Base	
mEk-an	'soft'	mEk	'soft'

Table 16: *an*-adjectives with an adjectival base and iotation

an-adjective		Base	
mlAđ-An	'young'	mlAAd	'young'

Table 17: *at*-adjectives with nominal bases and no segmental changes

at-adjective		Base	
nOg-At	'who has big feet'	nOgA	'foot'
krAk-At	'who has big limbs'	krAAk	'limb'
Uh-At	'who has big ears'	Uho	'ear'
brAd-At	'who has a big beard'	brAAdA	'beard'
bRk-At	'who has a big mustache'	bRRk	'mustache'
rOg-At	'who has big horns'	rOOg	'horn'
glAv-At	'who has a big head'	glAAvA	'head'
gUz-At	'who has a big bottom'	gUUz	'bottom'
lEđ-At	'who has a big back'	lEEđA	'back'
plEć-At	'who has big shoulders'	plEćA	'shoulder'
nOs-At	'who has a big nose'	nOOs	'nose'
pRs-At	'who has a big chest'	pRsa	'chest'
sIs-At	'who has big tits'	sIsa	'tit'
zUb-At	'who has big teeth'	zUUb	'tooth'
krIl-At	'who has big wings'	krIIlO	'wing'
rEp-At	'who has a big tail'	rEEp	'tail'

Table 18: *at*-adjectives with iotized nominal bases

at-adjective		Base	
kOšč-At	'who has big bones'	kOska	'bones'

Table 19: *en*-adjectives with monosyllabic nominal bases and no segmental changes

en-adjective		Base	
bAkr-En	'made of copper'	bAkAr	'copper'
bOrb-En	'related to fight'	bOrb-A	'fight'
tvOrb-En	'related to making'	tvOrb-A	'making'
drUštv-En	'sociable'	drUUštv-O	'society'
dvOjb-En	'related to dilemma'	dvOjb-A	'dilemma'
glAzb-En	'related to music'	glAzb-A	'music'
hImb-En	'pretentious'	hImb-A	'pretending'
kRzn-En	'made of fur'	kRRzn-O	'fur'
kOpn-En	'related to soil'	kOpn-o	'soil'
pApr-En	'related to pepper'	pApAr	'pepper'
jEčm-En	'made of barley'	jEčAm	'barley'
Ovs-En	'related to oat'	OvAs	'oat'
Ognj-En	'made of fire'	OgAnj	'fire'
plAtn-En	'made of canvas'	plAAtn-O	'canvas'
slUžb-En	'official'	slUžb-A	'service'
stAkl-En	'made of glass'	stAkl-O	'glass'
sUkn-En	'made of cloth'	sUUkn-O	'cloth'
svOjstvE-n	'characteristic'	svOOjstv-O	'property'
vApn-En	'made of limestone'	vAApn-O	'limestone'
vAtr-En	'made of fire'	vAtr-a	'fire'
dRv-En	'made of wood'	dRv-o	'wood'
glln-En	'made of clay'	glIIn-A	'clay'
gUm-En	'made of rubber'	gUm-a	'rubber'
lAn-En	'made of flax'	lAn	'flax'
lEd-En	'related to ice'	lEEd	'ice'
slAm-En	'made of straw'	slAm-a	'straw'
lIm-En	'made of tin'	lIm	'tin'
mEd-En	'made of honey'	mEEd	'honey'
svIl-En	'made of silk'	svIIl-A	'silk'
vOd-En	'made of water'	vOd-A	'water'
vUn-En	'made of wool'	vUn-a	'wool'
pUt-En	'fleshy'	pUUt	'flesh'

Table 20: *en*-adjectives with monosyllabic nominal bases and no segmental changes (cont'd)

en-adjective		Base	
sAn-En	'related to dream'	sAn	'dream'
mIsl-En	'related to thought'	mIIsao	'thought'
pAkl-En	'related to hell'	pAkAo	'hell'
cRkv-En	'related to church'	cRRkv-a	'church'
gRl-En	'related to throat'	gRl-o	'throat'
Igl-En	'related to needle'	Igl-A	'needle'
jEtr-En	'made of liver'	jEtr-a	'liver'
rAž-En	'made of rye'	rAAž	'rye'
kAzn-En	'related to punishment'	kAzn-a	'punishment'
kIčm-En	'related to spine'	kI(I)čm-a	'spine'
svAdb-En	'related to wedding'	svAdb-A	'wedding'
Usn-En	'related to lips'	Usn-a	'lip'
zdrAvstv-En	'related to health'	zdrAvstv-O	'health'
žAlb-En	'related to complaint'	žAlb-A	'complaint'
žRtv-En	'related to sacrifice'	žRRtv-a	'sacrifice'
kAv-En-ii/kAf-En-ii	'related to coffee'	kAv-A/kAf-A	'coffee'
zOb-En	'made of oat'	zOOb	'oat'
mArv-En	'related to cattle'	mAArv-a	'cattle'
pIsm-en	'literate'	pIIsm-O	'letter'
smIsl-en	'sensible'	smIIsao	'sense'

Table 21: *en*-adjectives with polysyllabic nominal bases and no segmental changes

en-adjective		Base	
božAnstv-En	'wonderful'	božAnstv-O	'deity'
jedInstv-En	'unique'	jedIInstv-O	'unity'
dostojAnstv-En	'with dignity'	dostojAnstv-O	'dignity'
veličAnstv-En	'great'	veličAnstv-O	'greatness'
knjigovOdstv-En-ii	'related to book-keeping'	knjigovOdstv-O	'bookkeeping'
prvEnstv-En	'primary'	prvEnstv-O	'priority'
ubIstv-En	'related to murder'	ubIIstv-O	'murder'
kOsItr-en	'made of tin'	kOsItar	'tin'
mOlItv-en	'related to prayer'	mOlItv-a	'prayer'
OdrEdb-en	'specificational'	OdrEdb-a	'specification'
pOrEdb-en	'comparative'	pOrEdb-a	'comparison'
IzlOžbe-n	'exhibitional'	IzlOžb-a	'exhibition'
OpOrb-en	'oppositional'	OpOrb-a	'opposition'

Table 22: *en*-adjective with a nominal bases and iotation of the base

en-adjective		Base	
gvOzd-En	'made of iron'	gvOOžđ-e	'iron'

Table 23: *en*-adjective with an adjectival base

en-adjective		Base	
mAl-En	'small'	mAAl-ii	'small'

Table 24: *en*-adjectives with phrasal bases

en-adjective		Base	
bezAzl-En	'harmless'	bez zlA	'without evil'
dvOsmIsl-en	'ambiguous'	dvAA smIIsla	'two meanings'
bEskIčm-en	'spineless'	bez kI(I)čmee	'without spine'
lakOmIsl-en	'impetuous'	lAka mIIsao	'light thought'
bEsmIsl-en	'senseless'	bez smIIsla	'without sense'
bEspOsl-en	'idle'	bez pOslA	'without job'
zApOsl-en	'employed'	za pOslOm	'for job'
UpOsl-en	'busy'	u pOslU	'in job'
jednOcIfr-en	'single-digit'	jEdnA cIfra	'one digit'
dvOcIfr-en	'double-digit'	dvEE cIfre	'two digits'

Chapter 13

Serbo-Croatian is developing stem-based prosody. Why so?

Marko Simonović
University of Nova Gorica

René Kager
Utrecht University

Evidence is presented that Serbo-Croatian, in contrast to other Slavic languages with lexical prosody, is developing a prosodic system in which stressed non-stem material is avoided and surface stress is becoming a property of the word stem. Five case studies are shown in which stress is moving from non-stem material to stems. We analyse the general move towards stem-stressed prosody as the final step in a chain of language changes initiated by the Neo-Štokavian retractions, which were automatic and contrast-preserving, but led to a massive removal of stress from inflectional endings. We discuss the general reasons behind this language change in terms of markedness and, more specifically, the constraints proposed within Optimality Theory. We propose an analysis of the change under consideration in terms of a promotion of the markedness constraint STEMSTRESS, which requires stems to be stressed independently of lexical prominence.

Keywords: lexical prosody, stem-controlled stress, inflection, paradigm uniformity, STEMSTRESS, Serbo-Croatian

1 Introduction

Standard Serbo-Croatian, in contrast to other Slavic languages with lexical prosody (e.g. Slovenian, Bulgarian, and Russian), is moving towards a system in which stressed inflectional material is avoided and stress is becoming restricted to the word stem. In this paper, we present evidence for this change from nominal

and verbal domains and consider different possible formalisations of the driving force behind this change.

This paper deals with prosodic prominence in morphologically complex words. Prosodic prominence can be a consequence of morpheme-specific lexical prominence, of general prosodic restrictions of a language or of an interaction between the two. For instance, when different morphemes within a word have conflicting lexical specifications or when a morpheme has a specification which clashes with a general prosodic requirement of the language, the prosodic pattern of the word is a result of an interaction of different prosodic preferences. Revithiadou (1999) presents a survey of typologically different systems with lexical prosody, showing that morphemes are not all equal in their influence on the resultant prosody of the word. Two observed tendencies are important for our purposes:

- Inflectional affixes lose when competing with derivational affixes or roots, and

- In cases where there is no preference (e.g. because no part of the complex word carries lexical prominence), prominence will go to the stem.

Working in Optimality Theory (OT) (Prince & Smolensky 1993), Revithiadou (1999) uses constraints to formalise these two observations. HeadFaith is a faithfulness constraint which protects lexical prominence of syntactic heads (derivational affixes and roots are argued to be syntactic heads, unlike inflectional affixes). HeadStress is a markedness constraint that militates against stress on non-heads. This constraint is violated whenever inflectional affixes are stressed.

The asymmetries described above make a prediction concerning diachronic change. Since prosodically prominent inflectional affixes are a marked option, they are expected to get lost in the course of language history. We argue that Serbo-Croatian is a case at hand, as will be illustrated by a number of case studies in which stress moved to the stem.

OT is a convenient tool for formalising this type of change. In OT, diachronic change is analysed as a promotion of a universal constraint, which therefore gets to play a more important role in the evaluations. The new/old pairs can then serve as ranking arguments. When looking for the constraint that got promoted, we are looking for a constraint which used to be violated by old forms but is not violated by the new ones. Often more than one constraint will fit this description, since differences between candidate forms in terms of constraint violations are rarely restricted to a single constraint. This is why we need to consider all candidate constraints, striving to find the best match for the change under consideration.

As will be discussed below, such a quest will leave us with three well-established constraints which correspond to three analyses of the driving force behind the change discussed here.

Before moving on, several remarks are in order concerning aspects in which we depart from the terminology used by Revithiadou (1999). First, prosodic prominence is often termed "accent" and morphemes carrying lexical prominence are called "accented". We avoid this term here and stick to "lexical prominence" because we reserve the term "accent" for the tonal contours which surface in Serbo-Croatian (discussed in §2). Second, since we are not concerned with headedness in general but only with the distinction between stems and non-stems, in order to improve the general readability of the paper, we will not refer to heads and non-heads but to stems and non-stem material. This also means that Revithiadou's HEADFAITH will continue its life in this paper under the name of STEMFAITH, whereas HEADSTRESS will be called STEMSTRESS. STEMFAITH is violated whenever in the domain of the stem there are differences between the input and the output form, but it is insensitive to any such differences outside of the stem. STEMSTRESS is violated whenever the stem is not stressed.

Now we can turn to the three constraints on which Serbo-Croatian seems to improve by moving stress from non-stem material to stems.

Option 1: STEMFAITH. Prosody becomes (more) stem-controlled in the sense that the lexical stem prominence wins (more often) when it is in conflict with other constraints. The promoted constraint is then STEMFAITH. Being a faithfulness constraint, this constraint is not a priori violated by any surface pattern. It tolerates stressed inflectional affixes whenever there is no lexical stem prominence (or, in models which allow this, when lexical stem prosody imposes stress away from the stem). This constraint family has a long history within OT (see also Alderete 2001).

Option 2: STEMSTRESS. Stress on the stem becomes an overall requirement. The relevant constraint is STEMSTRESS, a markedness constraint that requires prosodic prominence on the stem, regardless of whether it is lexically sponsored or not. Stressed non-stem material always incurs a violation of STEMSTRESS. While this constraint is readily derivable from Revithiadou's (1999) HEADSTRESS, we are the first to use it under this name and limited to stems.

Option 3: PARADIGMUNIFORMITY. Prosody becomes more uniform in morphologically related forms (within paradigms). The relevant constraint is PARADIGM-UNIFORMITY, which compares all the paradigm members and penalizes any differences between them (Burzio 1996; Kenstowicz 1996). PARADIGMUNIFORMITY can favour stressed stems, especially in languages which have null inflectional

affixes. Under the dominance of this constraint, stressed non-stem material is only tolerated if it is stressed in the whole paradigm.

The first option is relatively easily distinguishable from the other two: if STEM-FAITH gets promoted, the change should be restricted to words in which both stems and non-stem material are lexically prominent. The result of a promotion of STEMFAITH would be that in such words, non-stem material loses its influence on the prosody of the complex form. As will be clear from our discussion in §2 and §3, there are no compelling cases of such a constellation in the data illustrating the diachronic change discussed in this paper.

An issue that arises with respect to the two remaining constraints is that, in many cases, they have the same effect. For instance, with monosyllabic stems, any promotion of STEMSTRESS will automatically mean that the winning candidates also incur fewer violations of PARADIGMUNIFORMITY. We can illustrate this using a minimal toy language Tatitotu. Tatitotu only has one stem *ta* and three inflectional suffixes: *ti*, *to* and *tu*. The paradigm of the only Tatitotu word is *táti~táto~tatú*. If STEMSTRESS becomes undominated in Tatitotu, the new paradigm will be *táti~táto~tátu*. This language change would also be compatible with an analysis in which PARADIGMUNIFORMITY got promoted, since the number of violations of PARADIGMUNIFORMITY incurred by the paradigm went from 1 to 0.

However, the opposite is not true if PARADIGMUNIFORMITY gets promoted. This does not necessarily mean that the new paradigms incur fewer violations of STEMSTRESS. Returning to our Tatitotu example, a promotion of PARADIGMUNIFORMITY could also have the effect of *táti~táto~tatú* becoming *tatí~tató~tatú*, and the number of violations of STEMSTRESS would grow from 1 to 3.

In sum, STEMSTRESS offers a more restricted formalisation of a language change than PARADIGMUNIFORMITY but faced with data compatible with both a STEMSTRESS and a PARADIGMUNIFORMITY analysis, we cannot exclude either, at least as long as we are in the domain of monosyllabic stems. However, polysyllabic stems can help us decide. Consider Tatitotu's closest relative Tatetitotu, which has the disyllabic stem *tate* and the same three suffixes: *ti*, *to* and *tu*. Tatetitotu has the following paradigm: *táteti~tatéto~tatetú*. In such a paradigm, the only form that violates STEMSTRESS is the last one and therefore this is the only form that should change if STEMSTRESS gets promoted: it should become either *tatétu* or *tátetu*. So the paradigms which are compatible with the promotion of STEMSTRESS to an undominated position are *táteti~tatéto~tatétu* and *táteti~tatéto~tátetu*. As always, these are improvements on PARADIGMUNIFORMITY as well, but PARADIGMUNIFORMITY is also violated by

the difference between the first two forms, and, if promoted to an undominated position, ParadigmUniformity would crucially also level the stress pattern in the first two forms, yielding either *táteti~táteto~tátetu*, or *tatéti~tatéto~tatétu* or *tatetí~tatetó~tatetú*. Due to the described asymmetries between the predictions of the two accounts, we will not only focus on the changes we encounter but also on those that are predicted to happen by the less restrictive account (in this case: ParadigmUniformity) but fail to happen. This line of argumentation determines the way in which the paper is structured, as discussed in the next section.

We present five case studies of language change within standard Serbo-Croatian. In each case the stem is becoming more prosodically prominent at the expense of non-stem material. We consider several possible formal accounts along the lines hinted at above, adducing new data where necessary.

The rest of this contribution is organised as follows. §2 presents case studies from classical (pitch-accent) standard Serbo-Croatian. In §3 we consider the three analyses briefly presented above. We first establish that the data are compatible with a promotion of StemStress, which also means that a ParadigmUniformity analysis cannot be excluded, as discussed above. For this reason, we introduce some additional data from both verbal and nominal domains in order to pit the two analyses against each other. Finally we come to the conclusion that ParadigmUniformity, which does not make any distinction between stems and non-stem material, cannot account for the diachronic change under consideration. In §4 we discuss an additional data set from relatively new and uncharted stress-only standard varieties, which seem to have reached the endpoint of the described change: in these varieties, non-stem material receives virtually no prosodic prominence. §5 places the change discussed here in the broader context of prosodic changes in Serbo-Croatian. In §6 the conclusions are presented.

2 Case studies

2.1 Standard Serbo-Croatian prosody

Standard Serbo-Croatian is a pitch-accent system with distinctive vowel length. Four tonal accents are traditionally distinguished: two falling and two rising. Each accent is traditionally called long or short, depending of the length of its leftmost (or only) syllable. Falling accents (in which stress and HIGH TONE (H) co-occur) are monosyllabic and restricted to word-initial syllables (e.g. long-falling accent in *grâda* 'town.GEN', and short-falling in *grȁda* 'hail.GEN' cf. **gradà*, **gradâ*). The rising accents are traditionally analysed as spans of two adjacent

syllables which both have a H, whereas only the first syllable also carries stress (so *bèžati* 'flee' stands for *bé$_H$ža$_H$ti*). More recently Zsiga & Zec (2013) have argued that the first syllable of the rising accents only carries stress (so *bèžati* 'flee' stands for *béža$_H$ti*). As for the underlying contrast, all analyses assume that every word has one or zero underlying H's. Rising accents then originate from underlying H's on non-initial syllables, which spread onto the preceding syllable (so *bèžati* 'flee' is underlying /beža$_H$ti/). If an underlying H is on the first syllable, it gets realized as a falling accent because there is no room for it to spread. This explains why falling accents can only occur on the first syllable whereas the rising accents can occur on any syllable (but cannot start on a final syllable). Finally, if no H is available, a H gets assigned to the first syllable, so underlyingly toneless words surface with a falling accent as well.

The traditional notation of Serbo-Croatian tonal accents (used in all standard works in Serbo-Croatian, e.g. Stevanović 1979 and Klaić 2013) puts diacritics on all stressed syllables (and a macron on all long unstressed syllables). This feature makes them convenient for our purposes: the reader can immediately tell the location of the stress in the word. This is why we will use this traditional notation in this article. Note, however, that the function of the diacritics in this notation bears no relation to their function in IPA.

Table 1: Diacritics used for Serbo-Croatian accents and vowel length

	rising accent	falling accent	unstressed
long	á	â	ā
short	à	ȁ	a

The pitch-accent system described above has been the only official standard ever since the beginning of the standardisation in the 19th century. It was based on the prosody of the largest and most central dialect group called Neo-Štokavian. The implementation of the pitch-accent standard has been quite different in the Neo-Štokavian dialect areas, where speakers already spoke a four-accent variety natively, and in the remaining areas, where in most cases nowadays a stress-only version of the standard is spoken.

In this section, we focus on the dynamics within the Neo-Štokavian pitch-accent system (and we turn to the stress varieties in §4). We mostly rely on the normative manual by Klaić (2013) based on a manuscript that was "over 50 years old" at the moment this edition came out. Of special value for our purposes are the editor's notes, which often compare Klaić's description to modern usage, en-

abling the construction of new vs. old pairs that we use for documenting the language change. We furthermore presented the new vs. old pairs to five native speakers and report on their recognition and evaluation of the forms.

It should be pointed out that the distinction between Neo-Štokavian and Non-Neo-Štokavian areas that we are making in this article is not the same as the one that can be found in traditional dialect descriptions. This is because we are interested in the implementation of the standard Serbo-Croatian prosody, whereas traditional dialect descriptions usually reflect the non-standard language use of the oldest and least mobile inhabitants of an area. When it comes to the standard usage, rural areas are characterised by massive exposure to the vernaculars and the standard usage of the neighbouring cities (e.g. Kapović 2004 discusses the modern Croatian situation in this respect). As a consequence, certain rural areas where traditionally Non-Neo-Štokavian dialects are spoken can be and have to be included in the Neo-Štokavian area when it comes to standardisation. A case in point is Slavonia, the area in which Bratoljub Klaić was born. The dialect of Klaić's native village, Bizovac, has an Old-Štokavian prosodic system, without the Neo-Štokavian rising accents (Klaić 2007: 19) and is as such part of the traditional Old-Štokavian Slavonian dialect. However, Bizovac is at only 18 km from the Neo-Štokavian city of Osijek and the speakers from Bizovac adopt the prosody of Osijek when they speak standard Serbo-Croatian. Since there are no big cities where the Slavonian dialect is spoken, this dialect has not led to any specific features in the implementation of the standard prosody. This is why we include this dialect area, as well as the other dialect areas which, speaking in terms of Kapović (2004), "gravitate toward" big cities in the area to which the big city itself belongs.

2.2 From *bežíte* to *bèžīte*

This change targets stressed theme vowels. Stressed theme vowels used to be possible in verbs with a stem-final short-rising accent in the sg present tense. The type is illustrated by the form *bèž-ī-m* 'I flee', where *bež-* is the stem, *-i-* the theme vowel and *-m* is the 1sg suffix. In such verbs, 1pl and 2pl present-tense forms traditionally had the stress on the theme vowel: *bež-í-mo* 'we flee' and *bež-í-te* 'you(pl) flee'. Nowadays this pattern is virtually absent from standard Serbo-Croatian and only *bèž-ī-mo* and *bèž-ī-te* is possible (editor's remark in Klaić 2013: 379). The stress pattern is now the same in the whole paradigm and the stress always falls on the stem. The old and the new paradigm are illustrated in Table 2 and Table 3. The forms which underwent change are in bold.

Additional evidence for the unacceptability of the old paradigm is provided by

Table 2: The old present-tense paradigm of *bežati* 'flee'

	SG	PL
1	bèž-ī-m	**bež-í-mo**
2	bèž-ī-š	**bež-í-te**
3	bèž-ī	bèž-ē

Table 3: The new present-tense paradigm of *bežati* 'flee'

	SG	PL
1	bèž-ī-m	**bèž-ī-mo**
2	bèž-ī-š	**bèž-ī-te**
3	bèž-ī	bèž-ē

a recent social media hype after a Serbian politician uttered *Gdě bežíte?* 'Where are you running to?'. In the ensuing parodies of his statement, modern standard speakers applied the pattern to verbs in which it was never possible (e.g. *zviždíte* 'you(PL) whistle', for actual *zvíždīte*). This points at the conclusion that the old grammar (which produced stressed theme vowels only in 1PL and 2PL present-tense forms of those verbs which have a stem-final short-rising accent in the SG present tense) is no longer available to the speakers and that the modern Serbo-Croatian grammar only derives *bèžīmo*, *bèžīte*. Forms like *zviždíte* show that speakers are also able to form a generalised usage pattern based on their experience with another variety of the same language, similar to patterns which Janda et al. (1994) discuss as types of hypercorrection. This pattern is now applicable to all verbs.

