

---

*FASL* 23, 2014

---

Michigan Slavic Publications is a non-profit organization associated with the Department of Slavic Languages and Literatures of the University of Michigan. Its goal is to publish titles which substantially aid the study of and teaching of Slavic and East European languages and cultures. The present volume, based on a conference held at the University of California at Berkeley in May 2014, continues a series of conference proceedings devoted to formal approaches to Slavic languages.

Michigan Slavic Materials, 61

Series Editor  
Jindřich Toman  
ptydepe@umich.edu

**Annual Workshop on  
Formal Approaches  
to Slavic Linguistics**

---

**The  
First  
Berkeley  
Meeting**

Edited by

Małgorzata Szajbel-Keck  
Roslyn Burns  
Darya Kavitskaya

Michigan Slavic Publications  
Ann Arbor 2015

collection © Michigan Slavic Publications 2015  
individual contributions © authors

### Library of Congress Cataloging-in-Publication Data

Annual Workshop on Formal Approaches to Slavic Linguistics (23rd : 2014 : Berkeley, California)

Annual Workshop On Formal Approaches To Slavic Linguistics : the first Berkeley meeting / edited by Malgorzata Szajbel-Keck, Roslyn Burns, Darya Kavitskaya.

pages cm. -- (Michigan Slavic materials, 61)

"The present volume, based on a conference held at the University of California at Berkeley, in May 2014 continues a series of conference proceedings devoted to formal approaches to Slavic languages, including Bulgarian, Croatian, Macedonian, Polish, Serbian, and Russian. We are proud to call this volume The First Berkeley Meeting."

"Michigan Slavic Publications 2015."

"This volume contains selected papers presented at the 23rd Annual Meeting of the Formal Approaches to Slavic Linguistics held May 2/4, 2014 at the University of California, Berkeley."

ISBN 978-0-936534-16-9 (alk. paper)

1. Slavic languages--Grammar--Congresses. I. Szajbel-Keck, Malgorzata, editor. II. Burns, Roslyn, editor. III. Kavitskaya, Darya, editor. IV. Title. V. Title: Formal approaches to Slavic linguistics. VI. Title: First Berkeley Meeting, 2014. VII. Title: FASL 23, 2014.

PG13.M46 2014

491.8'045--dc23

2015018183

Michigan Slavic Publications  
Department of Slavic Languages and Literatures  
University of Michigan  
Ann Arbor, MI 48109-1275  
michsp@umich.edu

## PREFACE

This volume contains selected papers presented at the 23<sup>rd</sup> Annual Meeting of the Formal Approaches to Slavic Linguistics held May 2–4, 2014, at the University of California, Berkeley.

Thanks to the generous sponsorship of the Institute of Slavic, East European, and Eurasian Studies, the Department of Slavic Languages and Literatures, and the Department of Linguistics, this conference was able to bring together a wide range of scholarship whose expertise covered a variety of linguistic sub-fields. The conference program consisted of 32 talks, on aspect and tense, verb and predication, operators, clitics, word order, morphophonological issues and more. The discussed languages included Bulgarian, Croatian, Macedonian, Polish, Serbian, and Russian. We would like to thank Barbara Citko (University of Washington), Johanna Nichols (University of California, Berkeley), and Jerzy Rubach (University of Iowa / Uniwersytet Warszawski) for their readiness to participate as invited speakers.

Production of this volume would not have been possible without the dedicated work of our team of reviewers. We would like to thank Wayles Browne, Jakub Dotlacil, Hana Filip, Pavel Grashchenkov, Atle Grønn, Stephanie Harves, Tania Ionin, Katja Jasinskaja, Anna Kibort, Franc Marušič, Troy Messik, Ljiljana Progovac, Eugenia Romanova, Catherine Rudin, Radek Šimík, Luka Szucsich, Ludmila Veselovská, Jacek Witkoś Martina, and Gračanin-Yukse. Without their willingness to commit their time and energy to prepare thoughtful comments and critical review, this volume would not be what it is.

All papers were edited for style by the editors and authorized by the

authors before final submission. Finally, the review team would like to thank Małgorzata Szajbel-Keck and Jindřich Toman for the time and effort they spent in preparing the publication of this volume.

*The Editors*

Małgorzata Szajbel-Keck

Roslyn Burns

Darya Kavitskaya

University of California, Berkeley

## CONTENTS

<i>Polina Berezovskaya and Vera Hohaus</i> The Crosslinguistic Inventory of Phrasal Comparative Operators: Evidence from Russian	1
<i>Željko Bošković</i> On Multiple Left-Branch Dislocation: Multiple Extraction and/or Scattered Deletion?	20
<i>Barbara Citko</i> The Gapping that Could $\Delta$	36
<i>Steven Franks</i> Speculations on DP-structure: Macedonian vs. Bulgarian	56
<i>Allison Germain</i> Nullifying Null Expletives: Accounting for EPP in Russian Impersonal and Nominative in situ Constructions	77
<i>Julie Goncharov</i> About <i>many</i>	99
<i>Łukasz Jędrzejowski</i> On Tensed Modals in Polish	122
<i>Hakyung Jung and Krzysztof Migdalski</i> On the Degrammaticalization of Pronominal Clitics in Slavic	143

<i>Tatiana Luchkina and Tania Ionin</i> The Effect of Prosody on Availability of Inverse Scope in Russian	163
<i>Ora Matushansky and E. G. Ruys</i> 4000 Measure NPs: Another Pass Through the <i>шлюз</i>	184
<i>Denis Paperno</i> An Alternative Semantics for Negative Conjunction in Russian	206
<i>Agnieszka Patejuk</i> Coordinated Wh-words in Polish: Monoclausal or Multiclausal?	222
<i>Ljiljana Progovac</i> The Absolutive Basis of Middles and the Status of vP and UTAH	242
<i>Adam Przepiórkowski</i> A Weakly Compositional Analysis of Distance Distributivity in Polish	262
<i>María Luisa Rivero and Vesela Simeonova</i> The Inferential Future in Bulgarian: An Evidential Modal Proposal	282
<i>Katie Sardinha</i> Locus of Causation and <i>by itself</i> Phrases: A Case Study of Russian <i>sam po sebe</i>	302
<i>Aida Talić</i> On Clitics, Their Place in the Prosodic Structure, and Accent	322

<i>Egor Tsedryk</i> Deriving Null Pronouns: A Unified Analysis of Subject Drop in Russian	342
<i>Ksenia Zanon</i> Russian Anaphoric Possessive in Context	362
<i>Yulia Zinova and Hana Filip</i> Scalar Implicatures of Russian Verbs	382



## **The Crosslinguistic Inventory of Phrasal Comparative Operators: Evidence from Russian\***

*Polina Berezovskaya*  
Eberhard Karls Universität Tübingen

*Vera Hohaus*  
Eberhard Karls Universität Tübingen

In and across languages, there are often multiple compositional roads to the same meaning. We can thus arrive at identical truth conditions in very different fashions. This paper discusses just such a case, namely variation in the lexical inventory of comparative operators: In Russian as well as crosslinguistically, one and the same comparison is arrived at by very different lexical and structural means. More specifically, we argue that genitive-marked synthetic comparatives in Russian provide evidence for the phrasal comparative operator proposed in Kennedy (1997). We also show that this operator does not always have to be interpreted *in situ*, contrary to the claims in Beck, Hohaus & Tiemann (2012).

We go about this as follows: In the first section of the paper, we provide some necessary background. We briefly introduce some key features of the semantic analysis of the comparative and point out in how

---

\* We thank Nadine Bade, Sigrid Beck, Verena Hehl, Anna Howell, Tania Ionin, Ora Matushansky, Konstantin Sachs, and Sonja Tiemann for feedback and discussion. We are very grateful to Natalia Berezovskaya, Natalia Chmelnsky, Olga Daragan, Larissa Kaminskaya, Katja Leimann, Tatiana Lyubimkova, Sergei Primenko, Vlada Riftina, Zinaida Touraeva, Maria Yelenevskaya and Natalia Zubko for native speaker judgments. Thank you also to the anonymous reviewer whose comments helped improve this paper. Research for this paper was conducted within Project C1 of the Collaborative Research Center 833, for which funding is provided by the German Research Foundation DFG.

far this analysis is subject to crosslinguistic variation. For the purposes of this paper, we are most interested in the variation regarding the choice of comparative operator and its empirical consequences. This section also introduces the reader to Russian comparatives. In the second section, we discuss the genitive-marked comparative in Russian in greater detail and show that the empirical evidence is only compatible with a certain type of phrasal analysis. We explore the consequences of this analysis in the third section. Conclusions are offered in the fourth section.