2.3 From *gradóvā* to *grȁdōvā*, from *gradòvima* to *grȁdovima*

This change targeted stressed plural augments, which were possible in a large class of regular long-vowelled monosyllabic masculine nouns (type *grâd* 'town') and a small class of disyllabic masculine nouns that have a falling accent and a stem-final unstressed syllable (type *gȍlūb* 'pigeon'). Such nouns typically take the augment *-ov-* in the plural forms. While most of the forms had and still have the stress on the stem (e.g. in *grȁd-ov-i* 'towns', where *grad* is the stem, *-ov-* is the augment and *-i* is the NOM.PL suffix), the augment was traditionally stressed in

GEN.PL (e.g. *grad-óv-ā*) and in DAT.INS.LOC.PL (e.g. *grad-òv-ima*). Klaić (2013: 19–20) acknowledges the forms like *grad-óv-ā* and *grad-òv-ima* are losing ground to stem-stressed *gràd-ōv-ā* and *gràd-ov-ima*, which leads to prosodically perfectly regular paradigms. Klaić lists the old and new paradigms as in Table 4.

Table 4: The plural paradigm of *grad* 'town'

	Old	New
Nominative/Vocative	gràd-ov-i	gràd-ov-i
Accusative	gràd-ov-e	gràd-ov-e
Genitive	**grad-óv-ā**	**gràd-ōv-ā**
Dative/Instrumental/Locative	**grad-òv-ima**	**gràd-ov-ima**

For modern-day speakers, the pattern with the stressed augment *-ov-* has a different status for the two case endings. Forms like *grad-óv-ā* are judged familiar but 'extremely archaic', whereas forms like *grad-òv-ima* are judged unfamiliar. The pattern in *-óvā* seems to have lexicalised to index archaic language use, but also generalised: it is now also possible (in parodies) in short-vowelled stems, where it was not possible before, e.g. for *rȍb* 'slave' speakers accept the parody *robóvā* 'slave.GEN.PL', for actual *ròbōvā*. As such, *-óvā* seems to have become what Janda et al. (1994) termed "hyper-archaism", comparable to the use of once 2SG *-st* for all persons in English.

2.4 From *kamènu* to *kȁmenu*, from *poglédu* to *pȍglēdu*

This change affects some dozens of polysyllables with an initial falling accent and a long final syllable in NOM.SG that refer to something inanimate, e.g. *kȁmēn* 'stone' and *pȍglēd* 'gaze'. Such nouns generally kept and keep the falling accent throughout the paradigm (e.g. in the GEN.SG and DAT.SG forms *pȍglēd-a* and *pȍglēd-u*), but traditionally the LOC.SG forms displayed a stem-final rising accent (e.g. *pogléd-u*). Klaić (2013: 30) mentions the general tendency for the locative to take up 'the dative forms' (which means that the new LOC.SG form is *pȍglēd-u*) and the editor adds that forms like *poglédu* only survived in some fixed expressions. Both paradigms are illustrated in Table 5.

As for the status of the old LOC.SG forms for modern-day speakers, the forms with a short-rising accent have a different status from those with a long-rising accent. The short-rising pattern (*kamènu*) is unfamiliar to most speakers, whereas the long-rising pattern (*poglédu*) is impossible outside fixed expressions (as

Table 5: The singular paradigms of *kamen* 'stone' and *pogled* 'gaze'

	Old	New	Old	New
Nominative/Accusative	kàmēn	kàmēn	pòglēd	pòglēd
Genitive	kàmen-a	kàmen-a	pòglēd-a	pòglēd-a
Instrumental	kàmen-om	kàmen-om	pòglēd-om	pòglēd-om
Dative	kàmen-u	kàmen-u	pòglēd-u	pòglēd-u
Locative	**kamèn-u**	**kàmen-u**	**pogléd-u**	**pòglēd-u**

pointed out by Klaić's editor as well). For instance, speakers accept the form *poglédu* in the phrase *u tom poglédu* 'in that regard', but they reject it in *u tvojem pòglēdu *poglédu* 'in your gaze'. Simonović & Arsenijević (2015b) discuss the reasons for the survival of the long-rising pattern in fixed expressions, showing that the long-rising penultimate accent pattern appears in a range of similar contexts and claiming that this pattern should be considered the default in Serbo-Croatian in forms and paradigms with overt endings.

Before we move on to the next data set, a caveat is in order regarding the productivity of the rising LOC.SG pattern in the reconstructed older stage of Serbo-Croatian. There is a long tradition of representing this pattern as once fully productive (and Klaić seems to follow this tradition). Assuming a stage of full productivity does simplify the analysis of the stress shift, this analysis can then go along the following lines. All underlyingly toneless nouns (which generally surface with falling accents) had a stress shift in combination with the LOC.SG $-u_H$, which, as its representation shows, had a H. This used to produce NOM.SG~LOC.SG alternations *kàmēn~kamènu* and *pòglēdu~poglédu*, but also *grâd~grádu* 'town'. The latter alternation is still active in modern-day Serbo-Croatian, where $-u_H$ now marks the unified dative/locative inanimate case (Simonović & Arsenijević 2015a). This idyllic picture of the older stage is unfortunately an oversimplification. As shown by Stevanović (1979: 219–220), there is no real evidence that polysyllables ever productively had a stress-shift in interaction with LOC.SG $-u_H$. Stevanović lists 35 nouns in which the stress shift is attested, pointing out that many are optional and not all of the forms are attested in the same dialect. Stevanović then goes on to point out that there are more nouns which have the same prosodic pattern but never display any stress shift (e.g. *mȉrīs* 'smell' and *jȁblān* 'black poplar'). On the other hand, monosyllabic stems have the fully productive tonal accent shift of the type *grâd~grádu* (but note that there is no stress shift in *grâd~grádu*). In sum, it seems that the shift *kàmēn~kamènu* and *pòglēdu~poglédu*

actually was a feature of a lexical class. An analysis would then be possible along the lines of Kager (2008) in which the words in this class used to have two stored allomorphs (e.g. /po$_H$gle:d~pogle:$_H$d/) and the first allomorph used to win in almost all cases because it carries the default prosody generally preferred by the markedness constraints (which is proven by the fact that it gets assigned to toneless nouns as discussed in §2.1). However, the second allomorph used to win in one case: when it also helped realise the H of the ending -u_H, which could only get realized as the second part of a rising accent. If this analysis is on the right track, then inflectional affixes with a H never had the ability to cause a stress shift except in the rare cases where two allomorphs were stored. If this is the case, the fact that the old locative forms disappeared can be accounted for as a result of a lexicon-internal simplification (/po$_H$gle:d~pogle:$_H$d/ became /po$_H$gle:d/, or even /pogle:d/), rather than as a consequence of a grammar change.

2.5 From *putòvati* to *pȕtovati*

This change seems to be restricted to standard Serbo-Croatian in Serbia. It targets approximately a dozen denominal verbs derived using the suffix *-ovati* from nouns which themselves display allomorphy in their paradigm. Traditionally, denominal verbs ending in *-ovati* displayed two possible prosodic patterns. One is extremely uniform and reserved for verbs derived from nouns which display no stem allomorphy: *rȁt~rȁta* 'war.NOM~GEN' derives *rȁtovati~rȁtujēm~rȁtovao* 'wage war.INF~PRS.1SG~PST'. The other prosodic pattern displays prosodic allomorphy and surfaces in verbs derived from nouns which themselves display prosodic allomorphy: *pût~púta* 'travel.NOM~GEN' derives *putòvati~pùtujēm~pȕtovao* 'travel.INF~PRS.1SG~PST' (for a detailed analysis see Simonović 2015). In most standard varieties in Serbia, the denominal verbs of the type *putòvati~pùtujēm~pȕtovao* are moving to the new type *pȕtovati~pùtujēm~pȕtovao*, as illustrated in Table 6.

Table 6: The paradigm of *putovati* 'travel'

	Old	New
Infinitive	**putòvati**	**pȕtovati**
Present tense (1SG)	pùtujēm	pùtujēm
Past participle	pȕtovao	pȕtovao

Simonović (2015) found that approximately one third of modern speakers from Serbia reject the old forms entirely, whereas the remaining two thirds accept both old and new forms to different extents. The new infinitive prosody is clearly copied from the past participle. However, this copying only targeted the infinitive form, which had the stress away from the stem. In other words, what did not happen is the creation of a perfectly regular paradigm *pŭtovati~pŭtujēm~pŭtovao, which would basically mean that the type putòvati~pùtujēm~pŭtovao moved to the existing type rătovati~rătujēm~rătovao. Foreshadowing our analysis from the following section, it seems like this did not happen because introducing the form pùtujēm would not improve anything from the perspective of StemStress, because the stem is already stressed in the existing form pùtujēm.

3 Choosing the right analysis: StemStress vs ParadigmUniformity

Out of the three options for a formal analysis presented in §1, StemFaith is clearly the worst match for the data presented. The case studies in §2.2, §2.3 and §2.5 show no compelling cases of lexically prominent endings. The data described in §2.4 do appear like a case of an accented ending (-u_H) that loses ground. However, all relevant data come from words with toneless stems, in which there is no lexical stem prominence to be protected by StemFaith. Moreover, as discussed in §2.4, upon closer scrutiny, there seems to be no convincing evidence that there ever was a system in which the case ending -u_H could cause a stress-shift in the stem unless supported by a stored allomorph of the stem (so essentially not violating StemFaith). In sum, -u_H never caused and never causes any violations of StemFaith. If it did, there would exist paradigms of the type illustrated below. Note that Jùpiter 'Jupiter' is a noun with a lexical H: /jupi$_H$ter/.

Table 7: The non-existent and existent paradigms of Jupiter 'Jupiter'

	Non-existent	Existent
Nominative/Accusative	Jùpiter	Jùpiter
Genitive	Jùpiter-a	Jùpiter-a
Instrumental	Jùpiter-om	Jùpiter-om
Dative	Jùpiter-u	Jùpiter-u
Locative	**Jupitèr-u**	**Jùpiter-u**

As hinted in §1 the difficult part is distinguishing between the remaining two options: ParadigmUniformity and StemStress. All four changes we have discussed so far can in principle be seen as improving the uniformity of the paradigms. This is evident from the uniform paradigms that have resulted from the changes, illustrated by Table 3, Table 4, and Table 5, in which the number of prosodic allomorphs of the stem is now 1. Even the new paradigm in Table 6 improves on ParadigmUniformity to a certain extent, since the new infinitive's prosody is shared with the past participle (see *putòvati → pǔtovati*) which makes the number of prosodic allomorphs of the stem drop from 3 to 2. However, the survival of the present tense *pùtujēm* shows that tonal paradigm uniformity is not enforced when stress is already on the stem. Indeed, based on what we have seen so far, the ParadigmUniformity constraint in charge would have to be one only considering the location of stress, while remaining blind to tonal information: ParadigmUniformity(Stress Location).

The question is then: would ParadigmUniformity(Stress Location) do the same job as StemStress? The answer is that in all the cases we have seen ParadigmUniformity(Stress Location) does indeed predict the right result and the fact that stress ends up on the stem simply follows from the fact that most paradigm cells had stressed stems already. However, ParadigmUniformity(Stress Location) seems to have a different problem: it predicts several changes that did not happen. Below we show two extremely frequent paradigms, which are also extremely stable throughout the history of standard Serbo-Croatian.

Table 8: The present-tense paradigm of *kopirati* 'copy'

	SG	PL
1	kòpīr-ā-m	kòpīr-ā-mo
2	kòpīr-ā-š	kòpīr-ā-te
3	kòpīr-ā	kopír-a-jū

An analysis based on a promotion of ParadigmUniformity(Stress Location) would predict such paradigms to regularise. One of the allomorphs should then spread to the whole paradigm, resulting either in forms like *kopír-ā-m 'I copy' (if the 3PL allomorph was to spread to other forms) or in forms like *kòpīr-a-jū 'they copy' (if the allomorph attested in all the other persons spread to 3PL). We have not found any attestation for such forms.

In sum, changes seem to only have happened if they rendered unstressed

Table 9: The singular paradigm of *delfin* 'dolphin'

Nominative	dèlfīn
Genitive/Accusative	delfín-a
Instrumental	delfín-om
Dative/Locative	delfín-u

stems stressed. Hence the driving force behind this change is a promotion of StemStress. Additional evidence comes from stress varieties of standard Serbo-Croatian presented in the next section.

4 Additional evidence: Stress varieties

Stress-only standard varieties emerged in cities outside the Neo-Štokavian area (Zagreb, Rijeka, Pula, Bor, Niš, etc.). They are virtually undescribed: all the data presented here come from our data collection in the five cities listed above. Stress-only standard varieties are generally seen as a hybrid between Neo-Štokavian and Non-Neo-Štokavian varieties. However, as shown by Simonović & Kager (2017), while indeed often combining some stress patterns from Neo-Štokavian and some from Non-Neo-Štokavian, these varieties do not simply allow both in free variation, but rather impose their own restrictions and generalisations. One such generalisation, absent from both Neo-Štokavian and Non-Neo-Štokavian dialects, is that stress always falls on the stem. This is quite surprising given the fact that most Non-Neo-Štokavian dialects have paradigms in which the stem is consistently unstressed. One such paradigm is illustrated in Table 10. (We only mark the stress because some of these dialects only have stress, and the tonal dialects may have different tonal patterns).

Table 10: The present-tense paradigm of *bežati* 'flee' in most Non-Neo-Štokavian dialects

	SG	PL
1	beˈž-i-m	beˈž-i-mo
2	beˈž-i-š	beˈž-i-te
3	beˈž-i	beˈž-e

Paradigms of the type illustrated above never made it into the stress-only standard varieties. Preserving such a paradigm in a stress variety should not present any problem in terms of prosodic constraints, because the stress varieties, unlike Neo-Štokavian, allow final stress (e.g. in *kriˈstal* cf. Neo-Štokavian *krìstāl*). PARADIGMUNIFORMITY cannot offer an account for the systematic selection of the paradigms with stem stress because the Non-Neo-Štokavian paradigm above is perfectly uniform. Again, the correct result is predicted by promoting STEMSTRESS. Stress-only standard Serbo-Croatian varieties are unique in Serbo-Croatian (and possibly also entire Slavic) for having lost any prosodic prominence on non-stem material while still having lexical prosody. In that sense, they seem to be spearheading the change to come. In the following section, we summarise what the change is and how it can be accounted for.

5 Historical context: From automatic retraction to promoting STEMSTRESS

The general move towards a system with stem stress is the final step in a chain of language changes initiated by the Neo-Štokavian retractions in the 15th century (Bethin 2006: 162–168) which created the rising accents. This first step amounted to an automatic leftward tone spreading from all non-initial syllables, thereby removing (a) all cases of final stress and (b) a huge portion of stress from inflectional endings.

Step 1: *bež-ī-tĕ* → *bež-í-te* (tone/stress shift + removal of final stress)

The remaining portion of stress on non-stem material was targeted by the changes discussed here.

Step 2: *bež-í-te* → *bèž-ī-te* (complete removal of non-stem stress)

There are several possible accelerators of step 2. Dialect contact played a role: Neo-Štokavian has been the standard since mid-19th century, which led to extensive dialect contact in the whole language area. More important than external factors is the issue of the driving force – the linguistic factor that set the change in motion and determined in which contexts it took place. As argued above, there are general reasons behind this language change in terms of markedness. In OT literature, this translates into constraint interactions favouring stem stress (e.g. Revithiadou 1999; Alderete 2001; Kager 2000). The most direct result is achieved

by promoting a single markedness constraint: STEMSTRESS. Step 2 in the considered language change seems to show a rather advanced stage in promoting STEMSTRESS. After step 1, stress had been removed from most unambiguous inflectional endings (-*te* in the example above only expresses phi-features). Next, in step 2, stress is removed from the remaining non-stem material (theme vowels, augments etc.).

6 Conclusions

In this contribution, we have presented an account of several diachronic changes within standard Serbo-Croatian, showing that there is the same underlying force behind all of them: strengthening stem prosody. We assumed OT constraints as a tool for formalising the force behind language change and using this tool proved to be useful, enabling us to distinguish between different processes (e.g. singling out the process discussed in §2.4). This inquiry resulted in identifying STEMSTRESS as the constraint that got promoted, leading to a number of seemingly unrelated changes in different morphological contexts. Crucially, considering any of the changes in isolation would not have revealed the presented generalisation.

Abbreviations

1	1st person	INS	instrumental
2	2nd person	LOC	locative
3	3rd person	NOM	nominative
DAT	dative	OT	Optimality Theory
GEN	genitive	PST	past
H	high tone	PL	plural
INF	infinitive	SG	singular

Acknowledgements

References

Alderete, John. 2001. *Morphologically governed accent in Optimality Theory*. New York/London: Routledge.

Bethin, Christina Y. 2006. *Slavic prosody: Language change and phonological theory*. Cambridge: Cambridge University Press.

Burzio, Luigi. 1996. Surface constraints versus underlying representations. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, 123–141. Salford: European Studies Research Institute, University of Salford.

Janda, Richard D., Brian D. Joseph & Neil G. Jacobs. 1994. Systematic hyperforeignisms as maximally external evidence for linguistic rules. In Susan D. Lima, Roberta Corrigan & Gregory Iverson (eds.), *The reality of linguistic rules*, 67–92. Amsterdam: John Benjamins. DOI:10.1075/slcs.26.08jan

Kager, René. 2000. Stem stress and peak correspondence in Dutch. In Joost Dekkers, Frank Van der Leeuw & Jeroen Van de Weijer (eds.), *Optimality Theory: Phonology, syntax, and acquisition*, 121–150. Oxford: Oxford University Press.

Kager, René. 2008. Lexical irregularity and the typology of contrast. In Kristin Hanson & Sharon Inkelas (eds.), *The nature of the word*, 396–430. Cambridge, MA: MIT Press. DOI:10.7551/mitpress/9780262083799.003.0017

Kapović, Mate. 2004. Jezični utjecaj velikih gradova [The linguistic influence of the big cities]. *Rasprave Instituta za hrvatski jezik i jezikoslovlje* 30. 97–105. https://hrcak.srce.hr/9464.

Kenstowicz, Michael. 1996. Base identity and uniform exponence: Alternatives to cyclicity. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods,* Salford: European Studies Research Institute, University of Salford.

Klaić, Bratoljub. 2007. *Bizovačko narječje [The dialect of Bizovac]*. Bizovac: Matica Hrvatska, ogranak Bizovac.

Klaić, Bratoljub. 2013. *Naglasni sustav standardnoga hrvatskoga jezika* [The standard Croatian pitch-accent system]. Božidar Smiljanić (ed.). Zagreb: Nova knjiga Rast.

Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar*. Tech. rep. 2. Rutgers University Center for Cognitive Science.

Revithiadou, Anthi. 1999. *Headmost accent wins: Head dominance and ideal prosodic form in lexical accent systems.* Leiden: Leiden University. (Doctoral dissertation). https://rucore.libraries.rutgers.edu/rutgers-lib/38604/PDF/1/play/.

Simonović, Marko. 2015. *Surface bases and lexical conservatism: The case of Serbo-Croatian -ova-.* Handout from the presentation in the 12th Old World Conference in Phonology. https://www.academia.edu/10459760/Surface_bases_and_Lexical_Conservatism_The_case_of_Serbo_Croatian_-ova.

Simonović, Marko & Boban Arsenijević. 2015a. *The prosody of Serbo-Croatian locatives: Some preliminary remarks.* Presentation at TIN-dag / Grote Taaldag 2015. https://www.academia.edu/10676987/The_prosody_of_Serbo-Croatian_locatives_Some_preliminary_remarks.

Simonović, Marko & Boban Arsenijević. 2015b. *The sound of lexicalisation: Post-lexical prosody in Serbo-Croatian.* Presentation at the 48th Annual Meeting of the Societas Linguistica Europaea. https://www.academia.edu/31806009/The_Sound_of_Lexicalisation_Post-Lexical_Prosody_in_Serbo-Croatian.

Simonović, Marko & René Kager. 2017. *From accent to stress: The verb prosody of Serbo-Croatian creolised standard varieties.* Presentation at Old World Conference on Phonology 2017 in Düsseldorf, Germany, February 20-22, 2017. https://www.academia.edu/31737025/From_accent_to_stress_the_verb_prosody_of_Serbo-Croatian_creolised_standard_varieties.

Stevanović, Mihailo. 1979. *Savremeni srpskohrvatski jezik: Gramatički sistemi i književnojezička norma [Modern Serbo-Croatian: The systems of grammar and the standard-language norm].* Beograd: Naučna knjiga.

Zsiga, Elizabeth C. & Draga Zec. 2013. Contextual evidence for the representation of pitch accents in Standard Serbian. *Language and Speech* 56(1). 69–104. DOI:10.1177/0023830912440792

Chapter 14

Several quantifiers are different than others: Polish indefinite numerals

Marcin Wągiel
Masaryk University in Brno

In this paper, I examine properties of two Polish indefinite quantifiers, namely *ileś* 'some, some number' and *kilka* 'several, a few'. I argue that they share morpho-syntactic properties with cardinal numerals rather than with vague quantifiers such as *mało* 'little, few' and *dużo* 'much, many' and propose that they should be modeled as involving a built-in classifier comprising both a measure function and choice function. The difference between the two indefinites boils down to the type of set the choice function selects a member from and the type of measure function that is employed.

Keywords: indefinites, numerals, quantifiers, choice functions, classifiers, Polish

1 Introduction

For some time, properties of different series of Slavic indefinites have been successfully explored (e.g., Błaszczak 2001, Testelets & Bylinina 2005, Yanovich 2005, Geist 2008, Pereltsvaig 2008, Eremina 2012, Dočekal & Strachoňová 2015, Richtarcikova 2015, Šimík 2015, Strachoňová 2016). However, one particular class of indefinite expressions seems to have been somewhat overlooked, namely indefinite quantifiers such as those exemplified in (1).

(1)   a.   neskol'ko   (Russian)

         b.   několik   (Czech)

         c.   nyakolko   (Bulgarian)

         d.   nekoliko   (Bosnian/Croatian/Serbian)

            several/a few

Remarkably, a similar gap is discernible in a long and prolific research on quantifiers since certain characteristics of expressions corresponding to English *several* remain surprisingly understudied (with the notable exception of Kayne 2007). In this paper, I will examine an alternation involving two types of Polish indefinite quantifiers, such as those seen in (2). In terms of terminology, I will follow the Polish descriptive tradition and refer to such expressions as INDEFINITE NUMERALS, a term which I take to be legitimate in view of the data discussed in the subsequent sections.

(2)   a.   ileś (tam)
           some/some number
      b.   kilka
           several/a few

Though the alternation does not hold in every Slavic language, it does not seem to be a Polish idiosyncrasy, as attested by the prima facie similar contrast between Russian *skol'ko-to* and *neskol'ko* 'several'. The approach developed here is intended to fit a broader research program dedicated to accounting for semantic properties of distinct types of Slavic numeral expressions (Dočekal 2012; 2013; Wągiel 2014; 2015; to appear; Dočekal & Wągiel 2018). Thus, the insights presented here might have wider applicability, at least within Slavic.

The paper is outlined as follows. In §2, I employ a battery of tests to determine morpho-syntactic and semantic properties of the Polish indefinite numerals *kilka* and *ileś*. In §3, I discuss additional data concerning the alternation in question including the evidence in favor of specificity. In §4, I introduce the basic machinery necessary for the analysis: i.e., choice functions, measure functions, and the intersective theory of cardinal numerals. In §5, I develop a morpho-semantic approach to account for the discussed data. Finally, §6 concludes the article.

2 Cardinals, indefinite numerals, and vague quantifiers

2.1 Polish indefinite series

Similar to other Slavic languages, there are several series of indefinite expressions in Polish and Table 1 gives the paradigm for the main ones. Based on morphological evidence, it seems straightforward to assume that Polish indefinites constitute derivationally complex expressions which can be decomposed into a wh-element and an indefinite suffix. In addition, the indefinites in the *-ś* series

can be followed by an optional pronoun *tam* 'there' which can express either a great level of ignorance, or depreciative attitude (cf. Bylinina 2010).

Table 1: Indefinite series in Polish

wh-word		-ś	-kolwiek	-bądź
kto	'who'	ktoś (tam)	ktokolwiek	kto bądź
co	'what'	coś (tam)	cokolwiek	co bądź
gdzie	'where'	gdzieś (tam)	gdziekolwiek	gdzie bądź
kiedy	'when'	kiedyś (tam)	kiedykolwiek	kiedy bądź
jak	'how'	jakoś (tam)	jakkolwiek	jak bądź
jaki	'what/which'	jakiś (tam)	jakikolwiek	jaki bądź
ile	'how much/many'	ileś (tam)	ilekolwiek	ile bądź

As the last row in Table 1 shows, the Polish wh-word *ile* 'how much/ many' can take the indefinite morpheme *-ś* as well as the free choice item (FCI) markers *-kolwiek* and *bądź*. Unlike other wh-words, it is incompatible with the negative prefix *ni-* (*nile vs. *nikt* 'no one') and the depreciative FCI element *byle* (*byle ile vs. *byle kto* 'anyone (someone considered unworthy)') but it can occur within grammaticalized expressions such as *bóg wie ile* 'God knows how much/many' and *chuj wie ile* 'who the fuck knows how much/many'. Despite the fact that the *ile* series is somewhat defective compared to other wh-words, *ileś* is a proper indefinite whose meaning could be probably best paraphrased in English as *some number* or *some amount*.

On the other hand, *kilka* seems to be semantically more restricted. According to the intuition of a majority of Polish native speakers it refers to a number between 3 and 9.[1] Unlike *ileś*, it does not seem to be derivationally complex. In terms of etymology, it emerged from the obsolete wh-word *koliko* 'how much/many' (compare, e.g., Czech *kolik* 'how much/many' ~ *několik* 'some/several') and the cluster *-il-* is arguably related to *ile* (see Bańkowski 2000). However, from a synchronic perspective this relationship is completely opaque and for simplicity I will assume that *kilka* is not a derived form and can only be decomposed into the stem *kilk-* and the following inflectional marker.

I will refrain here from discussing the FCIs *ilekolwiek* and *ile bądź* 'any amount/ number' and for the purposes of this paper I will assume that whatever approach

---

[1]Such an intuition is corroborated by the lexical entries in standard dictionaries of the contemporary Polish language though perhaps it might be subject to some extent to vagueness or interpersonal variation.

accounts for, e.g., the *kto* 'who' ∼ *ktoś* 'someone' ∼ *ktokolwiek* 'anyone' series (e.g., Kadmon & Landman 1993, Aloni 2007, Chierchia 2013), could also be applied to the *ile* 'how much/many' ∼ *ileś* 'some amount/number' ∼ *ilekolwiek* 'any amount/number' alternation. Therefore, in the following text I will focus exclusively on discussing novel data concerning the distribution as well as morpho-syntactic and semantic properties of *ileś* and *kilka.*

To begin with, I will assume that two justifiable hypotheses can be formulated with respect to the nature of the analyzed indefinites: (i) *ileś* and *kilka* are similar to other vague quantifiers or (ii) to cardinal numerals. I will confine my focus to testing properties of these expressions in comparison to *pięć* 'five' on the one hand and *mało* 'few/little' and *dużo* 'much/many' as two representatives of a wider class of vague quantifiers (including lexical items such as *sporo* 'much/many', *trochę* 'some', *niemało* 'quite a lot', *niedużo* 'not much/many', and *masę* 'plenty') on the other. Although due to some lexical idiosyncrasies not every representative of that class has all the discussed properties, e.g., *trochę*, *masę*, and *sporo* are not gradable, the general picture is roughly as presented below.

2.2 Inflection

I will start with the observation that in many respects Polish *kilka* and *ileś* pattern with higher cardinals (i.e., five and higher) rather than with vague quantifiers such as *mało* and *dużo*. Similar to *pięć*, both *kilka* and *ileś (tam)* agree in gender with a modified NP and display the well-documented virile vs. non-virile alternation (e.g., Miechowicz-Mathiasen 2011). On the other hand, *mało* and *dużo* display no reflex of gender agreement with a modified NP, and thus have virile forms, see (3)–(4).

(3)   a.   {Mało dziewczyn / mało chłopców} przyszło.
          few   girls.NV   few boys.V   came
          'A few {girls / boys} came.'

       b.   {Dużo dziewczyn / dużo chłopców} przyszło.
          many girls.NV   many boys.V   came
          'Many {girls / boys} came.'

(4)   a.   {Pięć   dziewczyn / pięciu chłopców} przyszło.
          five.NV girls.NV   five.V   boys.V   came
          'Five {girls / boys} came.'

       b.   {Kilka   dziewczyn / kilku   chłopców} przyszło.
          several.NV girls.NV   several.V boys.V   came
          'Several {girls / boys} came.'

  c. {Ileś dziewczyn / Iluś chłopców} przyszło.
   some.NV girls.NV some.V boys.V came
   'Some {girls / boys} came.'

Another morpho-syntactic similarity between indefinite numerals and cardinals is that, unlike *mało* and *dużo*, the indefinites *ileś* and *kilka* do not take a comparative and superlative, see (5)–(6).

(5) a. mało ~ mniej ~ najmniej
   few fewer fewest

  b. dużo ~ więcej ~ najwięcej
   much more most

(6) a. pięć ~ *{pięciej / bardziej pięć} ~ *{najpięciej / najbardziej pięć}
   five five.CMPR more five five.SPRL most   five

  b. kilka ~ *{kilkiej   / bardziej kilka} ~ *{najkilkiej /
   several several.CMPR more several several.SPRL
   najbardziej kilka}
   most   several

  c. ileś ~ *{ilesiej / bardziej ileś} ~ *{najilesiej / najbardziej
   some some.CMPR more some some.SPRL most
   ileś}
   some

In the following sections, I will test the grammaticality of *kilka* and *ileś* in multiple environments in comparison to cardinal numerals and the quantifiers *mało* and *dużo*. I will start with different types of modifiers.

2.3 Degree and numeral modifiers

One can distinguish between two types of modifiers that can combine with quantifiers: (i) degree modifiers such as *very (much)* and (ii) numeral modifiers such as *over (five)*.[2] Degree modifiers are compatible with quantifiers such as *mało* and *dużo* but cannot combine with cardinal numerals. On the other hand, numeral modifiers can target cardinals but fail to modify gradable quantifiers. Interestingly, the indefinite numeral *kilka* behaves exactly like cardinals. The examples in (7)–(10) illustrate the pattern.

---

[2]Nouwen (2010) further distinguishes between class A and B numeral modifiers. However, for the purpose of this paper a simplified view is entirely sufficient.

(7)　a.　bardzo mało
　　　　very　few

　　b.　dość　mało
　　　　rather few

　　c.　zbyt mało
　　　　too　few

　　d.　tak mało
　　　　so　few

　　e.　niemało
　　　　not.few

(8)　a.　* ponad mało
　　　　　over　few

　　b.　* najwyżej mało
　　　　　up.to　　few

　　c.　* około　mało
　　　　　around few

　　d.　* co najmniej mało
　　　　　at least　　few

　　e.　* od　mało do stu
　　　　　from few　to 100

(9)　a.　* bardzo {pięć / kilka}
　　　　　very　　five　several

　　b.　* dość　{pięć / kilka}
　　　　　rather five　several

　　c.　* zbyt {pięć / kilka}
　　　　　too　five　several

　　d.　* tak {pięć / kilka}
　　　　　so　five　several

　　e.　* niepięć / *niekilka
　　　　　not.five　　not.several

(10)　a.　ponad {pięć / kilka}
　　　　　over　five　several

　　b.　najwyżej {pięć / kilka}
　　　　up.to　　　five　several

　　c.　około　{pięciu　/
　　　　around　five.GEN
　　　　kilku}
　　　　several.GEN

　　d.　co najmniej {pięć / kilka}
　　　　at least　　　five　several

　　e.　od　{pięciu　/ kilku}
　　　　from　five.GEN　several.GEN
　　　　do stu
　　　　to 100

Similar to cardinals and *kilka*, the indefinite *ileś* is incompatible with degree modifiers, see (11). Nevertheless, unlike the expressions discussed above it seems to be degraded with most numeral modifiers. Notice, however, that despite this fact, the contrast between (12) and (8) is still detectable.

(11)　a.　* bardzo ileś (tam)
　　　　　very　　some

　　b.　* dość　ileś (tam)
　　　　　rather some

　　c.　* zbyt ileś (tam)
　　　　　too　some

　　d.　* tak ileś (tam)
　　　　　so　some

　　e.　* nieileś　(tam)
　　　　　not.some

(12) a. ? ponad ileś (tam)
         over   some

    b. ? najwyżej ileś (tam)
         up.to   some

    c. ? około   iluś (tam)
         around some.GEN

    d. ? co najmniej ileś (tam)
         at least   some

    e. od   iluś (tam) do stu
         from some.GEN to 100

I speculate that the reason that the acceptability of *ileś* with numeral modifiers is reduced is its high level of indefiniteness. Since such modifiers compare more or less defined values, at least some approximation with respect to the targeted set of numbers is required. Out of the blue (12a) sounds odd, but if a proper context sets a plausible range of possible values, it becomes perfectly acceptable, as attested in an example from the National Corpus of Polish (NCP) provided in (13).

(13) […] jeśli stan załogi wynosi ponad ileś tam   osób   […] powinien
        if   state crew   equals over some.number people   it.should
    być zespół muzyczny […]
    be band   musical
    '[…] if a crew amounts to more than some number of people […] there should be a music band arranged […]'

All in all, the discussed data seem to indicate the distinction between quantifiers *mało* and *dużo* on the one hand and cardinals and the indefinites *kilka* and *ileś* on the other. The next test will involve the (un)grammaticality of quantificational NPs where the quantifier is modified by the adjective or possessive pronoun.

2.4 Adjectival and pronominal modification

It has been observed that Polish cardinals are compatible with agreeing adjectival modifiers such as *dobre* 'good' in preposition, see (14a), (cf. Babby 1987 and Miechowicz-Mathiasen 2011).[3] As indicated by the translations, if the preceding AP employs the agreement strategy, it is the referent of the numeral that is modified and not the quantified entities; e.g., in (14a) it is the number of cookies that is

---

[3]An anonymous reviewer wonders whether *dobre* in examples such as (14) is in fact an adjective and whether it could be analyzed as an adverbial element. The case, gender, and number agreement point to the contrary and, as far as I can tell, there is no evidence for the adverbial nature of *dobre* in such examples. Furthermore, swapping the standard adverb *dobrze* 'well' for *dobre* results in ungrammaticality.

good, not necessarily the cookies themselves. Again, *kilka* and *ileś* pattern with cardinal numerals in this respect, whereas *mało* and *dużo* do not allow for adjectival modification, as witnessed by the ungrammaticality of (15).[4]

(14) a. dobre pięć ciasteczek
          good five cookies.GEN
          'a good five cookies'

      b. dobre kilka   ciasteczek
          good several cookies.GEN
          'a good several cookies'

      c. dobre ileś (tam) ciasteczek
          good some   cookies.GEN
          'a good number of cookies'

(15) a. * dobre mało ciasteczek
            good few cookies.GEN
            Intended: 'a good few cookies'

      b. * dobre dużo ciasteczek
            good many cookies.GEN
            Intended: 'a good many cookies'

Similarly, both cardinals and indefinite numerals allow for pronominal modification employing the agreement strategy (cf. Miechowicz-Mathiasen 2011), while expressions such as *mało* and *dużo* do not, as witnessed by the contrast between (16) and (17).

(16) a. {te   / moje} pięć ciasteczek
           these my   five cookies.GEN
           'these / my five cookies'

      b. {te   / moje} kilka   ciasteczek
           these my   several cookies.GEN
           'these / my several cookies'

      c. {te   / moje} ileś (tam) ciasteczek
           these my   some   cookies.GEN
           'these / my cookies some number worth'

---

[4]It seems that there is a dialectal variation since some Polish speakers judge examples such as those in (14) as ungrammatical and accept only APs which agree with the noun to precede the quantificational NP. However, to my knowledge for such speakers the use of the genitival form *dobrych* 'good' in (15) is still impossible.

(17) a. * {te   / moje} mało ciasteczek
          these my   few cookies.GEN
          Intended: 'these / my few cookies'

   b. * {te   / moje} dużo   ciasteczek
          these my   many cookies.GEN
          Intended: 'these / my many cookies'

The (in)compatibility with different types of modifiers appears to be a reliable diagnostic for the classification of quantifiers and it suggests a distinction between cardinals and *kilka* and *ileś* on the one hand and *mało* and *dużo* on the other. Yet another test will explore the acceptability of the expressions in question in contexts involving universal quantification and markers forcing obligatory distributive readings.

2.5 Universal quantification and distributivity

It is a well-known fact that Slavic numerals can co-occur with the universal quantifier within one phrase (cf. Corbett 1978, Gvozdanović 1999, and Miechowicz-Mathiasen 2011). Examples such as those in (18) show that, similar to cardinals, the indefinites *kilka* and *ileś* are also licit in such an environment. However, expressions such as *mało* and *dużo* do not allow for modification by a universal quantifier, see (19).

(18) a. wszystkie pięć ciasteczek
          all   five cookies.GEN
          'all the five cookies'

   b. wszystkie kilka   ciasteczek
          all   several cookies.GEN
          'all the several cookies'

   c. wszystkie ileś (tam) ciasteczek
          all   some   cookies.GEN
          'all the cookies (where there are some cookies)'

(19) a. * wszystkie mało ciasteczek
            all   few cookies.GEN
            Intended: 'all the few cookies'

   b. * wszystkie dużo   ciasteczek
            all   many cookies.GEN
            Intended: 'all the many cookies'

Another contrast relates to distributivity. As observed by Safir & Stowell (1988) and discussed by Borer (2005), English binominal *each* does not allow the dis-

tributive share expressed by DPs involving the quantifier *some*. Similarly, there are a number of restrictions on arguments of the distributive preposition *po* in Polish which excludes a collective reading of a sentence in which it occurs (Przepiórkowski 2008). Interestingly, phrases headed by quantifiers such as *mało* and *dużo* are not acceptable as complements of *po*, see (21), unlike *kilka* and *ileś* which again pattern with cardinals, see (20).

(20) a. Dałem im   po   pięć ciasteczek.
          I.gave them DISTR five cookies.GEN
          'I gave them five cookies each.'

     b. Dałem im   po   kilka   ciasteczek.
          I.gave them DISTR several cookies.GEN
          'I gave them several cookies each.'

     c. Dałem im   po   ileś (tam) ciasteczek.
          I.gave them DISTR some   cookies.GEN
          'I gave some cookies to each of them.'

(21) a. * Dałem im   po   mało ciasteczek.
            I.gave them DISTR few   cookies.GEN
            Intended: 'I gave few cookies to each of them.'

     b. * Dałem im   po   dużo ciasteczek.
            I.gave them DISTR many cookies.GEN
            Intended: 'I gave many cookies to each of them.'

It seems that the contrasts discussed here cannot simply stem from, e.g., distinct ranges of vagueness or other superficial differences between the indefinites *kilka* and *ileś* as compared to *mało* and *dużo*. Rather, the data suggest that a much more essential disparity is involved and the expressions in question should be treated as belonging to two distinct classes.

2.6 Uncountable NPs

So far, we have discussed environments in which cardinals pattern both with *kilka* and *ileś*. However, another division can be drawn based on the interaction with uncountable nominals such as mass nouns and pluralia tantum. While cardinals and *kilka* cannot combine directly with such expressions[5] and require either a measure word or a specialized classifier suffix, (22), *ileś* patterns in this

---

[5]I put aside cases where the mass denotation is shifted to the count domain by means of the universal packager or the universal sorter.

respect with quantifiers such as *mało*.[6] In particular, it is compatible both with mass nouns and pluralia tantum as well as measure and classifier constructions involving such expressions and cannot take the classifier suffix, as presented in (23).

(22) a. * {pięć / kilka}   wody
           five   several water.GEN

   b. {pięć / kilka} butelek   wody
       five   several bottles.GEN water.GEN

   c. * {pięć / kilka}   nożyczek
           five   several scissors.GEN

   d. {pięć / kilka} par   nożyczek
       five   several pairs.GEN scissors.GEN

   e. % {pięcioro / kilkoro}   drzwi
         five.CL   several.CL door.GEN.PL

(23) a. {mało / ileś (tam)} wody
          little   some   water.GEN

   b. {mało / ileś (tam)} butelek   wody
       few   some   bottles.GEN water.GEN

   c. {mało / ileś (tam)} nożyczek
       few   some   scissors.GEN

   d. {mało / ileś (tam)} par   nożyczek
       few   some   pairs.GEN scissors.GEN

   e. * {małoro / ilesioro (tam)} drzwi
         few.CL   some.CL   door.GEN.PL

Before we move on to discussing more contrasts regarding *kilka* and *ileś*, let us recapitulate the findings so far.

2.7 Data summary

Table 2 summarizes morpho-syntactic and distributional properties of Polish indefinite numerals as compared to cardinals.

---

[6]The use of forms such as *pięcioro* and *kilkoro* with pluralia tantum seems to be fading, especially in younger generations. Some speakers, however, still use such expressions and the plurale tantum noun *drzwi* 'door' ranks in 11th place as a collocation candidate for the lemma *kilkoro* in the NCP. For more details concerning different uses of suffixed numerals such as *pięcioro* see Wągiel (2014; 2015).

Table 2: Morpho-syntactic properties of cardinals and indefinite numerals

	pięć 'five'	*kilka* 'several'	*ileś (tam)* 'some'	*mało* 'few/little'
degree modifiers	*	*	*	✓
comparison	*	*	*	✓
mass nouns	*	*	✓	✓
pluralia tantum	*	*	✓	✓
virile vs. non-virile	✓	✓	✓	*
universal quantifier	✓	✓	✓	*
distributive PO	✓	✓	✓	*
adjectival modifiers	✓	✓	✓	*
pronominal modifiers	✓	✓	✓	*
numeral modifiers	✓	✓	?	*

As Table 2 shows, three patterns can be distinguished within an axis extending over poles constituted by compatibility with numeral modifiers on the one hand and degree modifiers on the other. Based on the battery of tests applied in this section, cardinal numerals and the indefinite numeral *kilka* appear to form a logical class which contrasts with the class of vague quantifiers such as *mało* and *dużo*. On the other hand, the indefinite numeral *ileś* seems to somewhat fall in between the two categories. Although it shares a number of key properties with cardinals, it is not subject to the distributional constraints concerning direct modification of uncountable expressions.

I conclude that *kilka* is essentially a cardinal in disguise, whereas *ileś* seems to be a numeral augmented with some more general semantic features. In the next section, I will provide more data that shed new light on the core of the discussed alternation.

3 Some intriguing contrasts

3.1 Predicate position

As illustrated in (24), Polish cardinals and indefinite numerals have yet another property in common, namely they both can appear in predicate position.

(24)   a.   Tych dziewczyn było pięć.
           these girls   were five
           'The girls were five in number.'

        b.   Tych dziewczyn było kilka.
           these girls   were several
           'The girls were several in number.'

        c.   Tych dziewczyn było ileś (tam).
           these girls   were some
           'The girls were in some number.'

At this point, it might be tempting to analyze *kilka* and *ileś* essentially on a par with *pięć*. However, this is not the whole story. In the following sections, I will focus on some non-trivial differences between cardinals and indefinite numerals on the one hand and *kilka* and *ileś* on the other. By examining this distinction more closely, we can provide a proper semantic account of Polish indefinite numerals.

3.2 Reference to number concepts

One could attempt to analyze indefinite expressions such as English *several* in terms of existential quantification over numbers of a certain size. However, it appears that there is a serious problem with the existential quantification approach (see Schwarzschild 2002). In particular, indefinite numerals differ from cardinal numerals in that they cannot be used to name number concepts and do not fit contexts calling for numerical arguments, see (25). Furthermore, consider the mathematical statement in (26a). A natural way to paraphrase it making use of the existential quantifier is given in (26b). Nonetheless, similar statements involving indefinites in (27a) and (28a) are not felicitous despite the fact that their intended meaning can be easily paraphrased in terms of existential quantification, as provided in (27b) and (28b) respectively.

(25)   a.   liczba   {pięć / *kilka   / *ileś}
           number   five   several   some

        b.   Jaś umie policzyć do {pięciu   / *kilku   / *iluś}.
           Jaś can   count.up to   five.GEN   several.GEN   some.GEN
           'Jaś can count up to five.'

(26)   a.   Cztery plus pięć to   mniej niż   dziesięć.
           four   plus five   this less   than ten
           'Four plus five is less than ten.'

        b.   There is a number $n = 5$ such that $4 + n < 10$.

(27) a. # Cztery plus kilka   to   mniej niż   dziesięć.
          four   plus several this less   than ten
          Intended: 'Four plus several is less than ten.'

    b. There is a number $n \geq 3 \wedge \leq 9$ such that $4 + n < 10$.