### 1 The Crosslinguistic Inventory of Comparative Operators

Say my friend Mary is taller than her wife, Sue. Comparing the two individuals involves the following two ingredients apart from Mary and Sue themselves: The comparison is, first, along some dimension. In our example, this dimension is height. Mary and Sue need both be mapped to their degree of height. Second, these two degrees are being related, here by an exceed-relation. In the semantic analysis of the comparative, this first ingredient of the comparison, the dimension, is contributed by a gradable predicate such as English *tall*, with the lexical entry in (1). At the core of this lexical entry is a measure function of type  $\langle e, d \rangle$ , which maps an individual to its height degree.<sup>1</sup>

$$(1) \quad \llbracket tall \rrbracket = \lambda d_{\langle d \rangle}. \lambda x_{\langle e \rangle}. HEIGHT(x) \geq d$$

Our second ingredient, the greater-than relation between two degrees, is contributed by a degree operator, which in English, for instance, is morphologically realized as *-er*. Besides, degree operators are in charge of putting all the ingredients of a comparison together at Logical Form. The way they do so differs, however. This fact is reflected in the different entries for the comparative operator in (2) to (4).

$$(2) \quad \llbracket -er_{\text{clausal}} \rrbracket = \lambda D'_{\langle d, t \rangle}. \lambda D_{\langle d, t \rangle}. MAX(D) > MAX(D')$$

$$(3) \quad \llbracket -er_{\text{Heim(1985)}} \rrbracket = \lambda y_{\langle e \rangle}. \lambda R_{\langle d, \langle e, t \rangle \rangle}. \lambda x_{\langle e \rangle}. \\ MAX(\lambda d. R(d)(x)) > MAX(\lambda d'. R(d')(y))$$

---

<sup>1</sup> We refer the reader to von Stechow (1984) and Beck (2011) for a more comprehensive introduction to the syntax and semantics of comparison constructions. The paper is couched in an extended Heim & Kratzer's (1998) framework.

- (4)  $\llbracket -er_{\text{Kennedy}(1997)} \rrbracket = \lambda R_{\langle d, \langle e, t \rangle \rangle}. \lambda y_{\langle e \rangle}. \lambda x_{\langle e \rangle}. \text{MAX}(\lambda d. R(d)(x)) > \text{MAX}(\lambda d'. R(d')(y))$
- (5)  $\llbracket \text{MAX} \rrbracket = \lambda D_{\langle d, t \rangle}. \text{id} [\forall d' [D(d') \rightarrow d \geq d']]$

The operator in (2) differs from the operators in (3) and (4) in the structural environments in which it is employed. This operator is used in comparatives like the English examples in (6) and (7), in which the *than*-constituent is clausal, either overtly, as in (6), or underlyingly, as in (7).

- (6) Mary has more cats [than John has children].
- (7) The air was even smokier today [than ~~it was smoky~~ yesterday].

Logical Form:  $[[\text{DegP } -er_{\text{clausal}} [\langle d, t \rangle 1, \langle d \rangle [[\text{the air}] t_1, \langle d \rangle \text{ smoky yesterday}]]]$   
 $[\langle d, t \rangle 2, \langle d \rangle [[\text{the air}] \text{ was } t_2, \langle d \rangle \text{ smoky yesterday}]]]$

Interpretation:  $\text{MAX}(\lambda d. \text{SMOKE}_{\text{today}}(\text{the.air}) \geq d) > \text{MAX}(\lambda d'. \text{SMOKE}_{\text{yesterday}}(\text{the.air}) \geq d')$

Not all comparatives are amendable to such a clausal analysis, however. (See *e.g.* Hofstetter 2009, Bhatt & Takahashi 2011 as well as Beck, Hohaus & Tiemann 2012.) The standard of the comparison might as well just be a phrase, which is where the operators in (3) and (4) come in. What is now the difference between these two operators? The only but crucial difference between (3) and (4) lies in the order in which they combine with their arguments, that is they are schönfinkled differently. Beck, Hohaus & Tiemann (2012) show that this difference in Schönfinkelization matters though: As illustrated in Figure 1, the phrasal operator attributed to Heim (1985) can be used to analyze a wider range of constructions compared to the operator from Kennedy (1997).

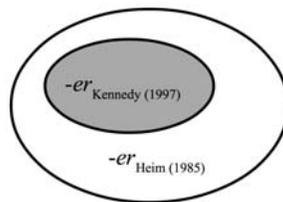


Fig. 1: Applicability of the Two Phrasal Comparative Operators

More specifically, while both operators do fine in the case of a simple predicative (-PRED-) comparative like (8),  $-er_{\text{Kennedy}(1997)}$  only derives one

of the readings available for an attributive phrasal comparative like the English example in (9).

(8) Mary is taller than Sue.

Logical Form #1: [Mary<sub><e></sub> [[-er<sub>Heim(1985)</sub> [<sub><e></sub> than Sue]] tall<sub><d,<e,t>></sub>]]

Logical Form #2: [Mary<sub><e></sub> [[-er<sub>Kennedy(1997)</sub> tall<sub><d,<e,t>></sub>] [<sub><e></sub> than Sue]]]

Interpretation:  $\text{MAX}(\lambda d. \text{HEIGHT}(\text{Mary}) \geq d) > \text{MAX}(\lambda d'. \text{HEIGHT}(\text{Sue}) \geq d')$

(9) Mary bought a faster computer than John.

As is illustrated in Figure 2, attributive comparatives such as (9) are ambiguous between an external reading (-ATTR(EXT)-), in which comparison is between Mary's and John's computer, and an internal reading (-ATTR(INT)-), in which Mary's computer is being compared with John (Lerner & Pinkal 1995). Albeit implausible for (9), the internal reading is the preferred reading in the case of the example in (9') below.

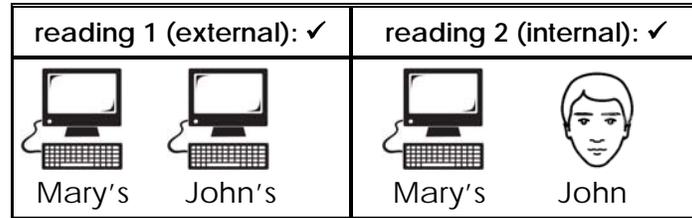


Fig. 2: The Internal and the External Attributive Readings

(9') Mary bought a faster computer than her old one.

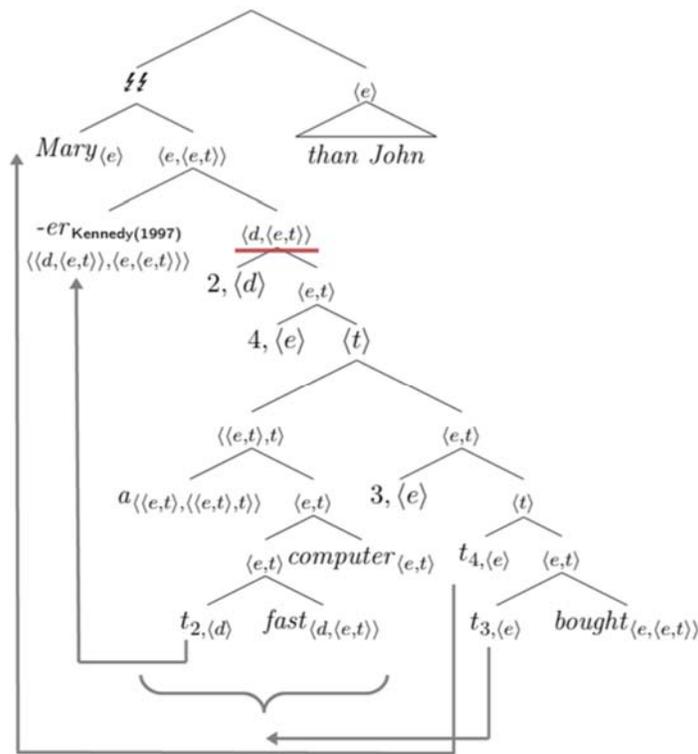
Of the two phrasal operators introduced above, only  $-er_{\text{Heim}(1985)}$  is suited to generate the external readings of attributive comparatives, as Beck, Hohaus & Tiemann (2012) point out. Deriving -ATTR(EXT)- requires a degree relation of type  $\langle d, \langle e, t \rangle \rangle$  other than the one lexically provided in the example, namely the one in (10). This relation between speed degrees and the owner of a recently bought computer has to be syntactically derived at Logical Form. If we try to do so in an attempt to generate the external reading with  $-er_{\text{Kennedy}(1997)}$ , we fail, as illustrated in (11).