(28) a. # Cztery plus ileś (tam) to   mniej niż   dziesięć.
          four   plus some   this less   than ten
          Intended: 'Four plus some number is less than ten.'

    b. There is a number n such that $4 + n < 10$.

The facts described above suggest that Polish indefinite numerals cannot be modeled in terms of existential quantification over numbers. The following section will provide additional evidence calling for an alternative treatment.

3.3 Specific reading

To my knowledge, it is a novel observation that Polish indefinite numerals can have a so-called specific reading, i.e., an interpretation corresponding to the widest scope in the sentence (cf. Fodor & Sag 1982 and Kratzer 1998).[7] For instance, (29) can be interpreted with *każdy* 'each' scoping over *kilka*: i.e., for each teacher there is some indefinite number of which they know that that many students were called before the dean. Such an interpretation is sometimes referred to as a quantificational reading. However, (29) can also mean that in a given context there is a certain number of my students, say five, and each teacher knows that the number of my students that were called before the dean is that number. The same applies to *ileś*, as illustrated in (30).

(29) Każdy nauczyciel wie,   że kilku moich studentów wezwano   do
      each teacher   knows that several my   students were.called to
      dziekana.
      dean
      'Each teacher knows that several students of mine had been called before
      the dean.'

    a. each > kilka   quantificational reading
    b. kilka > each   specific reading

---

[7]Fodor & Sag (1982) call it a "referential interpretation". I will stick to the term SPECIFIC though.

(30)   Każdy nauczyciel wie,   że   iluś (tam) moich studentów wezwano   do
       each   teacher   knows that some   my   students   were.called to
       dziekana.
       dean
       'Each teacher knows that some students of mine had been called before
       the dean.'

   a. each > ileś   quantificational reading
   b. ileś > each   specific reading

Alongside the ability to escape islands and insensitivity to various operators, the capability to take the widest scope is considered to be one of the diagnostics to detect specific indefinites such as *a certain word* in (31).

(31)   There is a certain word that I can never remember.

In a certain way the evidence seems to steer in the opposite directions. On the one hand, indefinite numerals appear to be 'referential' in the sense that they can indicate a specific, though indefinite, number. On the other hand, however, they are infelicitous in contexts calling clearly for numerical arguments such as terms in mathematical equations.

3.4 Referential restrictions

Another contrast concerns referential restrictions that apply to *kilka*. While *ileś* can be used to denote any real (or perhaps even complex) number, *kilka* seems to be restricted to a subset of integers, specifically the set {3, 4, 5, 6, 7, 8, 9}.[8] Notice that a constraint regarding natural numbers seems to apply also to cardinals. For instance, in a scenario where there are four and a half apples on the table and it is conspicuous that the half does not count as a whole apple, it is rather odd to utter (32a).[9] In such a context, it is also strange to use (32b). However, (32c) seems perfectly felicitous. A similar contrast is given in (33). Since π is an irrational number, it can be associated with the co-referential *ileś* in the main clause, however using *kilka* in such a sentence is impossible.

---

[8]Some speakers may include 10 while others may restrict the set even further by excluding 3. I acknowledge that this issue might be subject to some degree to idiolectal variation but for the sake of simplicity I will ignore this fact in the following analysis.

[9]An example of such a scenario would be a cooking event in which one bakes stuffed apples. In such a context half an apple is useless and simply does not count.

(32) a. # Na stole leży pięć jabłek.
          on table lies five apples
          Intended: 'There are five apples on the table.'

     b. # Na stole leży kilka   jabłek.
          on table lies several apples
          Intended: 'There are several apples on the table.'

     c. Na stole leży ileś (tam) jabłek.
          on table lies some   apples
          'There are some apples on the table.'

(33) a. # Pole koła to   kilka   razy   r^2, a   dokładnie π razy r^2.
          area circle this several times r^2 and precisely π times r^2
          Intended: 'The area of a circle is several times r^2, precisely π times r^2.'

     b. Pole koła to   ileś (tam) razy r^2, a   dokładnie π razy r^2.
          area circle this some   times r^2 and precisely π times r^2
          'The area of a circle is some number times r^2, precisely π times r^2.'

The data suggest yet another distinction between indefinite numerals. Similar to cardinals, *kilka* makes reference to natural numbers whereas *ileś* is not restricted in such a way. Rather, it is apt to denote any number associated with a particular plurality or quantity.

3.5 Cardinal suffixes

The final data point to be discussed in this section concerns an interesting fact that unlike, e.g., English *several* (Kayne 2007), the Polish indefinite *kilka* can take cardinal suffixes, as illustrated in (34b). On the other hand, *ileś* is significantly degraded with cardinal suffixes, see (34c).[10]

(34) a. pięć ~ piętnaście ~ pięćdziesiąt ~ pięćset
          five   fifteen   fifty   five.hundred

     b. kilka ~ kilkanaście ~ kilkadziesiąt ~ kilkaset
          several several.teen   several.ty   several.hundred

---

[10] Although such forms are definitely not part of standard Polish and many speakers judge them as ungrammatical, for some speakers they are marginally acceptable. However, the balanced NCP subcorpus which contains more than 240 million tokens returns no hits for the forms *ileś-naście* and *ileśdziesiąt* and six hits for *ileśset*, two of which are from the prose of a linguistically very creative author. Therefore, I will assume that such forms are not well-formed expressions of Polish.

c. ileś ~ *ileśnaście ~ *ileśdziesiąt ~ *ileśset
    some   some.teen   some.ty   some.hundred

Interestingly, the interpretation of the suffixed indefinite numerals seems to be derived from the meaning of *kilka*. For instance, at least for some speakers *kilka-naście* does not mean a number between 11 and 19 but rather it seems to exclude the values 11 and 12, hence {13,..., 19}. Similar, it would be awkward to refer to a plurality including approximately twenty members using *kilkadziesiąt*; for a collection of around thirty entities it would be felicitous though. In spite of the fact that such intuitions may not be shared by all native speakers and I suspect some interpersonal variation here, my judgments as well as the judgments of the informants I have consulted are quite clear with respect to this issue and I will assume them to hold in general.

3.6 Data summary

Although Polish indefinite numerals pattern with cardinals such as *pięć* 'five' rather than with vague quantifiers such as *mało* 'few/little' and *dużo* 'much/many', there are a number of respects in which they differ. In particular, though both cardinals and indefinite numerals can occur in predicate position and can have a specific reading, *kilka* and *ileś* cannot be used to name numbers, i.e., to refer to abstract concepts, and do not fit clearly numerical contexts. On the other hand, *ileś* differs from cardinals and *kilka* in that it cannot take cardinal suffixes and is not restricted to natural numbers: i.e., unlike *kilka* it can be used to talk about any real and possibly even complex number. Table 3 summarizes the similarities and contrasts discussed in this section.

Table 3: Semantic properties of cardinals and indefinite numerals

	pięć 'five'	*kilka* 'several'	*ileś (tam)* 'some'
predicate position	✓	✓	✓
specific reading	✓	✓	✓
cardinal suffixes	✓	✓	*
restricted to integers	✓	✓	*
names of numbers	✓	*	*
numeric contexts	✓	*	*

I conclude that a neat classification developed here calls for a more elaborate analysis of numerical expressions than usually assumed. In particular, a proper treatment of numerical expressions should account for the semantic differences between the class of cardinals and two types of indefinite quantifiers, namely *kilka* and *ileś*.

Before we move on to spelling out the semantics for indefinite numerals that will capture the discussed patterns and contrasts, it will be useful to introduce several theoretical tools. In the next section I will sketch a framework within which the proposed analysis will be grounded.

4 Setting the stage

4.1 Choice functions

Following Reinhart (1997) and Kratzer (1998) as well as subsequent cross-linguistic research on specific indefinites (see Alonso-Ovalle & Menéndez-Benito 2003 for Spanish *algún*, Kratzer & Shimoyama 2002 for German *irgendein*, Yanovich 2005 for Russian indefinite series, and Matthewson 1998 for indefinites in St'át'imcets), I model *ileś* and *kilka* as choice functions (CF): i.e., operators selecting a member from a set. On the adopted view, CF indefinites are not existentially quantified. Instead, the CF variable remains free at LF and its value is provided by the context. In particular, I embrace the approach that CFs provide a null pronominal element of type $\langle\langle\tau, t\rangle, \tau\rangle$, where τ is a generalized primitive type, see (35a) and (35b) for entities.

(35)   a.   For any $f_{\langle\langle\tau,t\rangle,\tau\rangle}$ and any $P_{\langle\tau,t\rangle}$, f is a CF if $P(f(P)) = 1$
         b.   For any $f_{\langle\langle e,t\rangle,e\rangle}$ and any $P_{\langle e,t\rangle}$, f is a CF if $P(f(P)) = 1$

If a CF f is applied to a set of, e.g., sleeping individuals, it will yield a specific sleeper relative to a particular context. Similar, when applied to a set of natural numbers, it will return a relevant integer. In this way, one can account for the referential flavor of specific indefinites without employing existential quantification.

4.2 Measure functions

Following Krifka (1989), I model quantification in numeral and measure constructions in terms of extensive measure functions (MF), i.e., operations that map a plurality of individuals or quantity of substance onto a real number corresponding to the number of individuals or units making up the plurality or quantity. Such MFs are additive, see (36a) and have the Archimedean property, see (36b).

In addition, assuming the remainder principle for ⊔ guarantees monotonicity, see (36c) (cf. Schwarzschild 2002).

(36)   a. μ is an additive MF with respect to ⊔ iff for any x_e and any y_e,
       $\neg x \circ y \rightarrow [\mu(x \sqcup y) = \mu(x) + \mu(y)]$

   b. μ is an Archimedean MF iff for any x_e and any y_e,
       $[\mu(x) > 0 \land y \sqsubseteq x] \rightarrow \mu(y) > 0$

   c. μ is a monotonic MF with respect to ⊑ iff for any x_e and any y_e,
       $x \sqsubset y \rightarrow \mu(x) < \mu(y)$

Counting is therefore modeled as a form of measuring. For instance, the MF LITER returns the integer 3 if there are three liters of an entity in question, see (37a). Similar, the MF # can be introduced which would yield 3 if a number of individual members of a plurality it is applied to equals 3, see (37b).[11] Let us assume that # is defined in such a way that it takes only a plurality of atomic individuals, i.e., entities that do not have proper parts, and returns a number of atoms making up that plurality. Such a restriction guarantees its incompatibility with mass nouns unless their denotation is shifted to the count domain, e.g., via the universal packager or the universal sorter.

(37)   a. ⟦three liters of juice⟧ $= \lambda x[\text{JUICE}(x) \land \text{LITER}(x) = 3]$

   b. ⟦three apples⟧ $= \lambda x[\text{APPLE}(x) \land \#(x) = 3]$

Furthermore, to account for the compatibility of *ileś* with both countable and uncountable NPs I will follow Bale & Barner (2009) in assuming a generalized context-dependent MF μ. Such an approach posits a mechanism of contextual conditioning along the lines defined in (38).

(38) μ is interpreted as one of the MFs m_z in the series $\langle m_1, m_2, m_3 \ldots m_n \rangle$ such that the argument for μ is in the range of m_z; furthermore, contextually m_z is preferred to m_y if $z < y$

A contextually conditioned MF can cover the meanings of both pure measure constructions such as (37a) and counting expressions like those in (37b). In particular, μ is interpreted as an MF counting units of, e.g., volume, when combined with a mass term denoting a substance and as an MF counting atomic entities when combined with expressions denoting individuated semi-lattices such as count nouns and pluralia tantum.

---

[11]Here I depart from Krifka's (1989) original proposal. In his system, the NU operation (for 'natural unit') is postulated which when applied to a property returns a number of natural units associated with that property.

4.3 Cardinals

Rothstein (2013; 2017) distinguishes between several functions of numerals. In a non-classifier language such as English cardinals can be used as (i) nominal modifiers, (ii) predicates, and (iii) names of concept numbers. When used in attributive and predicate position numerals are cardinal predicates of the same type as adjectives (Landman 2003), see (39a), whereas when used as names of numbers, they refer to abstract objects of a primitive semantic type n, see (39b). On this view, cardinal predicates denote sets of plural entities with a specific cardinality, i.e., $\{x : \#(x) = n\}$, and have standard intersective semantics.[12] For instance, *three apples* denotes a set of pluralities that are both in the denotation of *apples* and have the property *three*, i.e., a set of triples of apples. Rothstein assumes that cardinal properties are basic, whereas their individual correlates, i.e., names of number concepts, are derived and building on Fregean property theory (Chierchia 1985) postulates shifting operations $^\cup$ and $^\cap$ which allow for switching freely between the two.

(39) a. $[\![\text{three}_{\langle e,t\rangle}]\!] = \lambda x[\#(x) = 3]$
      b. $[\![\text{three}_n]\!] = 3$

In the system described above, complex numerals such as *twenty-three* are derived by means of a null + operator which works as illustrated in (40).

(40) a. $[\![+]\!] = \lambda m\lambda n[m + n]$
      b. $[\![\text{twenty-three}]\!] = \lambda m\lambda n[m + n](20)(3) = \lambda n[20 + n](3) = 20 + 3$

With all the ingredients in place, let us now see what they can account for and how they interact. In the following section, I will provide an analysis of the Polish indefinite numerals *kilka* and *ileś* which captures their similarities with cardinals as well as accounts for the discussed differences.

5 Putting the pieces together

5.1 Adaptations and extensions

Within the patch-work framework adopted here there are several adjustments and developments I will make. First of all, unlike Rothstein, I assume that the

---

[12]Both Landman and Rothstein use the symbol | . . . | instead of #. I have replaced it for the sake of notational uniformity and clarity.

use of cardinals as names of numbers is the basic one. In particular, I posit that numerals are complex expressions involving the numeral root which is an expression of type n, the Numeral head which introduces gender, and optionally the classifier element CARD (for 'cardinal property') which takes a number and returns a set of atomic individuals whose cardinality equals that number, see (41). Proper counting is guaranteed by the # MF and presupposition of atomicity incorporated into the semantics of CARD. In other words, cardinals are born as names of numbers (cf. Scha 1981) and by adding additional structure can be converted to cardinal properties at type $\langle e, t \rangle$. I assume that in a language such as English or Polish CARD has no overt exponent. However, in classifier languages it is introduced by the classifier (see Sudo 2016 for a similar proposal).

(41) $[\![\text{CARD}]\!] = \lambda n \lambda x . \text{ATOM}(x) [\#(x) = n]$

Furthermore, I posit yet another classifier element, namely QUANT (for 'quantificational property') which also shifts number concepts to sets of entities but unlike CARD it employs the contextually conditioned MF μ which can either measure, e.g., volume or count individuals depending on a context. Such conditioning makes QUANT compatible with both countable and uncountable NPs.

(42) $[\![\text{QUANT}]\!] = \lambda n \lambda x [\mu(x) = n]$

Finally, I propose that in Polish suffixed numerals there is no covert + operation but rather cardinal suffixes are number operators of type $\langle n, n \rangle$ themselves. They take the denotation of the numeral root and yield a number enlarged via addition or multiplication, see (43), which can be then shifted by CARD. Notice, however, that the cardinal suffixes incorporate a special presupposition that makes them compatible only with natural numbers. Such a move will explain the behavior of *ileś*, but it is also independently motivated by the fact that cardinal suffixes are not compatible with expressions denoting fractions, as shown by the contrast in (44).

(43)   a. $[\![\text{-naście}]\!] = \lambda n . \text{INTEGER}(n) [n + 10]$
       b. $[\![\text{-dziesiąt}]\!] = \lambda n . \text{INTEGER}(n) [n \times 10]$
       c. $[\![\text{-set}]\!] = \lambda n . \text{INTEGER}(n) [n \times 100]$

(44)   a.   dziesięć i   pół   b.   * półnaście
            ten   and half   half.teen

Let us now examine how the proposed semantics accounts for Polish cardinals and indefinite numerals.

5.2 Composition of cardinals

I argue that Polish cardinal numerals are complex expressions. First, let us consider cardinals in numerical contexts such as (26b) where they are used as names of abstract mathematical concepts. In general, I take numeral roots to be category-free, as often claimed (e.g., Halle & Marantz 1993). Due to the fact that Polish cardinals can be used not only as modifiers and predicates, but also as names of numbers and can be modified by agreeing adjectives, I assume that in a sense they have some nominal-like properties. Therefore, I posit that a gender value is always associated with the Numeral head which forges the cardinal. Let us consider the derivation of the non-virile numeral *pięć* 'five', see Figure 1. The category-free root $\sqrt{pięć}$- is a name of the natural number 5, i.e., an expression of a primitive type *n*. Though the Numeral head has a crucial structural role, i.e., it assigns the [NV] (for 'non-virile') gender value and forms the numeral, it lacks any particular semantic contribution, and the resulting expression is therefore simply the name of number 5.[13]
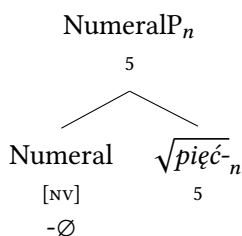
Figure 1: Derivation of the number name *pięć* 'five'

However, the structure in Figure 1 can be further augmented with the silent node which introduces the CARD operation, see Figure 2. As a result, the number 5 is shifted to the set of atomic individuals whose cardinality equals 5. Such an expression can be used both as a nominal modifier and in predicate position.

Finally, a derivationally complex numeral such as *piętnaście* 'fifteen' can be obtained by incorporating the node associated with the cardinal suffix in the structure. Specifically, I posit that it is not until the cardinal suffix attaches to the root and yields an enlarged number that the Numeral head applies and forms the NumeralP which can serve as an argument for CARD. The tree in Figure 3 gives the structure for the non-virile cardinal *piętnaście*; the derivation of other suffixed cardinals is analogous.[14]

---

[13]In the case of the form *pięciu*, the Numeral head assigns the [V] (for "virile") value.

[14]Notice that *pięć-* and *pięt-* are allomorphs, similar to the suffixes *-naści-* and *-nast-*, as in the virile form *piętnastu*. I take *-nast-* to be the basic form and assume that it alternates with *-naści-* in contexts preceding *-e*.

NumeralP$_{\langle e,t \rangle}$
$\lambda x.\text{ATOM}(x)\,[\#(x) = 5]$

CARD$_{\langle n,\langle e,t \rangle \rangle}$
$\lambda n \lambda x.\text{ATOM}(x)\,[\#(x) = n]$

NumeralP$_n$
5

Numeral
[NV]
-∅

$\sqrt{pięć\text{-}}_n$
5
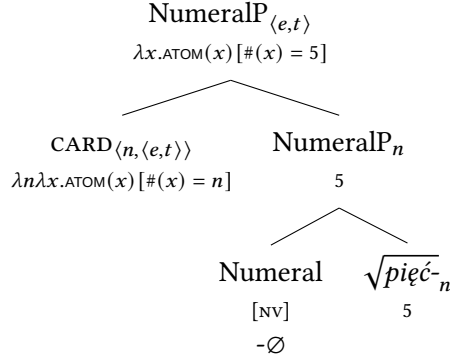
Figure 2: Derivation of the cardinal predicate *pięć* 'five'

NumeralP$_{\langle e,t \rangle}$
$\lambda x.\text{ATOM}(x)\,[\#(x) = 15]$

CARD$_{\langle n,\langle e,t \rangle \rangle}$
$\lambda n \lambda x.\text{ATOM}(x)\,[\#(x) = n]$

NumeralP$_n$
15

Numeral
[NV]
-*e*

n
15

-*nast*-$_{\langle n,n \rangle}$
$\lambda n.\text{INTEGER}(n)\,[n + 10]$

$\sqrt{pięć\text{-}}_n$
5
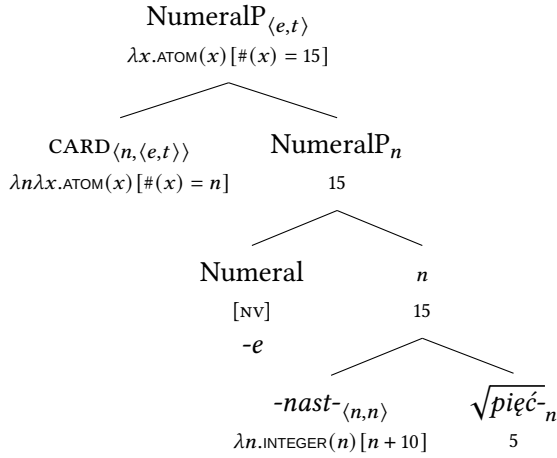
Figure 3: Derivation of the cardinal predicate *piętnaście* 'fifteen'

With the proposed mechanism of deriving Polish cardinals in place, let us now move to the semantics of indefinite numerals. The next section is dedicated to explaining the composition of *kilka* and *ileś*.

5.3 Composition of indefinite numerals

5.3.1 Deriving *kilka*

I will start with the structure for *kilka* 'several', see Figure 4. I presume that the root $\sqrt{kilk}$- involves a built-in CF that applies to the restricted set of alternatives, namely the set of natural numbers {3, 4, 5, 6, 7, 8, 9}, and yields a specific value in a given context. The root then combines with the Numeral head which assigns the [NV] gender value. However, unlike in the case of cardinals, the Numeral head does have a semantic contribution. In particular, it introduces the CARD opera-tion which shifts the indefinite number to the cardinal property. The resulting expression is of type $\langle e, t \rangle$, and thus it is illicit in contexts calling for numeric arguments, as already illustrated in (27a). Furthermore, the fact that the MF # re-quires atomic denotations explains why *kilka* is incompatible with mass terms.

$$\text{NumeralP}_{\langle e,t \rangle}$$
$$\lambda x.\text{ATOM}(x)[\#(x) = f_{\langle \langle n,t \rangle,n \rangle}(\lambda n[\text{INTEGER}(n) \wedge n \geq 3 \wedge n \leq 9])]$$

Numeral
$$[\text{NV}]$$
$$\text{CARD}_{\langle n, \langle e,t \rangle \rangle}$$
$$\lambda n \lambda x.\text{ATOM}(x)[\#(x) = n]$$
$$\text{-}a$$

$$\sqrt{kilk}\text{-}_n$$
$$f_{\langle \langle n,t \rangle,n \rangle}(\lambda n[\text{INTEGER}(n) \wedge n \geq 3 \wedge n \leq 9])$$
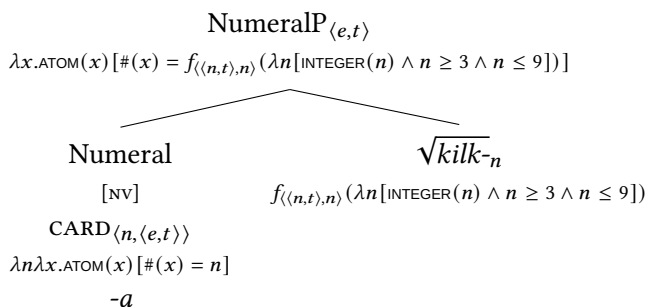
Figure 4: Derivation of *kilka* 'several, a few'

The proposed semantics also accounts for the fact that *kilka* can combine with cardinal suffixes. Since the number selected by the CF f is a natural number, it can serve as an argument for the cardinal suffixes, as defined in (43).