(10)  $\lambda d_{\langle d \rangle}. \lambda x_{\langle e \rangle}. \exists y[\text{computer}(y) \ \& \ \text{bought}(y)(x) \ \& \ \text{SPEED}(y) \geq d]$

While we are able to derive the required degree relation in (11) by

movement of the subject and parasitic movement of the degree operator (cf. e.g. Nissenbaum 2000, Beck & Sauerland 2000 for discussion), the operator then cannot combine with its two  $\langle e \rangle$ -type arguments in the intended order and the derivation fails. The degree operator from Kennedy (1997) is unable to undergo parasitic movement and must be interpreted *in situ*. When it is, however, we only derive -ATTR(INT)-. Beck, Hohaus & Tiemann (2012) conclude from this that this particular phrasal operator is scopally not mobile. We will argue in section 3.1 that this generalization is not quite right and that the operator has a limited scopal mobility (-SCOPMOB-) after all.

(11)



Where are we? We have seen, albeit only briefly, that the order in which the two phrasal operators from (3) and (4) combine with their arguments

has repercussions for their applicability. The results are summarized in Table 1.

	-PRED-	-ATTR(INT)-	-ATTR(EXT)-	-SCOPMOB-
<i>-er</i> <sub>Heim(1985)</sub>	Yes.	Yes.	Yes.	Yes.
<i>-er</i> <sub>Kennedy(1997)</sub>	Yes.	Yes.	No.	No.

Table 1: Applicability of the Two Phrasal Comparative Operators

Given the multitude of degree operators introduced for the comparative in (2) to (4), the following questions arise: Is there indeed empirical evidence that we need that many operators? Or is this just a finger exercise for the enthusiastic semanticist? These questions are particularly pressing for *-er*<sub>Kennedy(1997)</sub>, an operator which lacks the empirical coverage of *-er*<sub>Heim(1985)</sub>, and which we might therefore be tempted to remove from our inventory of degree operators. Data from first language acquisition (Hohaus & Tiemann 2009, Tiemann, Hohaus & Beck 2012, Hohaus, Tiemann & Beck 2014) as well as crosslinguistic research (Merchant 2009, 2011, 2012; Beck, Hohaus & Tiemann 2012) suggests that such a move would be hasty: First, English *than*-phrases are acquired considerably earlier than their German equivalents, which suggests that they receive a simpler analysis in English than they do in German, namely an analysis with *-er*<sub>Kennedy(1997)</sub>. Second, Greek has two different phrasal comparative constructions which differ in the way the standard is realized, examples of which are provided in (12) and (13).<sup>2</sup> The unacceptability of (13) derives from the fact that the genitive-marked comparative in Greek only allows for -ATTR(INT)- but lacks -ATTR(EXT)-. An analysis with *-er*<sub>Kennedy(1997)</sub> thus suggests itself.

- (12) O Giannis exi perisotera periodika [apo mena].  
 the Giannis has more magazines from me  
 ‘Giannis has more magazines than I have.’  
 (Merchant 2012: 6)

<sup>2</sup> Abbreviations used in glosses are: ACC = accusative, COMP = comparative morphology, GEN = genitive, INSTR = instrumental, PERF = perfective, and PREP = prepositional.

- (13) #O Giannis exi perisotera periodika [mu].  
 the Giannis has more magazines me<sub>GEN</sub>  
 ‘Giannis has more magazines than I am.’  
 (Merchant 2012: 6)

Then again, if we take a closer look at the crosslinguistic picture, in Table 2, Greek and English are the only languages for which such an analysis is plausible.

<b>English</b>	- <i>er</i> <sub>clausal</sub> - <i>er</i> <sub>Kennedy(1997)</sub> - <i>er</i> <sub>Heim(1985)</sub>
<b>German</b>	- <i>er</i> <sub>clausal</sub>
<b>Hindi</b> (Bhatt & Takahashi 2011)	- <i>er</i> <sub>Heim(1985)</sub>
<b>Turkish, Thai</b> (Hofstetter 2009, 2010)	- <i>er</i> <sub>Heim(1985)</sub>
<b>Persian, Tajiki, Ishkashimi</b> (Karvovskaya 2013)	- <i>er</i> <sub>Heim(1985)</sub>
<b>Greek</b> (Merchant 2009, 2011, 2012)	- <i>er</i> <sub>clausal</sub> ( <i>ap’oti</i> -clause) - <i>er</i> <sub>Heim(1985)</sub> ( <i>apo</i> -phrase) - <i>er</i> <sub>Kennedy(1997)</sub> (genitive-marked)

Table 2: The Crosslinguistic Inventory of Comparative Operators

If we want to hold on to *-er*<sub>Kennedy(1997)</sub>, it would therefore be nice to have evidence from another language for this operator. We argue that Russian is just such a language: The Russian genitive-marked comparative is best analyzed as employing *-er*<sub>Kennedy(1997)</sub><sup>3</sup>. Before we look at the data in favor of such an analysis, let us briefly take stock of the inventory of Russian comparatives.

<sup>3</sup> For an exploration of the acquisitional predictions of this analysis for Russian, we refer the reader to Berezovskaya (2014).

### 1.1 *The Inventory of Russian Comparatives*

Russian has both, a clausal and a phrasal comparative construction at its disposal. Comparatives as in (14a) and (14b), in which the standard of comparison is introduced by *čem*, are analyzed as being derived from a clausal source via an obligatory reduction operation (Pancheva 2006, Beck et al. 2009). The word *čem* is a *wh*-phrase in the instrumental case. Evidence for a clausal analysis comes from examples such as (14b), which contain a tensed auxiliary. The Russian *čem*-comparative is thus analyzed with the clausal operator from (2), repeated below:

- (14) a. Oleg umnee [čem Tolja].  
           Oleg clever<sub>COMP</sub> what<sub>INSTR</sub> Tolja  
           ‘Oleg is cleverer than Tolja is clever.’  
       b. Oleg umnee [čem byl Tolja v ego vozraste].  
           Oleg clever+COMP what<sub>INSTR</sub> was Tolja in his age  
           ‘Oleg is cleverer than Tolja was when he was his age.’  
 (2)  $\llbracket -er_{\text{clausal}} \rrbracket = \lambda D'_{\langle d,t \rangle}. \lambda D_{\langle d,t \rangle}. \text{MAX}(D) > \text{MAX}(D')$

Another possibility of expressing a comparison in Russian is in (15). Here, the standard of comparison is marked by the genitive case.

- (15) Oleg vyše Toli.  
       Oleg tall<sub>COMP</sub> Tolja<sub>GEN</sub>  
       ‘Oleg is taller than Tolja.’

Russian genitive-marked comparatives are best analyzed as employing the somewhat less powerful *-er*<sub>Kennedy(1997)</sub> because (i) they do not allow for clausal standards, (ii) they do only allow for *in situ* readings when used attributively, and (iii) they do not exhibit scope ambiguities unlike their clausal siblings.

## 2 Evidence for Kennedy’s Operator from Russian

Let us look at the relevant data in turn.

2.1 *Unavailability of Clausal Standards.* The first piece of data that suggests that genitive-marked comparatives are best analyzed with

*-er*<sub>Kennedy(1997)</sub> is that they, quite expectedly, do not allow for clausal standards. Consider the minimal pair in (16) and (17).

- (16) Maša pela gromče [čem Katja svistela].  
 Masha sang loud<sub>COMP</sub> what<sub>INSTR</sub> Katja whistled  
 ‘Masha sang louder than Katja whistled.’
- (17) \*Maša pela gromče [Kati svistela].  
 Masha sang loud<sub>COMP</sub> Katja<sub>GEN</sub> whistled  
 ‘Masha sang louder than Katja whistled.’

The example in (17) is only compatible with a phrasal analysis.

2.2 *Only in situ Readings for Attributive Uses.* When used attributively, Russian genitive-marked comparatives lack -ATTR(EXT)- and only allow for the *in situ* interpretation, -ATTR(INT)-. Consider the examples in (18) and (19). In (18), the internal reading is the preferred reading as computers do not own computers in our world. In (19), comparison is thus most likely between Mary’s and Vanja’s computer. However, this interpretation is unavailable and only the implausible, internal reading is available (as indicated by the hash). Figure 3 summarizes the observed pattern of available readings.

- (18) Maša kupila [kompjuter [AP moščnee  
 Masha bought computer<sub>ACC</sub> powerful<sub>COMP</sub>  
 ètogo kompjetera]].  
 this<sub>GEN</sub> computer<sub>GEN</sub>  
 ‘Masha bought a more powerful computer than this computer.’
- (19) #Maša kupila [kompjuter [AP moščnee Vani]].  
 Masha bought computer<sub>ACC</sub> powerful<sub>COMP</sub> Vanja<sub>GEN</sub>  
 ‘Masha bought a computer more powerful than Vanja.’

reading 1 (external): ✖	reading 2 (internal): ✔
  Maša's      Vanja's	  Maša's      Vanja

Fig. 3: The Internal and the External Attributive Readings in Russian

Under an analysis of Russian genitive-marked comparatives as employing *-er*<sup>Kennedy(1997)</sup>, this pattern is expected as the operator cannot undergo the parasitic movement needed to derive the external reading.

Before we move on, let us briefly consider the syntactic status of the phrases which we label as Adjective Phrases (APs) in (18) and (19). Considering that attributive APs in Russian occur both, post- as well as pre-nominally, as in the contextual comparative in (20), it could be objected that, in both, (18) and (19), these APs are contained within a reduced relative clauses with the structure in (21). This is also the syntactic analysis which Matushansky (2002) assumes.

- (20) a. Maša kupila [[<sub>AP</sub> bolee moščnyj]<sup>4</sup> kompju<sub>ter</sub>].  
 Masha bought more powerful computer<sub>ACC</sub>  
 ‘Masha bought a more powerful computer  
 (compared to a contextually salient other computer).’
- b. Maša kupila [[kompju<sub>ter</sub>] [<sub>AP</sub> bolee moščnyj]].  
 Masha bought computer<sub>ACC</sub> more powerful  
 ‘Masha bought a more powerful computer  
 (compared to a contextually salient other computer).’
- (21) [<sub>NP<e,t></sub> [<sub>N'</sub> computer<sub><e,t></sub> [<sub>RelCl<e,t></sub>  $\emptyset$  <sub>1,<e></sub> [<sub>t<sub>1,<e></sub></sub> more.powerful [...]]]]]

If (21) is indeed the underlying structure for these examples, the unavailability of -ATTR(EXT)- might be simply an island effect: The

<sup>4</sup> The attentive reader might have noticed that we employ the analytic form of the comparative, *bolee moščnyj* (‘more powerful’) in this example. The synthetic form sounds off here. We are aware of the fact that there are certain restrictions on the distribution of the synthetic vs. analytic comparative forms (cf. e.g. Matushansky 2002). However, we think that the synthetic/analytic-distinction is not relevant to our question. What is important here is that Adjectival Phrases can occupy both, the pre-nominal and the post-nominal position in Russian.

derivation of the relevant reading requires movement of the degree operator out of the relative clause. Such movement might be blocked if relative clauses, even in their reduced form, constitute syntactic islands in Russian (and are thus not a reflex of the choice of the degree operator). Until this syntactic question has received the closer attention it deserves, we are unable to decide whether a reduced-relative-clause analysis is any more plausible (or any less stipulative) than the AP-analysis we assume above. Two pieces of data might however point in favor of our AP-analysis. We discuss them in turn.

First, the genitive-marked comparative is to a certain degree acceptable even in the pre-nominal position, as is illustrated in the example in (22), both of which are certainly not entirely ungrammatical.

- (22) a. <sup>??</sup>Maša kupila [NP [AP moščnee] kompjuter] Vani.  
 Masha bought powerful<sub>COMP</sub> computer<sub>ACC</sub> Vanja<sub>GEN</sub>  
 ‘Masha bought a more powerful computer than Vanja.’  
 b. <sup>??</sup>Maša kupila [NP [AP moščnee] Vani kompjuter].  
 Masha bought powerful<sub>COMP</sub> Vanja<sub>GEN</sub> computer<sub>ACC</sub>  
 ‘Masha bought a more powerful computer than Vanja.’

For both (22a) and (22b), the plausible reading -ATTR(EXT)- is however also unavailable. Second, temporal adverbial phrases, which would constitute evidence for more structure beyond AP, are ungrammatical in the post-nominal genitive-marked comparative in Russian. They are however grammatical in the corresponding relative-clause construction, as is illustrated in (23).

- (23) a. \*Maša včera kupila cvetok [segodnja eščjo krasivee].  
 Masha yesterday bought flower today even pretty<sub>COMP</sub>  
 ‘Yesterday, Masha bought a flower even more beautiful today.’  
 b. Maša včera kupila cvetok [kotoryj segodnja eščjo  
 Masha yesterday bought flower which today even  
 krasivee].  
 pretty<sub>COMP</sub>  
 ‘Yesterday, Masha a bought a flower which today is even more  
 beautiful.’

We leave further exploration of this important question to future research, and move on to our last piece of evidence in favor of an analysis of genitive-marked comparatives as employing *-er*<sup>Kennedy(1997)</sup>.

2.3 *Absence of Scope Ambiguities.* As expected under such an analysis, genitive-marked comparatives in Russian do not exhibit scope ambiguities. In this respect they are unlike their clausal siblings with the *čem*-marked standard, for which Krasikova (2007) and Beck et al. (2009) observe ambiguities between the Degree Phrase, which hosts the comparative operator, and other quantificational elements such as modals. Consider the phrasal comparative in (24), for example, which only has a surface scope reading. The inverse scope reading, under which comparison is between what both girls desire (and which requires movement of the comparative operator above the propositional attitude verb), is absent. Both readings are, however, available for the clausal comparative in (25).

- (24) Katja xočet byt' vyše Maši.  
 Katja wants be tall<sub>COMP</sub> Masha<sub>GEN</sub>  
 =‘Katja wants to be taller than Masha is tall.’ [want>-er]  
 ≠‘Katja wants to be taller than Masha wants to be tall.’ [-er >want]
- (25) Katja xočet byt' vyše čem Maša.  
 Katja wants be tall<sub>COMP</sub> what<sub>INST</sub> Masha  
 =‘Katja wants to be taller than Masha is tall.’ [want>-er]  
 =‘Katja wants to be taller than Masha wants to be tall.’ [-er >want]

Considering everything we have discussed above, genitive-marked Russian comparatives are best analyzed with a phrasal operator with limited scopal mobility.

	-CLAUSAL STANDARDS-	-ATTR(INT)-	-ATTR(EXT)-	-SCOPMOB-
Standard(GEN.)	No.	Yes.	No.	No.

Table 3: Genitive-Marked Standards in Russian

The pattern summarized in Table 3 is expected under such an analysis. In addition to Greek and English, Russian also provides evidence for

assuming  $-er_{\text{Kennedy}(1997)}$  in addition to  $-er_{\text{Heim}(1997)}$  and  $-er_{\text{clausal}}$ . We move on to exploring some further repercussions of the analysis.

### 3 Consequences of the Analysis

A comparison between genitive-marked standards of comparatives in Russian and Greek turns out to be worthwhile as it brings to light an interesting distinction between the two languages: Only in Russian are adverbial genitive-marked comparatives grammatical.