5.3.2 Deriving *ileś*

As discussed in §2.1, the indefinite *ileś* 'some number' is a complex expression involving a wh-word and the indefinite suffix *-ś*. In general, I assume that wh-elements denote properties. In this case, the wh-root $\sqrt{il}$- denotes a property of

type $\langle n, t \rangle$, namely a property of being a real number.[15] Furthermore, I adopt the view that indefinite suffixes in Slavic introduce a generalized CF of type $\langle \langle \tau, t \rangle, \tau \rangle$, see (45), which can attach to any wh-element to yield an indefinite expression (Yanovich 2005).

(45) $\lambda P_{\langle \tau, t \rangle}[f_{\langle \langle \tau, t \rangle, \tau \rangle}(P)]$

I propose that the composition of *ileś* proceeds as in Figure 5. The indefinite suffix *-ś* combines directly with the wh-root $\sqrt{il}$- so that the CF f yields a specific real number relative to a particular context.[16] Similar to *kilka*, the Numeral head not only assigns the gender value, but also introduces the classifier element. However, in this case it is not CARD but QUANT.

NumeralP$_{\langle e,t \rangle}$
$\lambda x[\mu(x) = f_{\langle \langle n,t \rangle, n \rangle}(\lambda n[\text{REAL}(n)])]$

Numeral
[NV]
QUANT$_{\langle n, \langle e,t \rangle \rangle}$
$\lambda n \lambda x[\mu(x) = n]$
-e-

n
$f_{\langle \langle n,t \rangle, n \rangle} \lambda n[\text{REAL}(n)]$

indef.suffix$_{\langle \langle \tau, t \rangle, \tau \rangle}$
$\lambda P[f(P)]$
-ś

$\sqrt{il}$-$_{\langle n,t \rangle}$
$\lambda n[\text{REAL}(n)]$
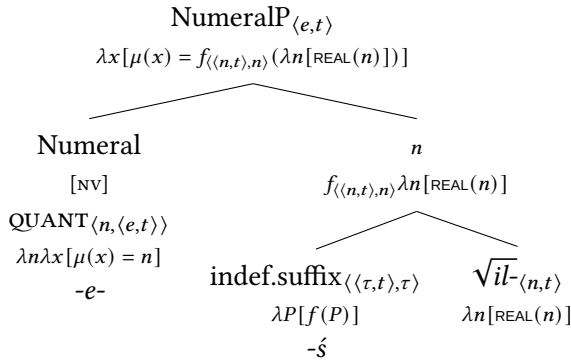
Figure 5: Derivation of *ileś* 'some, some number'

The type of the NumeralP is again $\langle e, t \rangle$ which does not allow *ileś* to refer to number concepts in clearly numeric environments. On the other hand, the contextually conditioned MF μ accounts for the fact that *ileś* is compatible both with count and mass terms. In the first case, it simply returns the number of atomic

---

[15]Arguably, it might be even a complex number. However, since I remain agnostic with respect to the question whether the concept of complex numbers is part of the semantics of natural language, I will stick to reals.

[16]The surface order of morphemes in Figure 5 is derived by (phrasal) movement of the root $\sqrt{il}$- to the left of the two functional heads *-e-* and *-ś*, which remain in the base order. As pointed out by an anonymous reviewer, this is not a frequent phenomenon, and it goes against traditional accounts of morpheme order based on head movement (e.g., Baker 1988), which would lead to a mirror-image order such as **il-ś-e*. However, the type of movement needed for Figure 5 has been argued independently to be necessary for various morpheme orders within words as well (e.g., Koopman & Szabolcsi 2000, Julien 2002; see also Caha 2017 for discussion).

individuals making up a plurality whereas in the latter it yields the amount of substance. Finally, the fact that the indefinite number is not necessarily an integer makes *ileś* incompatible with cardinal suffixes.

The last issue concerns how to ensure that the Numeral head gets the correct semantics in combination with a particular root, i.e., $\sqrt{kilk}$-, $\sqrt{il}$-, and cardinal roots such as $\sqrt{pięć}$-. For this purpose, I postulate the interface instructions as provided in (46).

(46) *Interpretation of the Polish Numeral head at LF*
 Numeral $\Leftrightarrow$ CARD / [__ [$\sqrt{kilk}$-]]
 $\Leftrightarrow$ QUANT / [__ [$\sqrt{il}$-]]
 $\Leftrightarrow$ $\varnothing$ / elsewhere

Given the standard elsewhere principle, the application of a specific operation overrides the application of a more general rule, and thus what happens at LF is as follows. The Numeral head is interpreted as CARD only in case it dominates the root $\sqrt{kilk}$- and as QUANT if and only if its complement is the root $\sqrt{il}$-. In all other cases, i.e., when Numeral combines with the cardinal root, it is semantically vacuous. The proposed mechanism guarantees adequate interpretations of the structures postulated for number-denoting cardinals and indefinite numerals in Figure 1, Figure 4, and Figure 5. Insertion of an additional null CARD node higher in the tree, see Figure 2, gives rise to a cardinal predicate which can be used as a nominal modifier and in predicate position.

6 Conclusion

In this paper, I presented novel data concerning the distribution as well as semantic properties of the Polish indefinite quantifiers *kilka* 'several, a few' and *ileś (tam)* 'some, some number'. Based on a number of tests, I concluded that such indefinites pattern with cardinal numerals rather than with vague quantifiers such as *mało* 'little, few'. Moreover, I posited that *kilka* and *ileś* should be treated as specific indefinites since they can have a 'referential' reading in an embedded clause, i.e., they can scope over a quantifier in a matrix clause. Therefore, I proposed that Polish indefinite numerals essentially share the core choice-functional semantics and argued that they should be analyzed as having a built-in classifier involving a measure function. The difference between the two results from the fact that *kilka* employs a cardinality function which is compatible only with atomic denotations and yields a value from the set of natural numbers $\{3, \ldots, 9\}$, whereas *ileś*

introduces a contextually conditioned measure function which, depending on a context, returns a real number corresponding either to a cardinality of a plurality or to a measure calibrated in relevant units.

Further research should focus on cross-linguistic investigations related to indefinite numerals both within Slavic and beyond as well as the behavior of FCIs such as *ilekolwiek* 'any number' in Polish. An open issue concerns the exact nature of the mapping between semantics and morphology in the case of the discussed indefinites from a typological point of view.

Abbreviations

BCS	Bosnian/Croatian/Serbian	V	virile
NCP	National Corpus of Polish	NV	non-virile
LF	Logical Form	CMPR	comparative
FCI	free choice item	SPRL	superlative
CF	choice function	GEN	genitive
MF	measure function	DISTR	distributivity marker
CARD	cardinal property	CL	classifier
QUANT	quantificational property	PL	plural

Acknowledgements

References

Aloni, Maria. 2007. Free choice, modals, and imperatives. *Natural Language Semantics* 15(1). 65–94. DOI:10.1007/s11050-007-9010-2

Alonso-Ovalle, Luis & Paula Menéndez-Benito. 2003. Some epistemic indefinites. In Makoto Kadowaki & Shigeto Kawahara (eds.), *NELS 33: Proceedings of the 33rd Annual Meeting of the North East Linguistic Society*, 1–12. Amherst, MA: GLSA Publications.

Babby, Leonard H. 1987. Case, prequantifiers, and discontinuous agreement in Russian. *Natural Language & Linguistic Theory* 5(1). 91–138. DOI:10.1007/BF00161869

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Bale, Alan C. & David Barner. 2009. The interpretation of functional heads: Using comparatives to explore the mass/count distinction. *Journal of Semantics* 26(3). 217–252. DOI:10.1093/jos/ffp003

Bańkowski, Andrzej. 2000. *Etymologiczny słownik języka polskiego*. Vol. 1. Warsaw: Wydawnictwo Naukowe PWN.

Błaszczak, Joanna. 2001. *Investigation into the interaction between the indefinites and negation*. Berlin: Akademie Verlag.

Borer, Hagit. 2005. *In name only*. Oxford: Oxford University Press.

Bylinina, Elizaveta. 2010. Depreciative indefinites: Evidence from Russian. In Gerhild Zybatow, Philip Dudchuk, Serge Minor & Ekaterina Pshehotskaya (eds.), *Formal studies in Slavic linguistics: Proceedings of Formal Description of Slavic Languages 7.5*, 191–207. Frankfurt am Main: Peter Lang.

Caha, Pavel. 2017. Suppletion and morpheme order: Are words special? *Journal of Linguistics* 53(4). 865–896. DOI:10.1017/S0022226717000196

Chierchia, Gennaro. 1985. Formal semantics and the grammar of predication. *Linguistic Inquiry* 16(3). 417–443. DOI:10.1016/0024-3841(78)90042-6

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford: Oxford University Press. DOI:10.1093/acprof:oso/9780199697977.001.0001

Corbett, Greville G. 1978. Universals in the syntax of cardinal numerals. *Lingua* 46(4). 355–368. DOI:10.1016/0024-3841(78)90042-6

Dočekal, Mojmír. 2012. Atoms, groups and kinds in Czech. *Acta Linguistica Hungarica: An International Journal of Linguistics* 59(1–2). 109–126. DOI:10.1556/ALing.59.2012.1-2.5

Dočekal, Mojmír. 2013. What do we count with numerals? Semantic analysis of Czech group-denoting and kind-denoting NPs. In Uwe Junghanns, Dorothee Fehrmann, Denisa Lenertová & Hagen Pitsch (eds.), *Formal Description of Slavic Languages: The ninth conference. Proceedings of FDSL 9, Göttingen 2011*, 87–105. Frankfurt am Main: Peter Lang.

Dočekal, Mojmír & Hana Strachoňová. 2015. Freedom to choose alternatives. 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*, 131–146. Frankfurt am Main: Peter Lang.

Dočekal, Mojmír & Marcin Wągiel. 2018. Event and degree numerals: Evidence from Czech. In Denisa Lenertová, Roland Meyer, Radek Šimík & Luka Szucsich (eds.), *Advances in formal Slavic linguistics 2016*, 77–108. Berlin: Language Science Press. DOI:10.5281/zenodo.2554021

Eremina, Olga. 2012. *The semantics of Russian indefinite pronouns: Scope, domain widening, specificity, and proportionality and their interaction*. Michigan State University, East Lansing, (Doctoral dissertation).

Fodor, Janet & Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5(3). 355–398. DOI:10.1007/BF00351459

Geist, Ljudmila. 2008. Specificity as referential anchoring: Evidence from Russian. In Atle Grønn (ed.), *Proceedings of Sinn und Bedeutung 12*, 151–164. Oslo: ILOS.

Gvozdanović, Jadranka. 1999. Some remarks on numeral morphosyntax in Slavic. In Jadranka Gvozdanović (ed.), *Numeral types and changes worldwide*, 187–196. Berlin/New York: Mouton de Gruyter.

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

Julien, Marit. 2002. *Syntactic heads and word formation*. Oxford: Oxford University Press.

Kadmon, Nirit & Fred Landman. 1993. Any. *Linguistics and Philosophy* 16(4). 353–422. DOI:10.1007/BF00985272

Kayne, Richard S. 2007. Several, few and many. *Lingua* 117(5). 832–858. DOI:10.1016/j.lingua.2006.03.005

Koopman, Hilda Judith & Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.

Kratzer, Angelika. 1998. Scope or pseudoscope? Are there wide-scope indefinites? In Susan Rothstein (ed.), *Events and grammar*, 163–196. Dordrecht: Kluwer.

Kratzer, Angelika & Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. In Yokio Otsu (ed.), *The proceedings of the Third Tokyo Conference on Psycholinguistics*, 1–25. Tokio: Hituzi Syobo. http://semanticsarchive.net/Archive/WEwNjc4Z/.

Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. In Renate Bartsch, Johan van Benthem & Peter von Emde Boas (eds.), *Semantics and contextual expression*, 75–115. Dordrecht: Foris Publications.

Landman, Fred. 2003. Predicate-argument mismatches and the adjectival theory of indefinites. In *From NP to DP: volume 1: The syntax and semantics of noun phrases*, 211–237. Amsterdam: John Benjamins.

Matthewson, Lisa. 1998. On the interpretation of wide-scope indefinites. *Natural Language Semantics* 7(1). 79–134. DOI:10.1023/A:1008376601708

Miechowicz-Mathiasen, Katarzyna. 2011. *The syntax of Polish cardinal numerals*. Manuscript, Adam Mickiewicz University. http://ling.auf.net/lingbuzz/001413.

Nouwen, Rick. 2010. Two kinds of modified numerals. *Semantics and Pragmatics* 3. Article 3:1–41. DOI:10.3765/sp.3.3

Pereltsvaig, Asya. 2008. Russian *nibud'*-series as markers of co-variation. In Natasha Abner & Jason Bishop (eds.), *WCCFL 27: Proceedings of the 27th West Coast Conference on Formal Linguistics*, 370–378. Somerville, MA: Cascadilla Proceedings Project. http://www.lingref.com/cpp/wccfl/27/paper1852.pdf.

Przepiórkowski, Adam. 2008. Generalised quantifier restrictions on the arguments of the Polish distributive preposition *po*. *Études Cognitives* 8. 159–177. https://www.ceeol.com/search/article-detail?id=93584.

Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy* 20(4). 335–397. http://www.jstor.org/stable/25001671.

Richtarcikova, Veronika. 2015. Epistemic indefinites in Slovak: Alternatives and exhaustification? 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*, 397–410. Frankfurt am Main: Peter Lang.

Rothstein, Susan. 2013. A Fregean semantics for number words. In Maria Aloni, Michael Franke & Floris Roelofsen (eds.), *Proceedings of the 19th Amsterdam Colloquium*, 179–186. Amsterdam: ILLC.

Rothstein, Susan. 2017. *Semantics for counting and measuring*. Cambridge: Cambridge University Press. DOI:10.1017/9780511734830

Safir, Ken & Tim A. Stowell. 1988. Binominal *each*. In James Blevins & Juli Carter (eds.), *NELS 18: Proceedings of the 18th Annual Meeting of the North East Linguistic Society*, 426–450. Amherst, MA: GLSA Publications.

Scha, Remko. 1981. Distributive, collective and cumulative quantification. In Jeroen A. G. Groenendijk, Theo M. V. Janssen & Martin B. J. Stokhof (eds.),

Formal methods in the study of language, 483–512. Amsterdam: Mathematisch Centrum.

Schwarzschild, Roger. 2002. The grammar of measurement. In Brendan Jackson (ed.), *SALT 12: Proceedings from the 12th Conference on Semantics and Linguistic Theory*, 225–245. Ithaca, NY: CLC Publications.

Šimík, Radek. 2015. Epistemic indefinites under epistemic modals 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, Leipzig 2013*, 425–442. Frankfurt am Main: Peter Lang.

Strachoňová, Hana. 2016. *Česká free choice indefinita*. Brno: Masaryk University. (Doctoral dissertation).

Sudo, Yasutada. 2016. The semantic role of classifiers in Japanese. *Baltic International Yearbook of Cognition, Logic and Communication* 11(1). 1–15. DOI:10.4148/1944-3676.1108

Testelets, Yakov & Elizaveta Bylinina. 2005. On sluicing-based indefinites. In Steven Franks, Frank Y. Gladney & Mila Tasseva-Kurktchieva (eds.), *Formal Approaches to Slavic Linguistics 13: The South Carolina Meeting 2004*, 355–364. Ann Arbor, MI: Michigan Slavic Publications.

Wągiel, Marcin. 2014. Boys, girls, and scissors: A semantic analysis of Polish NPs headed by the numeral *dwoje*. In Ludmila Veselovská & Markéta Janebová (eds.), *Nominal structures: All in complex DPs*, 69–84. Olomouc: Palacký University.

Wągiel, Marcin. 2015. Sums, groups, genders, and Polish numerals. 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*, 495–513. Frankfurt am Main: Peter Lang.

Wągiel, Marcin. to appear. Entities, events, and their parts: The semantics of multipliers in Slavic. In Peter Kosta & Teodora Radeva-Bork (eds.), *Current developments in Slavic linguistics: Twenty years after*. Frankfurt am Main: Peter Lang.

Yanovich, Igor. 2005. Choice-functional series of indefinite pronouns and Hamblin semantics. In Efthymia Georgala & Jonathan Howell (eds.), *SALT 15: Proceedings from the 15th Conference on Semantics and Linguistic Theory*, 309–326. Ithaca, NY: CLC Publications.   DOI:10.3765/salt.v15i0.2921

Chapter 15

Dative-marked arguments as binders in Polish

Jacek Witkoś

Adam Mickiewicz University in Poznań

Paulina Łęska

Adam Mickiewicz University in Poznań

Dominika Dziubała-Szrejbrowska

Adam Mickiewicz University in Poznań

This paper aims to account for peculiar binding properties of dative arguments in Polish: objects and dative object experiencers. Polish reflexive pronouns are (nominative) subject oriented, they can be bound by dative experiencers (Miechowicz-Mathiasen & Scheffler 2008; Witkoś 2007). At the same time, object experiencers, unlike nominative subjects, are also proper antecedents for both reflexive and pronominal possessives. This mixed behaviour poses a puzzle for traditional and novel formulations of binding theory, which assume complementarity between anaphors and pronominals and plainly states that the subject is the privileged binder in Slavic. We base our analysis on the concept of index raising, where the undifferentiated anaphoric/pronominal element is (covertly) moved and adjoined to v or T (Safir 2014; Nikolaeva 2014). The distribution of the two spellout forms of the anaphoric or pronominal elements is determined by its landing site and the case position of the binder.

Keywords: binding, psych verbs, dative antecedents, dative object experiencers, Polish

1 Introduction

A-binding has been a chief area of research in comparative linguistics since the early 80s (Chomsky 1981; 1986; Manzini & Wexler 1987; Belletti & Rizzi 1988; Rappaport 1986; Willim 1989; Burzio 1996; Hellan 1988; Progovac 1992; 1993; Avrutin 1994; among others) when the foundations for modern theory of binding were laid. It very soon became transparent that binding phenomena were subject to parametric differences involving such notions as the size of the binding domain, the morphology of the anaphoric element and the choice of the privileged binder. This paper touches upon the last aspect of the parametric difference, namely the strict subject orientation of anaphors in Polish (and other Slavic languages) as well as certain conditions which dative arguments must meet to qualify for binders; it turns out that even when dative arguments happen to be supreme arguments in particular structures, they do not fully mimic the behaviour of nominative subjects as binders.

Polish is a subject-oriented binding language, and objects, either dative or accusative-marked ones, cannot bind anaphors in other object/adjunct positions (Willim 1989; Reinders-Machowska 1991; Rappaport 1986 for almost identical data in Russian), as presented in (1).

(1) a. Jan$_1$ pokazał Marii$_2$ [{swoje$_{1,*2}$ / jej$_2$ / *jego$_1$} zdjęcie].
      Jan.NOM showed Maria.DAT self.POSS her   his   picture.ACC
      'Jan showed Maria his/her picture.'

   b. Jan$_1$ zawierzył córkę$_2$ [{swojej$_{1,*2}$ / jej$_2$ / *jego$_1$} patronce].
      Jan.NOM entrusted daughter.ACC self.POSS her   his   patron.DAT
      'Jan entrusted his daughter to his/her patron.'

   c. Jan$_1$ opowiedział Marii$_2$ [o   {sobie$_{1,*2}$ / ?niej$_2$} (samej) /
      Jan.NOM told   Maria.DAT about self   her   alone
      *nim$_1$].
       him
      'Jan told Maria about himself/her.'

Both the reflexive pronoun and the reflexive possessive seem to be oriented towards the nominative subject, while dative and accusative objects are infelicitous binders in (1).

In certain constructions referring to psychological states, dative arguments bind anaphoric pronouns but allow for optionality with pronominal/reflexive possessives.

(2) a. Marii$_1$ było żal {siebie$_1$ / *?jej$_1$} (samej).
     Maria.DAT was.3SG.N sorrow.3SG.M self.GEN her.GEN alone
     'Maria felt sorry for herself.'

  b. Marii$_1$ było żal {swojej$_1$ / jej$_1$} koleżanki.
     Maria.DAT was.3SG.N sorrow.3SG.M self.POSS her friend.GEN.F
     'Maria felt sorry for her female friend.'

The psychological predicate *podobać się* 'appeal to' shows variable behaviour: when bound, the possessive pronoun in the nominative argument is strongly preferred to the possessive reflexive, as in (3). However, Witkoś (2007; 2008) shows that a preverbal dative-marked argument can be involved in anaphoric binding into the nominative-marked constituent (cf. (4)):[1]

(3) Marii$_1$ spodobała się {?*swoja$_1$ / jej$_1$} nowa sukienka.
    Maria.DAT liked REFL self.POSS her new dress.NOM
    'Maria liked her new dress.'

(4) [Nowakom$_2$] spodobała się [nowa książka (Kowalskich$_1$) o
    Nowaks.DAT liked REFL new book.NOM Kowalskis.POSS about
    sobie$_{1,2}$]
    self
    'The Nowaks liked the new book (by the Kowalskis) about
    themselves/them.'

We address these issues by developing and updating an approach to binding based on Nikolaeva's (2014) INDEX RAISING (IR) and Despić (2013; 2015). In the view of the data in (2) and (4), our goal is to explain why dative antecedents in constructions with psychological verbs, (2b), allow for the option of binding both reflexive and pronominal possessives, while the nominative antecedent allows only for the reflexive possessive variant.[2] We submit that these different binding properties are due to different positions occupied by nominative and dative antecedents, namely SpecTP and Spec*v*P, respectively. We also claim that data such

---

[1] A detailed analysis of dative object experiencers of both verbal and non-verbal psychological predicates remains beyond the scope of this contribution. Recent analyses are presented in Jimenez-Fernandez & Rozwadowska (2016) and in Bondaruk (2017).

[2] We will not take into consideration reciprocal constructions in Polish, whose properties are markedly distinct from reflexives and identical to Russian reciprocals (Willim 1989; Reinders-Machowska 1991; Rappaport 1986 for Russian). For instance, in contrast to reflexives, reciprocals are not subject oriented and can be bound by the object as well. In terms of the IR-based analysis, reciprocal pronouns in Polish do not undergo IR.

as (3), though plentiful, are encumbered with an additional complicating factor in the form of the Anaphor Agreement Effect (AAE, Rizzi 1990; Progovac 1992; 1993; Woolford 1999; Reuland 2011) and they deserve a slightly different treatment. The most straightforward diagnostics for determining the binding potential of the dative argument involves cases when it binds (into) non-nominative elements (so (2) rather than (3)).