#### 3.1 Reinvestigating the Crosslinguistic Picture

When we turn back to the crosslinguistic picture we drew in section 1, Russian is like English and Greek in that it has  $-er_{\text{Kennedy}(1997)}$  at its disposal. Russian and Greek are however morphologically more transparent than English in that they indicate which operator a comparative employs by introducing the standard of comparison differently depending on the operator. The whole pattern is summarized in Table 4 below.

<b>Russian</b>	$-er_{\text{clausal}}$	<i>čem</i> -clause
	$-er_{\text{Kennedy}(1997)}$	genitive-marked phrase
<b>Greek</b> (Merchant 2009, 2011, 2012)	$-er_{\text{clausal}}$	<i>ap'oti</i> -clause
	$-er_{\text{Heim}(1985)}$	<i>apo</i> -phrase
	$-er_{\text{Kennedy}(1997)}$	genitive-marked phrase

Table 4: Comparison of the Inventory of Operators in Russian and Greek

#### 3.1 Adverbial Genitive-Marked Comparatives in Russian and Greek

The two languages differ, however, when it comes to adverbial comparatives in which the standard is marked by genitive case. Consider the minimal pair in (26) and (27). While the relevant adverbial comparative is ungrammatical in Greek, it is perfectly natural in Russian. Another example from Russian is in (28). How can we explain this contrast?

(26) **Greek:**

I Maria pezi kithara kalitera {apo mena/\***mu**}.  
 the Mary plays guitar better {from me/**me**<sub>GEN</sub>}  
 ‘Maria plays the guitar better than me.’  
 (Merchant 2012: 6)

(27) **Russian:**

Maša igraet na gitare lučše **menja**.  
 Masha plays on guitar<sub>PREP</sub> better **me**<sub>GEN</sub>  
 ‘Masha plays the guitar better than me.’

(28) Maša bežala bystree Vani.  
 Masha ran fast<sub>COMP</sub> Vanja<sub>GEN</sub>  
 ‘Masha ran faster than Vanja.’

In order to explain this contrast, let us first consider the semantics underlying adverbial comparatives. Adverbial comparatives actually require a slightly different analysis than the predicative and attributive cases discussed above, an analysis, which takes into account the fact that what is compared in (27) and (28) are events (music sessions, running). Their phrasal analysis thus requires a somewhat different operator. We are not aware of any such analysis in the literature. Here’s what we will therefore do: We will first suggest an adequate operator for phrasal adverbial comparatives, which is derived from but not identical to *-er*<sub>Kennedy(1997)</sub>. We then put it to work. Subsequently, we explain the contrast between Russian and Greek as a case of lexical variation: Russian decided to add this extended operator to its lexical inventory, while Greek did not.

If we consider the comparison in (28), the relation underlying this comparison is the one in (29) rather than the relation lexically provided by the adverbial, in (30). The standard of the comparison, Vanja, is mapped by (29) onto his running event, whose speed is then measured.

(29)  $\lambda d_{<d>}. \lambda z_{<e>}. \lambda e_{<v>}. \text{run}(e)(z) \ \& \ \text{SPEED}(e) \geq d$

(30)  $\lambda d_{<d>}. \lambda e_{<v>}. \text{SPEED}(e) \geq d$

We suggest in (31) a phrasal, adverbial operator (*-er*<sub>Kennedy(1997)-adverbial</sub>) which requires a relation such as (29) as its first argument. The operator is parallel to *-er*<sub>Kennedy(1997)</sub> as far as its argument structure is concerned, merely enriched with events. The type of semantic ellipsis we do in the case of phrasal comparatives requires that the adverbial operator introduce

and existentially close the event associated with the genitive-marked standard. In the case of (28), this is the running event associated with Vanja, for example. (It is also conceivable that the operator presupposes the existence of such an event rather than asserts it.)

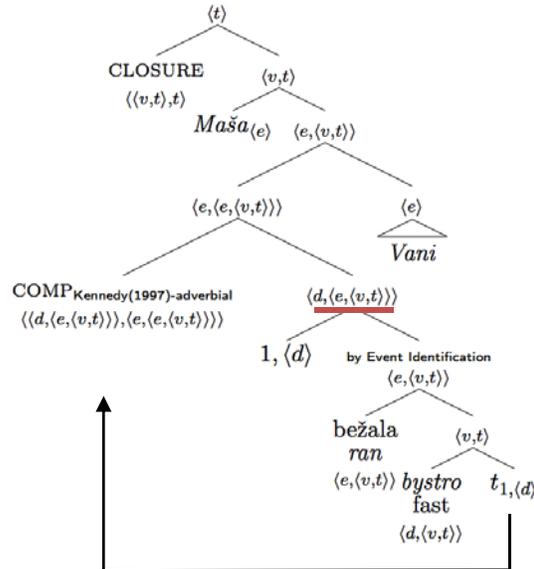
- (4)  $\llbracket -er_{\text{Kennedy}(1997)} \rrbracket = \lambda R_{\langle d, \langle e, t \rangle \rangle}. \lambda y_{\langle e \rangle}. \lambda x_{\langle e \rangle}. \text{MAX}(\lambda d. R(d)(x)) > \text{MAX}(\lambda d'. R(d')(y))$
- (31)  $\llbracket -er_{\text{Kennedy}(1997)\text{-adverbial}} \rrbracket = \lambda R_{\langle d, \langle e, \langle v, t \rangle \rangle \rangle}. \lambda y_{\langle e \rangle}. \lambda x_{\langle e \rangle}. \lambda e_{\langle v \rangle}. \exists e' [\text{MAX}(\lambda d. R(d)(x)(e)) > \text{MAX}(\lambda d'. R(d')(y)(e'))]$

Let's apply (31) to our example in (28), which has the Logical Form in (32). A couple of remarks on this LF might be helpful, bottom to top. As indicated in (30), we assume gradable adverbials to be of type  $\langle d, \langle v, t \rangle \rangle$ . The verb combines with the adverbial phrase via Event Identification (Kratzer 2003). Type mismatch forces the degree operator to move from its base position. It is this movement which creates the relation in (29). Note that while adverbial comparatives thus require movement at Logical Form, this movement is not parasitic, as it would have to be in the attributive case. (We will come back to this below.) We neglect the contribution of aspect and tense, and merely assume here an operator, CLOSURE, in (33), which existentially quantifies off the event argument (Heim 1982, 1988). The resulting truth conditions are in (34).

<b>Russian</b>	$-er_{\text{clausal}}$	<i>čem</i> -clause
	$-er_{\text{Kennedy}(1997)}$	genitive-marked phrase
	$-er_{\text{Kennedy}(1997)\text{-adverbial}}$	genitive-marked phrase
<b>Greek</b>	$-er_{\text{clausal}}$	<i>ap'oti</i> -clause
	$-er_{\text{Heim}(1985)}$	<i>apo</i> -phrase
	$-er_{\text{Kennedy}(1997)}$	genitive-marked phrase

Table 5: Revised Inventory of Operators in Russian and Greek

(32)

(33)  $\llbracket \text{CLOSURE} \rrbracket = \lambda P_{\langle v, t \rangle}. \exists e [P(e)]$ (34)  $\exists e, e' [\text{MAX}(\lambda d. \text{run}(e)(\text{Mary}) \ \& \ \text{SPEED}(e) \geq d) > \text{MAX}(\lambda d'. \text{run}(e')(\text{John}) \ \& \ \text{SPEED}(e') \geq d)]$ 

‘There are two events  $e$  and  $e'$  such that the maximal speed of Mary’s running event  $e$  exceeds the maximal speed of John’s running event  $e'$ .’

Back to the crosslinguistic picture. We suggest that the variation we observe between Russian and Greek is a case of lexical variation. The two languages differ in the inventory of phrasal operators which they have at their disposal, as outlined in Table 5. It thus appears that languages might choose whether or not to extend  $-er_{\text{Kennedy}(1997)}$  to the domain of eventualities.

Before we conclude, let us briefly comment on the movement observed in the Logical Form in (32). What we see is that Kennedy-style schönfinkeled operators are not generally banned from moving. These phrasal operators thus have some scopal mobility, contra Beck, Hohaus & Tiemann (2012). The VP-internal movement  $-er_{\text{Kennedy}(1997)\text{-adverbial}}$  undergoes in (32) is fine. Any parasitic movement, as we have seen for the attributive comparatives with  $-er_{\text{Kennedy}(1997)}$  in (11), is however not

possible. It is also not possible in the case of *-er*<sup>Kennedy(1997)-adverbial</sup>: Inverse-scope readings with modals always require parasitic movement and are thus expected to be also unavailable with adverbial phrasal comparatives. This expectation is borne out. An example is in (35).

- (35) Katja *xočet* *igrat'* na gitare *lučše* Ziny.  
 Katja wants play on guitar<sub>PREP</sub> better Zina<sub>GEN</sub>  
 = 'Katja wants to play the guitar better than Zina plays the guitar.' [want>-er]  
 ≠ 'Katja wants to play the guitar better than Zina wants to play the guitar.' [-er >want]

The distinction between Greek and Russian thus did not only prompt us to develop an analysis of adverbial comparatives with phrasal operators, it also allowed us to better understand the restrictions on their movement.

#### 4 Concluding Remarks

Let us retrace our steps: We started out with a brief investigation of different phrasal comparison operators which have been proposed in the literature. More specifically, we wondered whether keeping *-er*<sup>Kennedy(1997)</sup> in our inventory of degree operators was necessary as this operator has only limited applicability. Russian suggests it is: Genitive-marked comparatives in Russian are best analyzed with this operator because they (i) do not allow for clausal standards, (ii) only have an internal reading when used attributively, and (iii) do not exhibit scope ambiguities. In those respects, genitive-marked comparatives in Russian behave like their Greek counter-parts. However, the two languages differ with respect to the acceptability of adverbial comparatives with genitive-marked standards. We analyze this difference a variation in the functional lexicon: In addition to the individual-based phrasal operator *-er*<sup>Kennedy(1997)</sup>, Russian has an event-based phrasal operator schönfinkeled like *-er*<sup>Kennedy(1997)</sup>, while Greek has not.

The case of the Russian phrasal comparative also teaches us a lesson about crosslinguistic variation: The way languages compositionally arrive at truth conditions that are absolutely identical varies considerably, but systematically. Careful, theoretically motivated elicitation of cross-linguistic data can unmask these multiple roads to identical meanings.

## References

- Beck, Sigrid, Vera Hohaus & Sonja Tiemann. 2012. A note on phrasal comparatives. *Proceedings of Semantics and Linguistic Theory (SALT) 22*, 146-165.
- Beck, Sigrid. 2011. Comparison constructions. In Claudia Maienborn, Klaus von Stechow & Paul Portner (eds.), *Semantics: An International Handbook of Natural Language Meaning*, 1341-1389. Berlin: De Gruyter Mouton.
- Beck, Sigrid, Sveta Krasikova, Daniel Fleischer, Remus Gergel, Christiane Savelsberg, John Vandereelst & Elisabeth Villalta. 2009. Crosslinguistic variation in comparison constructions. *Linguistic Variation Yearbook 9*, 1-66.
- Beck, Sigrid & Uli Sauerland. 2000. Cumulation is needed: a reply to winter (2000). *Natural Language Semantics 8*, 349-371.
- Berezovskaya, Polina (2014). Acquisition of Russian degree constructions: a corpus-based study. *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 22*, 1-22.
- Bhatt, Rajesh & Shoichi Takahashi. 2011. Reduced and unreduced phrasal comparatives. *Natural Language and Linguistic Theory 29*, 581-620.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Malden, MA/Oxford, UK: Blackwell.
- Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts at Amherst.
- Heim, Irene. 1985. *Notes on Comparatives and Related Matters*. Ms., University of Texas at Austin.
- Heim, Irene. 1988. *The Semantics of Definite and Indefinite Noun Phrases*. New York: Garland Publishers.
- Hofstetter, Stefan. 2009. Comparison in Turkish: a rediscovery of the phrasal comparative. *Proceedings of Sinn und Bedeutung (SuB)13*, 191-205.
- Hofstetter, Stefan. 2010. Comparison in Thai. Poster, *Linguistic Evidence*, Tübingen.
- Hohaus, Vera, Sonja Tiemann & Sigrid Beck. 2014. Acquisition of comparison constructions. *Language Acquisition 21: 3*, 215-249.
- Karvovskaya, Lena. 2013. *Comparatives in Persian (Farsi), Tajiki and Ishkashimi: A Cross-linguistic Analysis*, Master thesis, Universität Potsdam.

- Kennedy, Chris. 1997. *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. Ph.D. thesis, University of California at Santa Cruz.
- Krasikova, Sveta. 2007. *Comparative Constructions in Russian*. Ms., Universität Tübingen.
- Kratzer, Angelika. 2003. *The Event Argument and the Semantics of Verbs*. Ms., University of Massachusetts at Amherst.
- Lerner, Jan & Werner Pinkal. 1995. Comparative ellipsis and variable binding. *Proceedings of Semantics and Linguistic Theory (SALT) 5*, 222-236.
- Matushansky, Ora. 2002. More of a Good Thing: Russian Synthetic and Analytic Comparatives. *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 10*, 143-162.
- Merchant, Jason. 2009. Phrasal and clausal comparatives in Greek and the abstractness of syntax. *Journal of Greek Linguistics 9*, 134-164.
- Merchant, Jason. 2011. *Genitives of Comparison in Greek*. Ms., University of Chicago.
- Merchant, Jason. 2012. *The Two Phrasal Comparatives of Greek*. Ms., University of Chicago.
- Nissenbaum, Jon. 2000. Covert movement and parasitic gaps. *Proceedings of NELS 30*, 541-555.
- Pancheva, Roumyana. 2006. Phrasal and clausal comparatives in Slavic. *Proceedings of Formal Approaches to Slavic Linguistics (FASL) 14*, 236-257.
- Tiemann, Sonja. 2009. *The Acquisition of Comparison: A Corpus-Based Study*. Magisterarbeit, Universität Tübingen.
- Tiemann, Sonja, Vera Hohaus and Sigrid Beck. 2012. Crosslinguistic variation in comparison: evidence from child language acquisition. In Britta Stollerfoht & Sam Featherston (eds.), *Empirical Approaches to Linguistic Theory: Studies of Meaning and Structure*, 115-146. Berlin: De Gruyter Mouton.
- von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics 3*, 1-77.

polina.berezovskaya@uni-tuebingen.de  
vera.hohaus@uni-tuebingen.de

## **On Multiple Left-Branch Dislocation: Multiple Extraction and/or Scattered Deletion?\***

*Željko Bošković*  
University of Connecticut

It is well known that Serbo-Croatian (SC) allows left-branch extraction (LBE), i.e. extraction of an NP edge. This paper focuses on examples with multiple NP-edges, as in (1). As shown in Bošković (in press a) and illustrated in (2), more than one NP-edge can be separated from the NP in addition to single NP-edge extraction (3). (I will refer to (2) as multiple left-branch dislocation (MLD)).

- (1) Prodaje onu staru kuću.  
sells that old house  
'He is selling that old house.'
- (2) Onu staru prodaje kuću.  
that old sells house
- (3) Onu<sub>i</sub> prodaje t<sub>i</sub> (staru) kuću.  
that sells old house

Bošković (in press a) examines such examples in some detail, but leaves several issues unresolved. The goal of this paper is to examine how MLD should be analyzed, investigating the viability of an analysis of MLD that was not considered in Bošković (in press a).

---

\*The paper is based upon work supported by the NSF under Grant BCS-0920888. For helpful comments and suggestions, I thank an anonymous reviewer and the participants of FASL 23 and my UConn seminars.

## 1 Multiple Left-Branch Dislocation Constructions

Bošković (2014a, in press a) analyzes (2) as involving multiple application of focus movement, with each application left-branch extracting one element, LBE being allowed in SC.

- (4) Onu<sub>i</sub> staru<sub>i</sub> prodaje t<sub>i</sub> t<sub>j</sub> kuću.  
 that old sells house
- (5) a. Onu<sub>i</sub> prodaje t<sub>i</sub> kuću.  
 b. Staru<sub>i</sub> prodaje t<sub>i</sub> kuću.

I will consider here the possibility of an alternative analysis where (2) involves a single application of focus movement and scattered deletion.