The paper is organized as follows. In §2, we present an outline of our theory of binding, with emphasis on our version of the IR theory articulated in Nikolaeva (2014), modified in line with Bošković (2005; 2012; 2013; 2014) and Despić (2011; 2013; 2015). §3 provides our account of binding in structures with dative arguments. We show why the dative argument of the ditransitive verb cannot bind reflexive elements, we analyse the position and binding option of the dative OBJECT EXPERIENCER (OE). §4 concludes the paper.

2 Components of the analysis

Our account of anaphoric binding in Polish follows from and draws from a triplet of sources: (A) approaches which stress the need for (covert) anaphor raising to some functional head position, usually Infl or T (Vikner 1985; Chomsky 1986; Pica 1987; 1991; Hestvik 1992; Avrutin 1994; Nikolaeva 2014), (B) approaches which stress the morphological impoverishment of the anaphoric elements (Burzio 1991; 1996; Safir 2014), and (C) approaches that recognize the notion of derived complementarity (Hellan 1988; Safir 2004; Boeckx et al. 2008).

(A) The most identity dependent form in (1), be it anaphoric, personal or possessive, is overlaid with lexical content late in the derivation, at Spell-Out. It is introduced into initial numeration as an undefined element, the most dependent form, called D-bound in Safir (2014), the index in Nikolaeva (2014), or root-pron in Heinat (2008). Safir (2014: 91–92) defines properties of D-bound/index in the following way:

(5) a. Always a variable: D-bound is the same object in SEM (the syntactic input to semantic interpretation) in all cases; it is interpreted as a bound variable regardless of its φ-features.

b. Always A-bound: the binder of D-bound (its antecedent) must c-command it from an A-position; that is, the D-bound form is A-bound. (We further narrow down the definition of A-position to the position where the antecedent has its case valued).

c. Always feature compatible: D-bound must be feature compatible with its antecedent (informally, this property may be termed antecedent agreement).

d. Spell-Out of the morphological shape of D-bound is potentially sensitive to whether A-binding is phase internal:
   – agreement compatible with morphological shape may be determined by phase internal factors locally distinct from antecedent agreement;
   – D-bound enters the derivation with φ-features arbitrarily assigned to it;
   – anywhere phase-internal shape is not required, D-bound receives default pronominal shape.

(B) The D-bound/index is impoverished in its feature composition, very much like the lexical anaphor in Polish, in that it has a [−var] feature.[3] The underspecification of this feature forces the index to move to a position where this interpretive impoverishment can be compensated for, in line with a similar procedure for semantically and morphologically deficient pronominal clitics in Béjar & Rezac (2003) and Franks (2017; Forthcoming).[4]

---

[3]The Polish reflexive pronoun and the reflexive possessive inflects for case but not for person, number and gender. The reflexive pronoun *siebie* 'self' also has a weak/clitic form *się* but we leave this issue aside in this paper.

[4]Franks (2017; Forthcoming) claims that clitics are deficient in three respects: prosodically, semantically and syntactically:

(i) The prosodic deficiency: Clitics cannot project prosodic feet. (Franks 2017: 147)

(ii) The semantic deficiency:
   a. Clitics cannot instantiate lexico-conceptual features.
   b. A clitic may not have [+person] features (either entirely or only subcomponents [Participant [Author]] of the 1st/2nd person).

(iii) The syntactic deficiency: Clitics cannot express syntactic complexity (they are heads).

In our analysis, the index does not show prosodic deficiency. Following Cardinaletti & Starke (1994) and Béjar & Rezac (2003), Franks (2017) proposes that [+person] must be licensed by entering into an agree relation with a functional category.

(C) The index moves from its thematic/case position to the head v/T, but it is not phonologically impoverished the way clitics are. This is why its movement forms a chain in which the copy is pronounced.[5],[6]

In short: we take the relation of binding to hold between the antecedent c-commanding D-bound/ index from its case position. The spell-out form of D-bound/index is determined by its movement to v/T.

Nikolaeva (2014), building on Chomsky (1986); Vikner (1985); Pica (1987; 1991); Hestvik (1992) and Avrutin (1994), proposes that the lexicalisation of D-bound/index depends on IR. We modify her original proposal as in Figure 1.

The diagrams in Figure 1 and Figure 2 show the placement of arguments in both the construction with ditransitive verbs and psychological predicates. In Figure 1 the direct object is the complement to V^0, the indirect object occupies SpecVP, from which it c-commands NP_1. The position of the dative experiencer in Spec$v$P corresponds to the dative in (2)–(4) with psychological predicates. As the diagrams above show, we assume two distinct positions for dative goals, bene- and malefactives (NP_2 in SpecVP) and dative OEs (Spec$v$P). We follow Larson (1988; 1990; 2014) for the placement of the former and Woolford (2006) for the placement of the latter. Two positions are reserved for the agentive subject: the

---

[5]A similar idea of an element raising (to the edge of the vP phase) and having its copy pronounced as reflexive is applied in an analysis of binding in German in Safir (2004) and in Lee-Schoenfeld (2008: 291). According to the latter source the licensing of *sich* 'self' co-indexed with *mother* requires covert movement:

(i) Die Mutter$_i$ lässt [$_{vP}$ die Kleine$_j$ {sich$_{?i/j}$ / ihr$_{i/*j}$} die Schokolade in den Mund
     the mother lets   the little-one self   her   the chocolate   in the mouth
     stecken].
     stick
     'The mother lets the little girl stick the chocolate in her mouth.'

[6]A reviewer for this volume expresses doubts as to whether a non-phonological clitic such as our D-bound/index should behave movement-wise like a clitic and pick the same landing site v^0/T^0. This reservation can be addressed in a number of ways. First, let us point out that in terms of their syntax non-clitic elements can be ambiguous between X^0/XP status and participate in head movement irrespective of their phonological properties; after all, clitic movement constitutes a subset of head movement. Second, in one of its multiple functions the Polish clitic *się* 'self' serves as the clitic form replacement of the reflexive pronoun *siebie* 'self'. Importantly, the distribution of this type of *się* 'self' fully overlaps with the distribution of clitic/weak pronouns and the span of the binding domain in Polish.

We claim that this overlap is not accidental but due to the same underlying operation: movement of D-bound/index and clitic/weak pronoun to the same functional head placed outside VP. Third, there are fruitful analyses of grammatical phenomena in Germanic (scrambling) and Romance languages, which link abstract (covert) clitic elements with overt non-clitic phenomena, such as Sportiche (1996) and the concept of "clitic voices".
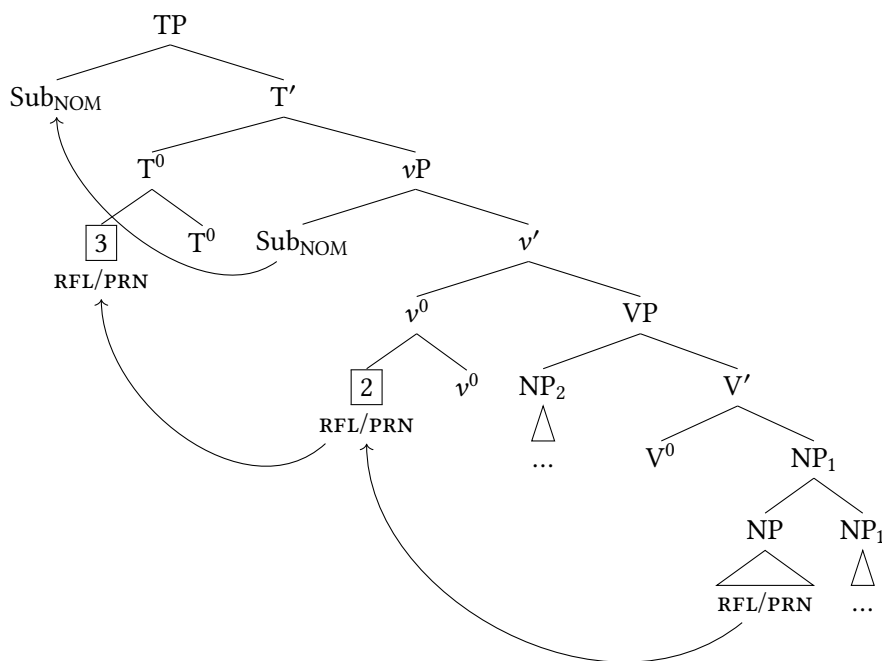
Figure 1: Index positions: a ditransitive predicate (the DOC type)

bottom of its A-chain in Spec$v$P and the top of its A-chain in SpecTP. The gist of the lexicalisation procedure is as follows (Nikolaeva 2014: 68):

(6) a. *Movement*: an index (marked as RFL/PRN in Figure 1–Figure 2) must undergo IR unless it is at a lexicalisation site or movement is no longer possible.

b. *Lexicalisation site*: an index is a sister to a node with label $D^0/v^0/T^0$ and is c-commanded by a specifier,

c. *Co-argumental Lexicalisation*: if an index is at a reflexivization site and is coindexed with a specifier which is its co-argument, the index has to be realized as reflexive.

d. *Lexicalisation at spell-out*: when the sentence is sent to spell-out, if an index is coindexed with a specifier of the projection to which it is adjoined, the index has to be realized as reflexive.

e. *Pronominal is an elsewhere condition*: if an index has not been realized as reflexive, it is realized as pronominal.
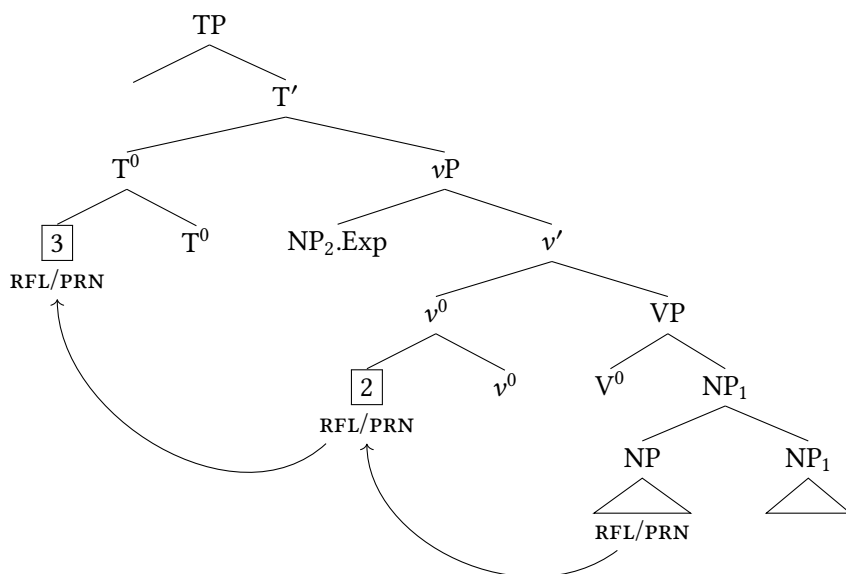
Figure 2: Index positions: A psychological predicate with ACC/DAT experiencer

As VP is not a lexicalisation site by definition, the overt position of the index (pronoun or anaphor) is mostly ignored in the calculation of its spell-out form.[7] IR is closely linked to ideas concerning clitic movement, see Sportiche (1996); Kayne (1985; 1991); Roberts (1992; 1993) in the GB tradition. Clause in (6e) clearly corresponds to the competition-based approach to binding, see Safir (2004), and the movement-based approach, see (Hornstein 2001; Boeckx et al. 2008), where the pronoun is the default 'elsewhere' option wherever the reflexive cannot be licensed.

---

[7]Exceptions include clause (6c) and co-argumental reflexivisation, where pronouns show not only strong anti-subject orientation but also anti-object orientation:
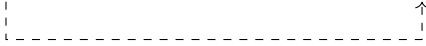
(i) a. * Mama$_2$ pokazała Marii$_1$ ją$_{1/2}$ (w lustrze).
          mother showed   Mary.DAT her.ACC in mirror
          Intended: 'Mother showed her to Maria (in the mirror).'

     b. * Mama$_2$ pokazała Marię$_1$ jej$_{1/2}$ (w lustrze).
          mother showed   Mary.ACC her.DAT in mirror
          Intended: 'Mother showed Maria to her (in the mirror).'

This issue remains beyond the scope of the current contribution but see Gogłoza et al. (Forthcoming) for a detailed analysis couched in the IR framework.

Let us sketch three derivations illustrating the mechanics of the system. Safir (2014) proposes the following derivation for an English example:

(7) a. John$_1$ praised {himself$_1$/*him$_1$/him$_2$}.

  b. [$_{vP}$ John [$_{v'}$ v [$_{VP}$ praise D-bound+3sg]]]

  c. [$_{TP}$ John [$_{T'}$ T [$_{vP}$ John [$_{v'}$ v [$_{VP}$ praise himself]]]]]
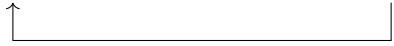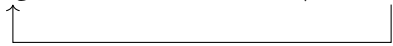
D-bound/index is merged in with unvalued φ-features assigned to it. *John* is the antecedent for D-bound/index and because *John* is the phase edge, the D-bound/index spells out in the shape indicating phase-internal dependency (the *-self* form in English). The major difference between Polish reflexive forms and the English ones is that the D-bound/index in Polish is impoverished in its feature composition, very much like the Polish lexical anaphor, in that it has an under-specified slot for [+φ] features. We take this underspecification to allow for the copying of the φ-features from the antecedent but not for their expression in situ. The expression of these features takes place only upon the movement of the D-bound/index to v and T (see Béjar & Rezac 2003 and Franks 2017 for a corresponding notion of clitic movement to a compensatory position for [+person] expression):

(8) a. Jan$_1$ zobaczył {siebie$_1$ /*jego$_1$}.
     Jan noticed   self.ACC   him.ACC
     'Jan noticed himself.'

  b. *Binding*
     [$_{vP}$ Jan$_1$ [$_{v'}$ [$_v$ +3sg$_1$ – v] [$_{VP}$ noticed D-bound[#φ]]]]

  c. *φ-expression*
     [$_{vP}$ Jan$_1$ [$_{v'}$ [$_v$ ^{ok}D-bound+3sg$_1$ – v] [$_{VP}$ noticed ~~D-bound~~[#φ]]]]

  d. *Spell-Out*
     [TP Jan$_1$ [T [$_{vP}$ ~~Jan~~ [$_{v'}$ [$_v$ ^{ok}D-bound+3sg$_1$ – v] [$_{VP}$ noticed self]]]]]

In (8a) the D-bound/index is bound in its base position (it copies the φ-features of its antecedent). In Witkoś et al. (Forthcoming), we treat A-binding as upward agree for feature [+variable], following Hicks (2009) rather than plain phi-feature copying. Here, the index meets Safir's (2014) condition of local antecedent agreement of (5c). The φ-features on D-bound/index cannot be expressed in its base

position and this morpho-syntactic deficiency forces the index to move to v^0 in (8b), forming a chain. At the point of spell-out of the vP D-bound/index is realized as the reflexive form *siebie* 'self' on the bottom copy of the chain in (8c).

The derivation of (1) follows a similar path. In (9c), D-bound/index is c-commanded by its antecedent and copies its φ-features under local antecedent agreement, yet it cannot express them, so it moves to v and forms a chain. The lexicalization of (covert) D-bound/index at v is determined by Nikolaeva's (6d) the NP content of the local Spec$v$P bears φ-features different from D-bound/index, so it is lexicalized as a pronoun at the bottom of its chain in (9e):

(9)   a.   Jan pokazał Marii jej zdjęcie.
          Jan showed Maria her picture
          'Jan showed Maria her picture.'

     b. $[_{vP}$ Jan.NOM$_1$ [showed $[_{VP}$ Maria.DAT$_2$ $[_{V'}$ V [her$_2$ picture]]]]

     c.   *Binding*
         $[_{vP}$ J$_1$ $[[_v$ +3sg.f2-v] showed $[_{VP}$ M$_2$ $[_{V'}$ V [D-bound[#φ]$_2$ [pic]]]]]

     d.   *φ-expression*
         $[_{vP}$ J$_1$ $[[_v$ +3sg.f2-v] showed $[_{VP}$ M$_2$ $[_{V'}$ V [D-bound[#φ]$_2$ [pic]]]]]

     e.   *Spell-Out*
         $[_{vP}$ J$_1$ $[[_v$ +3sg.f2-v] showed $[_{VP}$ M$_2$ $[_{V'}$ V [her$_2$ [pic]]]]]]

Interpretation-wise, the dative goal of a ditransitive verb can function as antecedent for the possessive in the accusative object, see (9d), but its case position is placed too low in the structure (it is VP-internal in a broad sense) to serve as a local antecedent for the index at the lexicalization site, see (9e).[8]

One of the consequences of IR is that the index moved via head movement and adjoined to v/T (positions [2] and [3]) should not c-command from the head adjoined position, as this would lead to undesirable principle C violations. While it is commonly believed that an adjunct to a maximal projection does c-command outside its adjunction host (see Kayne 1994 and subsequent work), there is less

---

[8]It seems that movement of the D-bound/index to a VP-external position is an inevitable step for any empirically adequate account of subject orientation, as it prevents one object from being antecedent of a possessive reflexive embedded in the other object. Even recent conceptually appealing accounts of binding (Reuland 2011; Zubkov 2018) take subject orientation for granted.

evidence for c-command following head adjunction. Nikolaeva (2014: 93–94): excludes this option by following the definition of c-command in Hestvik (1992: 574): "x c-commands y iff every node dominating x includes x and y, and x does not dominate y (where x includes y iff y is dominated by every segment of x, as proposed in May (1985)". Such a definition leaves the c-command domain of the adjunct undefined, as the node dominating the adjunct at the adjunction site does not include it. Citko et al. (2018) invokes the "word interpretation" notion from Chomsky (1995: 322) to prevent such unwelcome c-command: "at LF, X^0 is submitted to independent word interpretation processes WI, where WI ignores principles of the computational system within X^0". If c-command from within a complex head (a word) leading to a violation of binding principle C is such a "principle of the computational system" then it can be ignored.[9] Furthermore, Roberts (2009) develops a minimalist analysis of clitic climbing, to which IR corresponds, and observes that if clitics are taken to minimally constitute only the bundle of φ-features, moving them via excorporation from one head to another is very close to agree for φ-features.

We propose a particular structure for NPs including possessives, which captures ANTI-CATAPHORA EFFECTS (ACEs):

(10)   * Jan₁   pokazał jej₂   dyplom   koleżance Marii₂.
       Jan.NOM showed her.ACC diploma.ACC friend.DAT Maria's.DAT
       Intended: 'Jan showed her diploma to Marta's friend.'

R-expressions in Polish cannot be placed in positions following co-indexed pronouns, even if these pronouns apparently do not c-command them in an obvious manner. The grammar of Polish (as well as other Slavic languages) does not tolerate cataphoric relations. Despić (2011; 2013; 2015) develops an account of binding in Serbo-Croatian (SC) which relies to a large degree on the idea that adjectival possessives are adjuncts and therefore c-command outside the NP they are part of. In SC, the possessive c-commands from its adjoined position, on a theory of adjunction as in Kayne (1994), and thus causes a principle B effect, (12b), and a principle C effect, (12a), which does not occur in English examples, e.g. (11). Significantly, Polish shares with SC the fact that possessive pronouns trigger off

---

[9]Baker (1988) argues extensively that heads incorporated into other heads (where incorporation is a showcase example of head movement) cease to act upon elements they used to c-command before incorporation. So, head movement (incorporation) does not extend their c-domain, quite the contrary. For example, in Mohawk, the incorporated N no longer governs (under c/m-command) its possessor and does not license case on it, the verb as the incorporation host governs the possessor instead.

the ACEs (13a), although nominal possessives do not, see (13b), as discussed in Witkoś & Dziubała-Szrejbrowska (2015):

(11) a. His$_i$ latest movie really disappointed Kusturica$_i$.

     b. Kusturica$_i$'s latest movie really disappointed him$_i$.

(12) a. *Njegov$_i$ najnoviji film   je zaista razočarao   Kusturicu$_i$.
         his   latest   movie is really disappointed Kusturica.
         Intended: 'His new movie really dissapointed Kusturica.'

     b. *Kusturicin$_i$ najnoviji film   ga$_i$ je zaista razočarao.
         Kusturica's latest   movie him is really disappointed
         Intended: 'Kusturica's new movie really dissapointed him.'

(13) a. *Jego$_i$ siostra   bardzo pocieszyła   Janka$_i$.
         his   sister.NOM very   comfort.PAST Janek.ACC
         Intended: 'His sister comforted John very much.'

     b. Siostra   Janka$_i$ bardzo go$_i$ pocieszyła.
         sister.NOM Janek.GEN very   him.ACC comfort.PAST
         'Janek's sister comforted him very much.'

These authors conclude that Polish seems to employ two structures to represent nominal with possessives: the simpler bare NP-structure is used with pronominal and reflexive possessives, while the more complex structure involving possessive phrase and another functional projection (FP) on top of it is used with nominal possessives:

(14) a. [$_{NP}$ jego   [$_{NP}$ siostra]]
             his.GEN   sister.NOM
         'his sister'

     b. [$_{FP}$[$_{NP}$ siostra][$_{F'}$ F^0 [$_{PossP}$[$_{NP}$ Janka][$_{Poss'}$ Poss^0 [$_{NP}$ siostra]]]]]
             sister.NOM   Janek.GEN   sister.NOM
         'Janek's sister'

Nominals with pronominal possessors appear to be smaller, truncated versions of structures with nominal possessors.[10] The result is that only the pronominal

---

[10]In his analysis of English possessive constructions, Despić (2015) proposes a similar solution in that the pronominal possessor is placed at a lower level of the DP structure than the nominal possessor or the reciprocal possessor:

(i) [$_{DP}$ Mary/each other [$_{D'}$ [$_D$'s] [$_{PossP}$ my/their/her [$_{Poss'}$ Poss [$_{NP}$ friends]]]]]

possessives are expected to c-command outside the NP they modify, while nominal possessives do not. Significantly, the structure in (14a) has the following advantage: the pronominal c-commands outside its NP from its base position, the position where it has both its thematic role and case licensed, thus its A-position.[11]

3 Index raising in action

This section serves as an illustration of an application of the notion of IR to constructions with datives in Polish.