- (6) [Onu staru ~~kuću~~]<sub>i</sub> prodaje [~~onu—staru~~ kuću]<sub>i</sub>.  
 that old house sells

I will start the discussion by pointing out some potential problems for the multiple focus LBE analysis of MLD examples like (2) (though see section 2 for ways of dealing with the issues in question under the multiple LBE analysis).

First, while SC multiple wh-fronting constructions (MWF) like (7) have been argued to involve multiple focus movement (see Bošković 2002, Stjepanović 1999), which indicates that multiple focus movement is in principle allowed in SC, multiple focus movement of non-wh-phrases is generally disallowed (the judgment in (7b) holds for the multiple-focus reading).

- (7) a. [<sub>FocP</sub> Kome koga/koga kome [<sub>Foc'</sub> on predstavlja]]?  
 who<sub>DAT</sub> who<sub>ACC</sub> he is-introducing  
 ‘Who is he introducing to whom?’
- b. \* [<sub>FocP</sub> Petru Mariju/Mariju Petru [<sub>Foc'</sub> on predstavlja]].  
 Peter<sub>DAT</sub> Marija<sub>ACC</sub> he is-introducing  
 ‘He is introducing Marija to Peter.’

Second, as noted above, MWF has been argued to involve multiple focus movement. Based on MWF, Bošković (2002) shows that multiple focus

movement is free of ordering constraints; thus, either order is acceptable in (7a). This is not the case with the MLD construction in (8).

- (8) a. Onu staru prodaje kuću.  
       that old sells house  
       b. \*Staru onu prodaje kuću.

In some MWF languages, MWF is subject to ordering constraints, i.e. superiority effects. This is for example the case with Bulgarian (see Bošković 2002 for an account of the SC/Bulgarian difference regarding superiority). However, even in Bulgarian, in examples with three wh-phrases the second and the third wh-phrase are freely ordered (compare (9b) and (9d); see Bošković 2002 for an account of this selective superiority effect).

- (9) a. Kogo kakvo e pital Ivan?  
       whom what is asked Ivan  
       ‘Who did Ivan ask what?’  
       b. ?\*Kakvo kogo e pital Ivan?  
       c. Koj kogo kakvo e pital?  
       who whom what is asked  
       ‘Who asked who what?’  
       d. Koj kakvo kogo e pital? (Bulgarian)

However, with MLD strict ordering holds even for the cases with three dislocated left-branches. (10) gives the only allowed word order for *onog neozbiljnog mašinskog*. There is thus no selective superiority effect with MLD. More generally, the ordering effects with MLD do not correspond to those found with MWF.

- (10) On otpušta onog neozbiljnog mašinskog tehničara.  
       he is-firing that not-serious mechanical technican  
       (11) a. ?\*Onog mašinskog neozbiljnog otpušta tehničara.  
           that mechanical not-serious is-firing technican  
       b. ? Onog neozbiljnog mašinskog otpušta tehničara.  
           that not-serious mechanical is-firing technican

Next, a clitic (*je*) cannot follow a sequence of two fronted *wh*-phrases, which, as noted above, undergo independent focus movements. However, a clitic can follow fronted elements with MLD. Under the standard assumption that SC clitics follow either the first word or the first constituent of their sentence, this indicates that the elements preceding the clitic form a constituent in (13) but not in (12).

- (12)?\*Ko koga je vidio?  
 who whom is seen  
 ‘Who saw whom?’
- (13) Malu žutu je kupio kuću.  
 small yellow is bought house  
 ‘He bought a small, yellow house’

The above discussion raises potential issues for the focus movement treatment of MLD. There is also a potential argument that MLD does not involve LBE, more precisely, that MLD should not be treated in the same way as LBE. With simple LBE, the remnant can be placed either before or after the verb, as in (14) (most speakers in fact prefer (14a)). In MLD, the remnant needs to follow the verb, as shown by (15).

- (14) a. Žutu mu kuću pokazuje.  
 yellow him house is-showing  
 ‘He is showing him the yellow house.’  
 b. Žutu mu pokazuje kuću.
- (15) a.?\*Onu žutu mu kuću pokazuje.  
 that yellow him house is-showing  
 b. Onu žutu mu pokazuje kuću.

MLD thus does not behave like LBE in this respect.

Consider now the nature of the restriction that is responsible for the effect in (15), since it will be important for the scattered deletion analysis of MLD. Bošković (2014a) argues that what we are dealing with here is a discourse requirement on MLD; the fronted elements are interpreted as focalized, and the remnant is backgrounded. Backgrounded elements follow the verb in SC, hence the contrast in (15). Bošković also observes that this analysis can account for the contrast in (16)-(17), the

backgrounding requirement being the reason why intensifying/focalizing adverbs cannot occur in the remnant.

- (16) ?Onu tamnu prodaje plavu kuću.  
 that dark is-selling blue house
- (17) ?\*Onu tamnu prodaje izuzetno plavu kuću.  
 that dark is-selling extremely blue house

## 2 The Scattered Deletion Analysis

Having discussed potential problems for the multiple LBE analysis of MLD, in this section I examine the viability of the alternative, scattered deletion account of MLD.

While examples like (18) are standardly analyzed as involving subextraction of *malu*, there are alternative accounts of such examples (though, as discussed in the references cited below, they all face very serious problems). Thus, Fanselow and Čavar (2002) argue that (18) involves full NP fronting+scattered deletion; one part of the fronted NP being pronounced in the fronted and one part in a lower position, as in (19).<sup>1</sup>

- (18) Malu<sub>i</sub> je kupio [t<sub>i</sub> kuću]  
 small is bought house  
 ‘He bought a small house.’
- (19) [Malu ~~kuću~~] je kupio [~~malu~~ kuću]  
 small is bought house

What is of interest here is Franks’ (1998) claim that pronunciation of a lower copy is possible if and only if higher copy pronunciation would lead to a PF violation. There is ample motivation for this claim (see e.g. Bošković 2001 and Bošković and Nunes 2007), which also follows from independent mechanisms, as shown by Nunes (2004). While PF considerations typically force lower pronunciation of the full copy of the fronted constituent, there are cases where PF considerations require

---

<sup>1</sup>Another alternative is remnant movement, as in Abels (2003) and Franks and Progovac (1994); see Bošković (2005), Stjepanović (2010, 2011), and Talić (2013) for evidence against this analysis.

scattered deletion, as with some instances of cliticization in Bulgarian and Macedonian.

Consider the basic cliticization pattern in Bulgarian and Macedonian.

(20)					Bulgarian	Macedonian	
	a.	Petko	<i>mi go</i>	dade včera.	OK	OK	
		Petko	me <sub>DAT</sub> it <sub>ACC</sub>	gave yesterday			
		‘Petko gave me it yesterday.’					
	b.	Včera	<i>mi go</i>	dade Petko.	OK	OK	
	c.	<i>Mi go</i>	dade	Petko včera.	*	OK	
	d.	Dade	<i>mi go</i>	Petko včera.	OK	*	
	e.	Včera	dade	<i>mi go</i> Petko.	*	*	

In this context Macedonian clitics always precede the verb, while Bulgarian clitics precede the verb unless that ordering of clitics with respect to the verb would leave clitics sentence initial. In that case Bulgarian clitics follow the verb. Bošković (2001) proposes a lower copy pronunciation account of these facts based on Franks’ proposal regarding when lower copy pronunciation is allowed. In both Bulgarian and Macedonian the clitics move in front of the verb. Now, it is well-known that Bulgarian clitics are enclitics, and Macedonian clitics are proclitics (in this context). Nothing then goes wrong if the highest copy of the clitics is pronounced in Macedonian, which then must happen. In Bulgarian, this holds for the cases where something precedes the clitic in the raised position. If that is not the case, pronouncing the highest clitic copy would lead to a violation of their enclitic PF requirement. The lower copy of the clitic is then pronounced in this case, which then correctly gives us the V-clitic order only for the context where nothing precedes the verb.