3.1 The VP-internal dative antecedents

The study of ditransitive structures has gained a lot of prominence in Slavic linguistics and the discussion has typically involved two problem areas. Initially, the assumption was that there was one underlying structure for all ditransitive constructions and much of the debate centred around the issue of the basic order between the ACC and DAT objects: e.g. Willim (1989), Witkoś (1998; 2007; 2008), Tajsner (2008), and Citko (2011) for Polish, and Franks (1995), Dyakonova (2007; 2009) for Russian, claimed that the DAT–ACC was the basic order, while Bailyn (1995; 2010; 2012) and Antonyuk (2015) argued for ACC–DAT as the basic order. The argumentation was based on such tests as genitive of negation, distributive *po* constructions, binding of reciprocals, licensing of secondary predicates, idiom formation, focus propagation, and VP topicalization.[12]

The second general approach was funded on the conviction that ditransitive verb constructions are derived from two basic underlying structures, one corresponding to English DOCs (V-/ACC), as in (15) and the other to the so called *to*-dative construction (V-ACC- (to) DAT), in ((16), after Dvořák (2010)):

(15)   [$_{vP}$ Jan showed [$_{ApplP}$ Maria.DAT Appl^0 [$_{VP}$ V^0 her picture.ACC]]]

(16)   [$_{vP}$ Jan subordinated [$_{VP}$ a page.ACC V^0 [$_{PP}$ P^0(to) his knight.DAT]]]

---

[11]A reviewer raises the question of the origin of the thematic role and case for the possessor, an adjunct in syntax which functions like an argument in LF. We follow Bošković (2005; 2012) and Despić (2011; 2013; 2015) in this regard and assume that the thematic role for the possessive as adjunct is determined compositionally at LF upon the transfer of the nominal phase (NP). Its genitive case is inherent, determined straightforwardly by the thematic relation. A plausible alternative leading to identical consequences for c-command relations, which we do not consider here, would be to posit movement of the pronominal possessor from within an extended projection of NP and adjunction to its outer edge, cf. Cegłowski (2017) for a recent analysis of the internal composition of the Polish NP.

[12]We refer the Reader to the above-mentioned sources for a detailed discussion of these tests.

In this regard, two positions can be outlined. One holds that particular verbs project only one of the two underlying structures and further alternative word order permutations operate on them (Dvořák 2010). Other authors argue that all benefactive/recipient verbs can appear with any underlying structures (Gračanin-Yuksek 2006; Marvin & Stegovec 2012). The criteria used in distinguishing between the two construction types involve the obligatory presence of the dative argument (17), causative reading (18), VP-topicalization, nominalisation, quantifier scope, and the two-goal construction. For lack of space, we illustrate only a few of these tests. Our Polish examples are based on the examples given in Gračanin-Yuksek (2006), Dvořák (2010), and Marvin & Stegovec (2012) for other languages.

(17) a. Jan wysłał (Marii) paczkę już wczoraj.
Jan sent Maria.DAT package.ACC already yesterday
'Jan sent Maria a package yesterday.'

b. Jan powierzył Marię *(jej patronowi) już wczoraj.
Jan entrusted Maria.ACC her patron.DAT already yesterday
'Jan entrusted Maria to her patron yesterday.'

(18) a. Beethoven dał światu Czwartą Symfonię.
Beethoven gave world Fourth Symphony
'Beethoven gave the world the Fourth Symphony.'

b. # Beethoven dał Czwartą Symfonię światu.
Beethoven gave Fourth Symphony world
'Beethoven gave the Fourth Symphony to the world.'

From our perspective, all the above mentioned ditransitive constructions show a crucial property, namely the superior object cannot function as an antecedent for the reflexive possessive in the other object; it can only antecede a pronominal possessive:[13]

---

[13]It must be noted that this conclusion does not hold for all Slavic languages. For example, Marvin & Stegovec (2012) show that in Slovenian, a quantifier in the higher dative object can bind a reflexive possessive in the lower object, as in (i).

(i) Tat$_2$ je vrnil [vsakemu oškodovancu]$_1$ [svoj$_{1,2}$ avto].
thief$_j$ AUX return.PAST each.DAT victim.DAT his.ACC car.ACC
'The thief returned every victim his car.'/ 'The thief returned every victim his (the thief's) car.'

(19) a. Jan$_1$ pokazał Marii$_2$ {swoje$_{1/*2}$ / jej$_2$} zdjęcie.
         Jan showed Maria.DAT self.POSS   her picture.ACC
         'Jan showed Maria her picture.'

     b. Jan$_1$ pokazał Marię$_2$ {swojej$_{1/*2}$ / jej$_2$} przełożonej.
         Jan showed Maria.ACC self.POSS   her supervisor.DAT
         'Jan showed Maria to her supervisor.'

     c. Król$_1$ podporządkował giermka$_2$ {swojemu$_{1/*2}$ / jego$_2$} rycerzowi.
         king subordinated   page.ACC self.POSS   his   knight.DAT
         'The king subordinated the page to his knight.'

     d. Król$_1$ podporządkował rycerzowi$_2$ {swojego$_{1/*2}$ / jego$_2$} giermka.
         king subordinated   knight.DAT self.POSS   his   page.ACC
         'The king subordinated his page to the knight.'

This leads us to propose that in both patterns singled out, the Spell-Out form of the index in both constructions is determined by the fact that VP is not a reflexivization domain/site. So, both the accusative and the dative object of a regular ditransitive verb is placed too low in the structure to serve as a co-indexed antecedent for the index at the reflexivization site, defined as vP or TP, but not VP, see positions [2] and [3] in Figure 1 and Figure 2. As soon as the pronoun is not a co-argument of the object, IR applies and carries the index to the domain of vP/TP, out of the c-domain of the object, so despite their coindexation, no Condition B violation occurs:

(20) [$_{v\text{P}}$ Maria.NOM$_1$ [index$_2$-v showed [$_{\text{VP}}$ Jan.DAT$_2$ [$_{\text{V}'}$ V [index$_2$ [pictures]]]]]

Interestingly, the dative of possession seems to behave like a regular VP-internal dative object. Polish has a construction where the dative-marked nominal represents the thematic role of possessor, correctly captured in the English translation:

(21) a. Jan$_1$ złamał Tomkowi$_2$ {*swoją$_2$ / jego$_2$} ulubioną
         Jan.NOM broke Tomek.DAT self.POSS his   favourite
         kredkę.
         colour-pencil.ACC
         'Jan broke Tomek's favourite colour pencil.'

     b. Maria$_1$ wybiła   Tomkowi$_2$ {*swoją$_2$ / jego$_2$} nową złotą
         Maria.NOM knocked out Tomek.DAT self.POSS his   new golden
         plombę.
         filling.ACC
         'Maria knocked out Tomek's new golden filling.'

The fact that only the possessive pronoun is the correct co-indexed bound form indicates that dative of possession is placed only as high as SpecVP, as NP_2 in Figure 1, and the index is raised to attach to v/T, outside of its c-command domain:

(22) $[_{vP}$ Jan$_1$ index$_2$-v broke $[_{VP}$ Tomek.DAT$_2$ $[_{V'}$ V $[_{NP}$ index$_2$ [favourite pen]]]]]

In general, if accepted, our analysis can be used as a detector for the position in which a given antecedent is placed with respect to heads v/T; any antecedents placed below v, so within VP, are predicted not to be able to bind reflexive pronouns/reflexive possessives.

In this context, consider an example of the impersonal passive construction with a dative argument:

(23) Marii   pokazano   {*swoją   / jej} nową koleżankę.
      Maria.DAT shown.IMPRS self.POSS her new friend.ACC
      'Maria was shown her new friend.'

Despite the fact that the dative argument is placed in the left peripheral position in the clause and on many analyses, it occupies SpecTP, it can only function as antecedent to a pronominal possessive. We take this fact to indicate that the case position of this argument is really low, probably SpecVP, as any ordinary dative object of a ditransitive construction and its movement to T does not extend its binding domain, see (5b).[14] At the same time, the word order in (23) does not convey any information-structure related message and it can be used as an answer to a general 'what has happened?' question. So a position in SpecTopP or SpecFocP is not an option. We assume that the dative NP in (23) is either in SpecTP on account of checking only the [+EPP] property of T, which is not sufficient to extend its binding domain, or it is moved to a position that is technically an A-position but, crucially, not a case position, as proposed in Germain (2015) and Citko et al. (2018):[15]

---

[14]The same conclusion is reached in Moore & Perlmutter (2000) for Russian impersonal passives.

[15]Germain (2015) proposes that conflicting characteristics of this position find a natural explanation if feature inheritance is split and the phase head C (Fin in her account where Rizzi's (1997) split CP architecture is assumed (i). The head Fin passes on only φ-features to T but retains the [+EPP] property. Hence the nominative case can be valued under agree on the postverbal DP, while the non-nominative DP can move up to SpecFinP to satisfy the EPP-property.

(i) $[_{ForceP}$ Force $[_{TopP}$ Top $[_{FocP}$ Foc $[_{FinP}$ Fin]]]]

(Russian left periphery; Germain 2015: 428)

(24) [$_{\text{TP/FinP}}$ Maria.DAT (Fin) [$_{\text{TP}}$ index-T shown [$_{\text{VP}}$ ~~Maria.DAT~~ [$_{\text{V}'}$ V [$_{\text{NP}}$ index [new friend]]]]]]

3.2 The medial domain: Dative OEs in Spec*v*P

In this section we investigate (both verbal and non-verbal) dative OEs which bind anaphoric pronouns as co-arguments and optionally possessive reflexives as non-co-arguments. The successful antecedents to anaphoric pronouns are all placed in a clausal position higher than VP.

Psychological predicates with dative experiencers fall into two classes: non-verbal predicates and verbal ones. The chief source of differences between them in terms of binding properties of their dative arguments stems from the fact that only the latter allow for nominative T/SM (Target/Subject Matter) arguments and binding into these shows considerable speaker variation.

3.2.1 OEs in non-verbal psychological predicates

We start with non-verbal psychological predicates such as *było żal* 'was sorrow' or *było wstyd* 'was shame'. In (25), the anaphoric/pronominal object (the index) is the object of the predicate *żal* 'pity', so a co-argument of the dative experiencer. IR carries it to the *v*-adjoined position and no further, see (6c). This position is c-commanded by the dative NP:

(25) a. Marii$_1$ było   żal   {siebie$_1$ / *?jej$_1$} (samej).
          Maria.DAT was.3.SG.N sorrow.3.SG.M self   her   alone
          'Maria felt sorry for herself.'

  b. [$_{v\text{P}}$ Maria.DAT [$_{v'}$ index-*v* was [sorrow index]]]

In (26) the index is free to either head-adjoin to *v* or move on to head-adjoin to T, as it is not a co-argument to Maria. In the former case, clause (6d) forces lexicalisation as reflexive, in the latter, clause (6e) forces lexicalisation as pronominal:

(26) a. Marii$_1$ było   żal   {swojej$_1$ / jej$_1$} koleżanki.
          Maria.DAT was.3.SG.N sorrow.3.SG.M self.POSS   her   friend.SG.F.GEN
          'Maria felt sorry for her female friend.'

  b. [$_{\text{TP}}$ index-T [$_{v\text{P}}$ Maria.DAT [$_{v'}$ index-*v* was [sorrow [index friend].GEN]]]]

The two derivations above markedly differ from equivalent constructions with nominative subjects and a corresponding verbal predicate *żałować* 'feel pity' in a predictable manner:

(27) a. Maria₁   żałuje   siebie₁ (samej).
          Maria.NOM feels-pity self.GEN alone
          'Maria feels pity for herself.'

     b. Maria₁   żałuje   {swojej₁ / *jej₁} koleżanki.
          Maria.NOM feels-pity self.POSS   her friend
          'Maria feels pity for her friend.'

In (27a)–(27b), the subject occupies SpecTP, the highest A-position in the clause, so in both corresponding derivations the index must be spelled-out as a reflexive:

(28) a. [TP Maria.NOM index-T [vP Maria.NOM [v' index-v [feels-pity index]]]

     b. [TP Maria.NOM index-T [vP Maria.NOM [v' index-v [feels-pity [index friend]]]

Furthermore, other verbs with dative OEs can function as antecedents to both reflexive pronouns and reflexive possessives, for instance in selected PPs:

(29) Marii   nudziło   się w {swoim / jej} domu.
       Maria.DAT bored.3.SG.NEUT REFL in self's   her home.LOC
       'Maria was bored at home.'

Additionally, other factors support the idea of a higher placement in the clausal structure of dative OEs in comparison with dative goals/benefactives, for instance the applicative characteristics that the predicates with dative OEs display. Cuervo (2003) argues that the dative argument of *gustar*, the Spanish equivalent to *podobać się* 'appeal to' seems to be licensed by a high applicative in the sense of Pylkkänen (2002). This is because the nominative argument is not involved with it in any relation of possession or location which are typical of 'low' applicatives in (30a)–(30b), where Maria's habitual possession of the pen is implied, so (30b) can only mean that the pen was Mark's and it was in Maria's possession only temporarily. No possession or location is implied in (31), where the fancy of the pen (whosoever it is) has overcome Maria:

(30) a. Marii   złamał się długopis.
           Maria.DAT broke REFL pen.NOM
           'Maria broke a pen.'

     b. # Marii   złamał się długopis Marka.
           Maria.DAT broke REFL pen.NOM Mark.GEN
           'Maria broke Mark's pen.'

(31) a. Marii   podoba się   długopis Marka.
          Maria.DAT appeal   REFL pen.NOM Mark.GEN
          'Mark's pen appeals to Maria.'

      b. Marii   było   żal   długopisu Marka.
          Maria.DAT was.3.SG.NEUT sorrow pen   Marek.GEN
          'Maria was sorry about Mark's pen.'

In the context of the solution proposed here, the dative argument corresponding to the 'high' applicative is placed in Spec *v*P, while the one in the 'low' applicative in SpecVP. It appears that (31) stands apart from (30).

It is therefore more worrying that at first glance the binding capacity of verbal psychological predicates runs counter to what has been presented thus far.

3.2.2 Verbal psychological predicates and the idiosyncrasy of *podobać się* 'appeal to'

The most frequently researched verbal psychological predicate in Polish is *podobać się* 'appeal to' (see Miechowicz-Mathiasen & Scheffler 2008; Bondaruk & Szymanek 2007; Żychliński 2013; Jimenez-Fernandez & Rozwadowska 2016; Bondaruk et al. 2017). Its distinctive property is the fact that it selects for the experiencer in dative and the cause/target of emotion in nominative. It has also been noticed that the binding potential of its dative-marked argument differs from the dative of non-verbal psych-predicates:[16]

(32) Maria$_1$ podobała się sobie$_1$ w lustrze.
      Maria.NOM appealed REFL self.DAT in mirror
      'Maria appealed to herself in the mirror.'

---

[16]Franks (1995: 253) observes this contrast for Russian:

    (i)   Mne   zal'   sebja.
           me.DAT sorry self.ACC
           Intended: 'I feel sorry for myself.'

    (ii) * Mne   nadoedaet svoj ucebnik.
           me.DAT bore.3.SG   self's textbook.NOM
           Intended: 'My textbook bores me.'

    (iii) * Mne   dosazdaet svoj brat.
           me.DAT vex.3.SG   self's brother
           'My brother vexes me.'

(33) Maria₁   podobała się {swojemu₁ / *jej₁} koledze   z   ławki.
      Maria.NOM appealed REFL self.POSS   her friend.DAT from school.desk
      'Maria appealed to her school desk friend.'

(34) Marii₁   spodobała się nawet ?%ona₁   {(sama) / *sobie₁} (sama)
      Maria.DAT appealed REFL even   she.NOM alone   self.NOM alone
      w lustrze.
      in mirror
      'Even herself in the mirror appealed to Maria.'

(35) Marii₁   podobał się {*swój₁ / jej₁} kolega   z   ławki.
      Maria.DAT liked   REFL self's her friend.NOM from school.desk
      'Her school desk friend appealed to Maria.'

(36) Janowi₁ spodobały się listy   od   {swoich₁ / jego₁} fanek.
      Jan.DAT liked   REFL letters from self's   his   fans
      'Letters from his fans appealed to Jan.'

(37) Mi₁ się   swój₁ głos podoba.
      Me REFL my   voice like
      'I like my voice.'
              (Miechowicz-Mathiasen & Scheffler 2008: 107, ex. (62); corpus search)

In general, if the dative Experiencer of a psychological verb is in Spec*v*P, we ex-
pect to see optionality with the pronominal vs reflexive possessive, similar to
that with non-verbal psych-predicates, but this is not the case.[17] The data pie
can be partitioned into three uneven sections. Most native speakers asked for
judgements on the data prefer for the dative Experiencer to bind pronominal
possessives, see (35). Quite a few allow the dative Experiencer to bind a posses-
sive reflexive but only in contexts where the reflexive is embedded deep in the
nominative constituent and bears a different case.

---

[17]Other similar verbs in Polish are *dokuczać* 'tease, vex', *nudzić* 'bore', and *szkodzić* 'harm':

  (i) Marii₁   dokuczał {*swój₁ / jej₁} kolega   z   ławki.
       Maria.DAT teased   self's her friend.NOM from school.desk
       'Her school desk friend teased to Maria.'

  (ii) Janowi₁ szkodziły listy   od   {?swoich₁ / jego₁} fanek.
       Jan.DAT harmed   letters from self's   his   fans
       'Letters from his fans harmed Jan.'

These inconclusive data lead to conflicting views on the position of the dative OE and the position of the nominative argument. Jimenez-Fernandez & Rozwadowska (2016) assume that its A-position is in SpecVP (and the preverbal position is in an articulated CP area). Bondaruk & Szymanek (2007), Tajsner (2008), and Bondaruk et al. (2017) propose that the dative experiencer is in Spec$v$P, as its binding scope is different from dative goals. Miechowicz-Mathiasen & Scheffler (2008) claim that the dative Experiencer reaches as high as SpecTP:

For the sake of concreteness, we assume that the dative Experiencer occupies the position of Spec$v$P, though this view is not uncontroversial:[18]

(38) [$_{TP}$ index$_i$-T [$_{vP}$ DAT$_i$ [$_{v'}$ index$_i$-v [$_{VP}$ V [NOM index$_i$ N$_k$]]]]]]

Let us tentatively assume that the structure in (38) is correct for the Polish *podobać się* 'appeal to', with the nominative theme argument optionally raised to its case position in SpecTP in overt syntax.[19] The index is c-commanded by the dative OE in the v-adjoined position (corresponding to position [2] in Figure 2) but it is not so c-commanded when placed in the T-adjoined position (corresponding to position [3] in Figure 2).

With this general idea in mind, the derivations of (32)–(36) look as follows, with the dative experiencer fronted to an A-position in line with Germain (2015), see (35):

(39) [$_{FinP}$ Maria.DAT Fin [$_{TP}$ index-T [$_{vP}$ ~~Maria.DAT~~ [$_{v'}$ index-v [$_{VP}$ appealed REFL-V [even she.NOM/*self$_{*case}$ (self/alone)] in mirror]]]]]]

(40) [$_{FinP}$ Maria.DAT Fin [$_{TP}$ index-T [$_{vP}$ ~~Maria.DAT~~ [v' index-v [$_{VP}$ appealed REFL-V [?*self's/her school friend]]]]]]

(41) [$_{FinP}$ Jan.DAT Fin [$_{TP}$ index-T [$_{vP}$ ~~Jan.DAT~~ index-v [$_{VP}$ appealed REFL-V [$_{NP}$ [$_{NP}$ letters] [$_{PP}$ from [$_{NP}$ self's/his fans]]]]]]].

---

[18]For instance, Cuervo (2003) argues strongly for the view that the dative OE is placed in a higher position, as the nominative Theme occupies Spec$v$P:

(i) [$_{TP}$ index$_i$-T [$_{ApplP}$ DAT [$_{Appl'}$ APPL^0 [$_{vP}$ index [$_{v'}$ v^0 [NOM index] [$_{v'}$ v-BE [$_{VP}$ PSYCH VERB]]]]]]]]

We cannot discuss this issue in full for lack of space.

[19]An analogous structure is proposed in Klimek & Rozwadowska (2004).

(27) is easy to deal with, as Polish has no nominative reflexive pronouns. Due to lack of this form in the morphological paradigm, its closest equivalent is selected, in line with Safir's (2004) FORM TO INTERPRETATION PRINCIPLE (FTIP).[20]

(40) seems to be a problem indeed, but an unacceptable status of the reflexive possessive can be credited to what Rizzi (1990: 26) calls the ANAPHOR AGREEMENT EFFECT (AAE):[21]

(42)   Anaphors do not occur in syntactic positions construed with agreement.

Our discussion of Polish data reveals that consequences of the AAE are subject to considerable language variation: nominative reflexive possessives are typically avoided by most speakers, although they are construed with agreement only indirectly: they agree (in case and φ-features) with NP they modify and this NP agrees with the auxiliary/verb. Yet, the structure we propose for pronominal possessive NPs is shown in (43). It is only natural to extend it to cases of reflexive possessives:[22]

---

[20]FTIP, Safir (2004): If:

   a. X c-commands position Y,

   b. z is the lexical form or string that fills Y,

   c. w is a single form more dependent than z,

   d. both w and z could support the same identity-dependent interpretation if Y were exhaustively dependent on X, then (the referential value for) Y cannot be interpreted as identity dependent on X.

[21]Rizzi (1990: 32–33) reports the following contrast in Italian: a dative experiencer can bind an anaphor as long as it is not nominative, so since *importare* 'matter' takes a genitive theme, this theme can be bound, while a nominative argument of *interessare* 'matter' cannot:

   (i)   A loro   importa solo di se stessi.
          to them matters only of themselves
          'They matter only to themselves.'

   (ii)   * A loro   interessano solo se stessi.
          to them interest   only themselves
          Intended: 'They have interest only in themselves.'

Significantly, however, dative experiencers can function as binders once the AAE is controlled for, as in (i). The same picture obtains with Polish dative experiencers above.

[22]Note that the structure in (43) with a pronominal possessive is much less ambiguous than the one with the reflexive in (44) on account of the pronominal possessive bearing a different case (genitive) from the nominative of the NP it modifies.

(43) [NP jego   [NP siostra]]
          his.GEN   sister.NOM
      'his sister'

(44) [NP swoja   [NP siostra]]
          self.POSS   sister.NOM
      'self's sister'

This structure may be quite ambiguous when the AAE applies, as the possessive element is equidistant to T with the NP it modifies.

(45) Janowi nie spodobała się   {*swoja   / jego} siostra.
      Jan.DAT not appleal   REFL self.POSS his   sister.NOM
      'His sister did not appeal to Jan.'

(46) T.AGR$_{2/1}$ … Jan.DAT$_1$ … [NP self.NOM$_1$ [NP sister.NOM$_2$]]

The equidistance relationship in question may cause confusion as to what really agrees with Infl/T here, the modified NP (with no consequence for the AAE) or the possessive reflexive (violating the AAE in (44) and (45) above). In the latter case, from the perspective of the Binding Principles, the possessive forces its referential subscript to represent the subscript of the entire NP.[23] Now, this is quite similar to what Landau (2000: 109–111) observes for cases of Obligatory Control, where the controller (unexpectedly) does not c-command PRO but constitutes the specifier of a c-commanding DP:

(47) It would help Bill's$_1$ development [PRO$_1$ to behave himself$_1$ in public]

Landau proposes that a well-defined class of nouns denoting abstract notions reflecting the individuality of the controller ([X's NP]) allows for what he calls the logophoric extension of X:

(48) For the purpose of control, a logophoric extension [X's NP] is
      non-distinct from X: [X's$_1$ NP] → [X's NP]$_1$.

Thus, logophoric extension is a selective process that affects only one module of grammar and one aspect of interpretation: Control Theory. We would like to submit that an analogous process of reanalysis affects the adjunct/specifier structure:

---

[23]In languages where possessives are genuine specifiers rather than adjuncts, possessive reflexives are allowed, as shown in Woolford (1999: 273–274).

(49)   *Extended AAE:*
       Anaphors do not occur in syntactic positions construed with agreement
       directly (a) or indirectly (b):

   a.   Nominative anaphors do not exist in languages showing
        subject/verb agreement;
   b.   For the purpose of binding, an indexical extension [X's NP] is
        non-distinct from X:
        $[_{NP2}$ swój.NOM$_1$ $[_{NP2}$ name.NOM$_2$]] $\rightarrow$ $[_{NP1}$ swój.NOM$_1$ $[_{NP1}$
        name.NOM$_2$]]

Our notion of indexical extension differs from Landau's original on two counts:
first, it is not limited by the semantic (sub)class of N and second, it depends on
the structural position of X, which we have shown to act as an adjunct, following
Despić (2011; 2013; 2015).

But this does not seem to be enough to cover the whole spectrum of the data.
First, notice that index extension may be less local in the cases discussed by Lan-
dau. For instance, the controller for Obligatory Control PRO can also be placed
in a position embedded in a measure NP selecting for the 'logophoric NP':

(50)   ? It considerably helped $[_{NP^1}$ first stages of $[_{NP^2}$ her$_1$ music career]] [PRO$_1$
       to have an uncle in a record company]

So, it seems that (at least for some speakers) X from (50) need not be very close
to the edge of the NP to propagate its index to the maximal NP (here NP^1).

Once we allow for the less local propagation of the index in the cases of in-
dexical extension in definition (49) above, we can account for (4) above on the
assumption that the rule of the Extended AAE is subject to graded speaker vari-
ation:

(51)   * antecedent$_i$ ... $[_{AgrP}$ anaphor$_i$ agreement$_i$...]

(52)   For the purpose of binding, an indexical extension [X's NP] is
       non-distinct from X:
       $[_{NP2}$ self.NOM$_1$ $[_{NP2}$ name.NOM$_2$]] $\rightarrow$ $[_{NP1}$ self.NOM$_1$ $[_{NP1}$ name.NOM$_2$]]

(53)   For the purpose of binding, an indexical extension [X's NP] is
       non-distinct from X:
       $[_{NP3}$ N$_3$... $[_{NP2}$ self.NOM$_1$ $[_{NP2}$ name.NOM$_2$]]] $\rightarrow$ $[_{NP3}$ N1... $[_{NP1/2}$ swój.NOM$_1$
       $[_{NP1/2}$ name.NOM$_2$]]]

All speakers of Polish have (51) in their grammars, most speakers have (52) as a part of their grammars and exclude nominative reflexive possessives as a result of indexical extension, while the most conservative ones allow for non-local indexical extension and disallow reflexive possessives in cases other than nominative if they are embedded in nominative NPs, see (53). [24]

4 Concluding remarks

In the process of our investigation, we have raised a number of questions with respect to the data in (1)–(4). We conclude that, universally, there is one D-bound/index which is the most dependent form bound locally and lexicalized in two different forms: reflexive and pronominal, determined by IR. When the co-agreeing NP locally c-commands D-bound/ index in its landing site at v/T, it is spelled out as a reflexive form, otherwise it is spelled out as a pronoun. The chain of Index Raising exhibits copy pronunciation, i.e. the tail of the chain is pronounced. In Polish (Slavic), IR is driven by the need to compensate for impoverished [+person] feature on the D-bound/index. The subject orientation of reflexive pronouns and possessive reflexives comes out rather naturally in this account. As IR places the index in these positions, it is not surprising that pronouns and anaphors show complementary distribution only with respect to the subject but not the object. The picture becomes even more transparent when we take into account the distinction between co-argument and non-co-argument reflexivization, see (6c) vs. (6d)–(6e). The non-co-argument index covertly raised beyond the c-domain of the object is predicted to be spelled out as a pronominal possessive, although it is co-indexed with the object c-commanding it in overt syntax. We have shown that successful binders of reflexives and reflexive possessives in Polish need not occupy the position of SpecTP, which is reserved for nominative subjects only. Dative OEs occupy a lower position of Spec$v$P. [25] In view of the scope and reach of IR, these elements can bind and be co-indexed

---

[24]A similar effect arises for the ACE. Willim (1989: 82) reports that the following example is problematic, though many native speakers accept it as only mildly deviant:

    (i)   % Ta recenzja   książki mojego brata$_1$ zupełnie   go$_1$ załamała.
           this review.NOM of-book my   brother's completely him.ACC devastated
           'This review of my brother's book devastated him completely.'

A reviewer for this volume raises the issue of how the propagation of the index can be constrained. We presume that it is a matter of speaker variation but the extent of the propagation is difficult to gauge on accout of processing difficulties. Certainly, this issue deserves further empirical study.

[25]For arguments to this effect also see Citko et al. (2018).

with reflexives and reflexive possessives adjoined to v but they can also be co-indexed with pronominal possessives adjoined to T. The optionality of possessive forms co-indexed with them is thus explained.[26] We have also credited imperfect results of dative OEs binding into nominative themes to Extended AAE of Rizzi (1990) caused by a specific placement of possessives as adjuncts at the edge of the nominative NP. Such placement leads to ambiguity of representation and the probe/goal relations involving T and NP.NOM.

Abbreviations

1/2/3	first/second/third person	INF	infinitive
AAE	anaphor agreement effect	IR	index raising
ACC	accusative	LOC	locative
ACE	anti-cataphora effects	NEUT	neuter
AUX	auxiliary	NOM	nominative
CL	clitic	OE	object experiencer
COP	copula	PAST	past tense
DAT	dative	POSS	possessive
DOC	double object construction	PRES	present tense
F	feminine	REFL	reflexive clitic
FTIP	form to interpretation principle	SG	singular
GEN	genitive	SC	Serbo-Croatian
IMPRS	impersonal		

---

[26]Introduction of more structural content which does not block IR, specifically PRO and infinitive T with raising and control constructions, multiplies reflexivization sites and provides for more spell-out options for the index:

(i) Maria$_1$ kazała Piotrowi$_2$ patrzeć   na {siebie$_{1,2}$ / *niego$_2$ / nią$_1$}.
Maria.NOM told   Piotr.DAT look.INF at   self   him   / her
'Maria told Piotr to look at himself/her.'

The infinitive complement in (i) constitutes a binding domain for a co-argument index of PRO. Hence the co-argument must stop moving at the vP level and is spelled-out as a reflexive (this is interpretation with index$_2$, as PRO is controlled by the object). The pronoun co-indexed with PRO is clearly impossible here (*niego$_2$). However, a considerable number of cliticization sites implies that the index co-indexed with the subject of the control predicate has a few options and can be spelled out as either a reflexive (siebie$_1$'self') when it is raised to matrix v or T, or as a pronoun (nią$_1$'her') when it cliticizes to embedded v or embedded T:

(ii) [$_{TP}$ Maria$_1$ index$_1$-T [$_{vP}$ index$_1$-v-told [$_{VP}$ Piotr.DAT$_2$ [$_{V'}$ V [$_{CP}$ [$_{TP}$ PRO$_2$ index$_1$-T [$_{vP}$ PRO index$_{1,2}$-v-look [$_{VP}$ V [$_{PP}$ at index]]]]]]]]

For a more thorough discussion of the issues discussed in this contribution, see Witkoś et al. (Forthcoming).

Acknowledgements

References

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

Avrutin, Sergey. 1994. The structural position of bound variables in Russian. *Linguistic Inquiry* 25(4). 709–727. https://www.jstor.org/stable/4178882.

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

Bailyn, John F. 2010. What's inside VP? New (and old) evidence from Russian. In Wayles Browne, Adam Cooper, Alison Fisher, Esra Kesici, Nikola Predolac & Draga Zec (eds.), *Formal Approaches to Slavic Linguistics 18: The Second Cornell Meeting 2009*, 21–37. Ann Arbor, MI: Michigan Slavic Publications.

Bailyn, John F. 2012. *The syntax of Russian*. Cambridge: Cambridge University Press.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Béjar, Susana & Milan Rezac. 2003. Person licensing and the derivation of PCC effects. In Ana Teresa Perez-Leroux & Yves Roberge (eds.), *Romance Linguistics: Theory and Acquisition*, 49–62. Amsterdam: John Benjamins.

Belletti, Adriana & Luigi Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language & Linguistic Theory* 6(3). 291–352. DOI:10.1007/BF00133902

Boeckx, Cedric, Norbert Hornstein & Jairo Nunes. 2008. Copy-reflexive and copy-control constructions: A movement analysis. *Linguistic Variation Yearbook* 8(1). 61–100. DOI:10.1075/livy.8.03boe

Bondaruk, Anna. 2017. The syntax of accusative and dative experiencer in Polish. Paper presented at the LingBaW Conference at the Catholic Univeristy of Lublin, 18–19.10.2017.

Bondaruk, Anna, Bożena Rozwadowska & Wojciech Witkowski. 2017. Passivization of Polish experiencer verbs vs. the unaccusativity hypothesis (part 1). *Studies in Polish Linguistics* 12(2). 57–73. DOI:10.4467/23005920SPL.17.003.7021

Bondaruk, Anna & Bogdan Szymanek. 2007. Polish nominativeless constructions with dative experiencers: Form, meaning and structure. *Studies in Polish Linguistics* 4. 61–97.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia linguistica* 59(1). 1–45.   DOI:10.1111/j.1467-9582.2005.00118.x

Bošković, Željko. 2012. Phases in NPs and DPs. In Ángel J. Gallego (ed.), *Phases: Developing the framework*, 343–383. Berlin: Walter de Gruyter.

Bošković, Željko. 2013. Phases beyond clauses. In Lilia Schürcks, Anastasia Giannakidou & Urtzi Etxeberria (eds.), *The Nominal Structure in Slavic and Beyond*, 75–128. Berlin: Walter de Gruyter.

Bošković, Željko. 2014. More on the edge of the edge. In Cassandra Chapman, Olena Kit & Ivona Kučerová (eds.), *Formal Approaches to Slavic Linguistics 22: The McMaster Meeting 2013*, 44–66. Ann Arbor, MI: Michigan Slavic Publications.

Burzio, Luigi. 1991. The morphological basis of anaphora. *Journal of Linguistics* 27(1). 81–105.   DOI:10.1017/S0022226700012421

Burzio, Luigi. 1996. The role of the antecedent in anaphoric relations. In Robert Freidin (ed.), *Current issues in comparative grammar*, 1–45. Dordrecht: Kluwer.

Cardinaletti, Anna & Michal Starke. 1994. The typology of structural deficiency: On the three grammatical classes. *University of Venice Working Papers in Linguistics* 4. 41–109.

Cegłowski, Piotr. 2017. *The internal structure of nominal expressions: Reflections on extractability, fronting and phasehood.* Poznań: Wydawnictwo naukowe UAM.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrech: Foris.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

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

Citko, Barbara. 2011. *Symmetry in syntax: Merge, move and labels.* Cambridge: Cambridge University Press.

Citko, Barbara, Allison Germain & Jacek Witkoś. 2018. If you cannot agree, move on! On labels and non-nominative subjects. *Glossa* 3(1). 28. DOI:10.5334/gjgl.399

Cuervo, Maria Cristina. 2003. *Datives at large.* Cambridge, MA: MIT. (Doctoral dissertation).

Despić, Miloje. 2011. *Syntax in the absence of Determiner Phrase.* Storrs, CT: University of Connecticut. (Doctoral dissertation).

Despić, Miloje. 2013. Binding and the structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44(2). 239–270.   DOI:10.1162/LING_a_00126

Despić, Miloje. 2015. Phases, reflexives and definiteness. *Syntax* 18(3). 201–234. DOI:10.1111/synt.12031

Dvořák, Věra. 2010. On the syntax of ditransitive verbs in Czech. In Wayles Browne, Adam Cooper, Alison Fisher, Esra Kesici, Nikola Predolac & Draga Zec (eds.), *Formal Approaches to Slavic Linguistics 18: The Second Cornell Meeting 2009*, 161–177. Ann Arbor, MI: Michigan Slavic Publications.

Dyakonova, Marina. 2007. Russian double object constructions. *ACLC Working Papers* 2. 3–30. http://home.hum.uva.nl/oz/hengeveldp/publications/2007_hengeveld.pdf.

Dyakonova, Marina. 2009. *A phase-based approach to Russian free word order*. Amsterdam: University of Amsterdam. (Doctoral dissertation).

Franks, Steven. 1995. *Parameters of Slavic morphosyntax*. Oxford: Oxford University Press.

Franks, Steven. 2017. *Syntax and spell-out in Slavic*. Bloomington, IN: Slavica Publishers.

Franks, Steven. Forthcoming. PCC violations and their resolutions. In *Formal Approaches to Slavic Languages 26: The Urbana-Champaign Meeting*. Ann Arbor, MI: Michigan Slavic Publications.

Germain, Allison. 2015. Nullifying null expletives: Accounting for EPP in Russian impersonal and nominative in situ constructions. In *Formal Approaches to Slavic Linguistics 23: The Berkeley Meeting 2014*, 418–438. Ann Arbor, MI: Michigan Slavic Publications.

Gogłoza, Aleksandra, Paulina Łęska, Roland Meyer & Jacek Witkoś. Forthcoming. Binding by objects in Polish double object constructions – Experimental analysis of its acceptability and correlation with object order. In Tania Ionin & Jon MacDonald (eds.), *Formal Approaches to Slavic Linguistics 26: The Urbana-Champaign Meeting*. Ann Arbor, MI: Michigan Slavic Publications.

Gračanin-Yuksek, Martina. 2006. Double object construction in Croatian: Arguments against Appl°. In Richard Compton, Magdalena Goledzinowska & Ulyana Savchenko (eds.), *Formal Approaches to Slavic Linguistics 15: The Toronto Meeting 2006*, 94–112. Ann Arbor, MI: Michigan Slavic Publications.

Heinat, Fredrik. 2008. *Probes, pronouns and binding in the minimalist program*. Saarbrucken: VDM Verlag Dr Müller.

Hellan, Lars. 1988. *Anaphora in Norwegian and the theory of grammar*. Dordrecht: Foris.

Hestvik, Arild. 1992. LF movement of pronouns and anti-subject orientation. *Linguistic Inquiry* 23(4). 557–594. http://www.jstor.org/stable/4178790.

Hicks, Glyn. 2009. *The derivation of anaphoric relations*. Amsterdam: John Benjamins.

Hornstein, Norbert. 2001. *Move! A minimalist theory of construal.* Oxford: Blackwell.

Jimenez-Fernandez, Ángel Luis & Bożena Rozwadowska. 2016. The information structure of Dative Experiencer psych-verbs. In Bożena Cetnarowska, Marcin Kuczok & Marcin Zabawa (eds.), *Various dimensions of contrastive studies*, 100–121. Katowice: Wydawnictwo UŚ.

Kayne, Richard S. 1985. Principles of particle constructions. In Jean-Yves Pollock, Hans-Georg Obenauer & Jacqueline Guéron (eds.), *Grammatical Representation*, 101–140. Dordrecht: Foris.

Kayne, Richard S. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22(4). 647–686. http://www.jstor.org/stable/4178745.

Kayne, Richard S. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Klimek, Dorota & Bożena Rozwadowska. 2004. From psych adjectives to psych verbs. *Poznań Studies in Contemporary Linguistics* 39. 59–72. http://wa.amu.edu.pl/psicl/Volume_39.

Landau, Idan. 2000. *Elements of control: Structure and meaning in infinitival constructions.* Dordrecht: Kluwer Academic Publishers.

Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19(3). 335–391. http://www.jstor.org/stable/25164901.

Larson, Richard K. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21(4). 589–632. http://www.jstor.org/stable/4178697.

Larson, Richard K. 2014. *On shell structure.* London: Routledge.

Lee-Schoenfeld, Vera. 2008. Binding, phases and locality. *Syntax* 11(3). 281–298. DOI:10.1111/j.1467-9612.2008.00118.x

Manzini, Rita & Kenneth Wexler. 1987. Parameters, binding theory and learnability. *Linguistic Inquiry* 18(3). 413–444. http://www.jstor.org/stable/4178549.

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

May, Robert. 1985. *Logical form: Its structure and derivation.* Cambridge, MA: MIT Press.

Miechowicz-Mathiasen, Katarzyna & Paweł Scheffler. 2008. A corpus-based analysis of the peculiar behaviour of the Polish verb *podobać się.* In Jacek Witkoś & Gisbert Fanselow (eds.), *Elements of Slavic and Germanic Grammars: A comparative view*, 89–112. Frankfurt am Main: Peter Lang.

Moore, John & David M. Perlmutter. 2000. What does it take to be a dative subject? *Natural Language & Linguistic Theory* 18(2). 373–416. DOI:10.1023/A:1006451714195

Nikolaeva, Liudmila. 2014. *The secret life of pronouns*. Cambridge, MA: MIT. (Doctoral dissertation).

Pica, Pierre. 1987. On the nature of the reflexivization cycle. In Joyce McDonough & Bernadette Plunkett (eds.), *NELS 17: Proceedings of the 17th Annual Meeting of the North East Linguistic Society*, 483–500. Amherst, MA: GLSA Publications.

Pica, Pierre. 1991. On the interaction between Antecedent-Government and Binding: The Case of Long-Distance Reflexivization. In Jan Koster & Eric Reuland (eds.), *Long distance anaphora*, 119–135. Cambridge: Cambridge University Press.

Progovac, Ljiljana. 1992. Long-distance reflexives without movement. *Linguistic Inquiry* 23(4). 671–680. https://www.jstor.org/stable/4178796.

Progovac, Ljiljana. 1993. Long-distance reflexives: Movement to infl versus relativized subject. *Linguistic Inquiry* 24(4). 755–772. http://www.jstor.org/stable/4178839.

Pylkkänen, Liina. 2002. *Introducing arguments*. Cambridge, MA: MIT. (Doctoral dissertation).

Rappaport, Gilbert. 1986. On anaphor binding in Russian. *Natural Language & Linguistic Theory* 4(1). 97–120. DOI:10.1007/BF00136266

Reinders-Machowska, Ewa. 1991. Binding in Polish. In Jan Koster & Eric Reuland (eds.), *Long distance anaphora*, 137–150. Cambridge: Cambridge University Press.

Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1990. On the anaphor-agreement effect. *Rivista di Linguistica* 2(1). 27–42.

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

Roberts, Ian. 1992. A formal account of grammaticalisation in the history of Romance futures. *Folia Linguistica Historica* 13(1/2). 219–258. DOI:10.1515/flih.1992.13.1-2.219

Roberts, Ian. 1993. *Verbs and diachronic syntax*. Dordrecht: Kluwer Academic Publishers.

Roberts, Ian. 2009. *Agreement and head movement*. Cambridge, MA: MIT Press.

Safir, Ken. 2004. *The syntax of anaphora*. Oxford: Oxford University Press.

Safir, Ken. 2014. One true anaphor. *Linguistic Inquiry* 45(1). 91–124. DOI:10.1162/LING_a_00149

Sportiche, Dominique. 1996. Clitic constructions. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 213–276. Dordrecht: Springer.

Tajsner, Przemysław. 2008. *Aspects of the grammar of focus: A minimalist view*. Frankfurt am Main: Peter Lang.

Vikner, Sten. 1985. Parameters of binder and of binding category in Danish. *Working Papers in Scandinavian Syntax* 23. 1–61.

Willim, Ewa. 1989. *On word order: A government-binding study of English and Polish*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Witkoś, Jacek. 1998. *The syntax of clitics: Steps towards a minimalist account*. Poznań: Motivex.

Witkoś, Jacek. 2007. Polish and A-type scrambling. In Peter Kosta & Lilia. Schürcks (eds.), *Linguistic investigations into Formal Description of Slavic Languages*, 165–180. Frankfurt am Main: Peter Lang.

Witkoś, Jacek. 2008. On the correlation between A-type scrambling and lack of weak crossover effects. *Studia Anglica Posnaniensia* 44. 297–328.

Witkoś, Jacek & Dominika Dziubała-Szrejbrowska. 2015. A note on the genitive of quantification in Polish and derivational phases. *Poznań Studies in Contemporary Linguistics* 51(3). 433–462.

Witkoś, Jacek, Paulina Łęska, Aleksandra Gogłoza & Dominika Dziubała-Szrejbrowska. Forthcoming. *Bind me tender, bind me do! Dative and accusative arguments as antecedents for reflexives in Polish*. Frankfurt am Main: Peter Lang.

Woolford, Ellen. 1999. More on the anaphor agreement effect. *Linguistic Inquiry* 30(2). 257–287. DOI:10.1162/002438999554057

Woolford, Ellen. 2006. Lexical case, inherent case, and argument structure. *Linguistic Inquiry* 37(1). 111–130. DOI:10.1162/002438906775321175

Zubkov, Petr. 2018. *The grammar of binding: A study with reference to Russian*. Utrecht: Utrecht University. (Doctoral dissertation).

Żychliński, Sylwiusz. 2013. *On some aspects of the syntax of object experiencers in Polish and English*. Poznań: Adam Mickiewicz University. (Doctoral dissertation).

Name index

Language index

Language index

Subject index

Did you like this book?

This book was brought to you for free

Advances in formal Slavic linguistics 2017

Advances in Formal Slavic Linguistics 2017 is a collection of fifteen articles that were prepared on the basis of talks given at the conference Formal Description of Slavic Languages 12.5, which was held on December 7–9, 2017, at the University of Nova Gorica. The volume covers a wide array of topics, such as control verbs, instrumental arguments, and perduratives in Russian, comparatives, negation, n-words, negative polarity items, and complementizer ellipsis in Czech, impersonal se-constructions and complementizer doubling in Slovenian, prosody and the morphology of multi-purpose suffixes in Serbo-Croatian, and indefinite numerals and the binding properties of dative arguments in Polish. Importantly, by exploring these phenomena in individual Slavic languages, the collection of articles in this volume makes a significant contribution to both Slavic linguistics and to linguistics in general.