(21) *Bulgarian*: a. [X clitic<sup>i</sup> V elite<sup>i</sup>]                      b. [elite<sup>i</sup> V clitic<sup>i</sup>]

(22) *Macedonian*: [(X) clitic<sup>i</sup> V elite<sup>i</sup>]

Bošković (2001) shows that this analysis leads to scattered deletion in certain cases. Main verbs and auxiliary/pronominal clitics form a complex head in Bulgarian and Macedonian, so that the verb carries the clitics along when undergoing head-movement, as in the *li* construction. In (23a), this complex head left-adjoins to *li*, with the head of its chain

pronounced. This pronunciation is, however, not possible in Bulgarian (23b), since *si mu gi* as well as *li* are enclitics. The only way to satisfy the enclitic requirement here is via scattered deletion, as in (25b), which yields (24b). Since nothing goes wrong with full higher copy pronunciation in Macedonian (25a), this is then the only option, hence the ungrammaticality of (24a).

- (23) a. *Si mu gi dal li* parite? (Macedonian)  
 are him<sub>DAT</sub> them given Q the-money
- b. \**Si mu (gi) dal li* parite? (Bulgarian)  
 are him<sub>DAT</sub> them given Q the-money  
 ‘Have you given him the money?’
- (24) a. \**Dal li si mu gi* parite? (Macedonian)  
 given Q are him<sub>DAT</sub> them the-money
- b. *Dal li si mu (gi)* parite? (Bulgarian)  
 given Q are him<sub>DAT</sub> them the-money  
 ‘Have you given him the money?’
- (25) a. *[[si mu gi dal] li [si mu gi dal] parite]* (Macedonian)
- b. *[[~~si mu gi~~ dal] li [si mu gi dal] parite]* (Bulgarian)

What this indicates is that scattered deletion is in principle possible. There are, however, many well documented problems with the scattered deletion analysis of (18) which show that the analysis cannot be maintained: it simply does not hold up empirically (see also the discussion below). Thus, Bošković (2005) shows that the analysis has a very serious overgeneration problem, considerably overgenerating the available splits. Stjepanović (2010) shows that the analysis fails to account for the available readings of multiple questions involving LBE and Stjepanović (2011) shows that it does not account for crossing restrictions in negative concord constructions. The most glaring problem is that scattered deletion is basically a last resort mechanism. While it is in principle available, it is severely constrained: it takes place only if full deletion is not possible. This is e.g. the reason why it is disallowed in (26).

- (26) \**[That student]<sup>i</sup> was arrested [that student]<sup>i</sup>*

In (19), full deletion is obviously possible, hence scattered deletion should be disallowed. While this rules out the scattered deletion analysis

of simple LBE cases like (18) the problem actually does not arise with MLD: full deletion may in fact not be an option with MLD on the relevant reading.

- (27) [Onu žutu ~~kuću~~] je kupio [~~onu žutu~~ kuću]  
           that yellow is bought house

Consider (27) in light of the discourse requirement on MLD where one part of the NP is focalized and one part is backgrounded. The requirement cannot be met if *kuću* is pronounced in the focus position, where the full NP [*onu žutu kuću*] moves. *Kuću* may then be pronounced in its base position following the verb to meet the backgrounding requirement.

Recall now that under Franks' proposal, only PF considerations can sanction lower copy pronunciation. Stjepanović (1999) shows that stress assignment can also cause lower copy pronunciation. The relevant discourse properties have PF reflexes in terms of stress (emphatic stress vs normal stress vs distressing), which can motivate lower copy pronunciation here.

The scattered deletion analysis thus seems to be a viable option for analyzing MLD. In fact, it resolves all the potential problems for the multiple focus/left-branch extraction analysis, noted above. 1. Under the scattered deletion analysis, MLD does not involve otherwise disallowed multiple focus-movement of non-wh-phrases (cf. (7b)). 2. There is no superiority issue because there is no multiple movement. The fronted part then has to preserve the base-generated order ((8), (11)). 3. While under the multiple Spec analysis of MWF (see Koizumi 1994, Richards 2001), two separate constituents precede the clitic in (12), which is disallowed, only one precedes it in (13) under the scattered deletion analysis. 4. The contrast in (14)-(15) also follows from the scattered deletion analysis, where (14), but not (15), involves subextraction.

A question, however, arises here. As discussed in Bošković (2014a, in press a), it is actually very hard to block the multiple LBE analysis theoretically. Can the multiple LBE derivation, adopted in Bošković (2014a, in press a), then still be available for the MLD construction?

In fact, the discussion in Bošković (2014a, in press a) indicates that most of the issues noted above can be handled under the multiple LBE analysis, though with some additional assumptions that are not needed under the scattered deletion analysis. Thus, Bošković analyzes the con-

trast between (4) and (7b) as involving a semantic effect. In particular, Bošković claims that focalized elements in a multiple non-wh focus movement construction must have a single referent, which is the case in (4), but not (7b). It is in fact clear that there are additional pragmatic/semantic requirements on MLD, e.g., deicticity, as shown below ((28) actually improves with pointing).

- (28)?\*Malu plavu mu pokazuje kuću.  
 small blue him<sub>DAT</sub> is-showing house  
 ‘She is showing him a small blue house.’
- (29) Onu malu plavu mu pokazuje kuću.  
 that small blue him<sub>DAT</sub> is-showing house
- (30) \*Male plave ga ne zanimaju kuće.  
 small blue him<sub>ACC</sub> not interest house  
 ‘Small blue houses don’t interest him.’

Regarding Superiority, Bošković (2014a, in press a) follows the standard assumption that what is responsible for Superiority effects (i.e. free/fixed order of fronted wh-phrases) with MWF is Attract Closest. However, he argues that what is responsible for the fixed order of fronted elements in MLD, i.e. (8), is the Phase-Impenetrability Condition, given Bošković’s proposal that in phases with multiple edges, only the outmost edge counts as the phasal edge for the purpose of the PIC.<sup>2</sup> Further, Bošković argues that just like traces do not count as interveners for relativized minimality (see Chomsky 1995, Bošković 2011), they do not count as edges for the purpose of the PIC. Consider in this respect (31).<sup>3</sup>

- (31) a. Onu<sub>i</sub> prodaje t<sub>i</sub> staru kuću.  
 that sells old house  
 b. \*Staru<sub>i</sub> prodaje onu t<sub>i</sub> kuću.

<sup>2</sup>For additional evidence for the proposal, see Wurmbrand (2013), Zanon (in press), and Yoo (2015).

<sup>3</sup>The underlying assumptions in the following discussion are that SC lacks DP, as a result of which demonstratives as well as adjectives are NP-adjoined in SC (see Bošković 2012), and that the highest projection in the extended domain of N (in fact any lexical category) functions as a phase (see Bošković 2014b), which makes NP a phase in SC (due to the absence of DP).

The second NP-adjoined element, *staru*, in (31) is not at the edge of the NP, hence cannot move, until the first element moves. After *onu* moves, *staru* can move without violating the PIC, tucking in under *staru* (see Richards 2001), which results in fixed word order in (4)/(8). The same holds for (11).

As for the potential problem for the multiple LBE analysis noted above regarding (14)-(15), the issue here may simply be the discourse requirement on MLD. MLD and simple LBE have different discourse requirements, which can be implemented as a filtering effect in the case of MLD that rules out in semantics/pragmatics certain constructions (namely (15a)) that are syntactically well-formed.

The clitic placement issue is, however, real. Given the nature of SC cliticization, where what precedes the clitic must be a constituent, clitics force constituency on the fronted elements in MLD. Accommodating the contrast in (12)-(13) under the multiple LBE analysis then becomes non-trivial. Here is one possibility: Rudin (1988) argues that multiple movement to the same projection found in MWF constructions involves right-adjunction of the element that moves second to the first fronted element. Koizumi (1994), on the other hand, argues that such cases involve multiple specifiers. Given that only the first analysis treats the fronted elements as a syntactic constituent, if we assume straightforward syntax-phonology mapping here which preserves syntactic constituency it may be that both the Rudin option and the Koizumi option are available, with MWF involving the latter and MLD the former. Since the fronted elements are then a constituent only with MLD, placing a clitic following the fronted sequence is then possible only with MLD.

Another option could be to adopt Rudin's (1988) treatment of SC MWF where the first fronted wh-phrase is located in SpecCP and the second one in a lower position below the CP projection, which can be the focus position as in Bošković's (2002) analysis. Both fronted elements would then be located in the focus position in the MLD case, since there is obviously no wh-movement here. In fact, the MWF construction could involve multiple focus movement, just like the MLD construction, followed by wh-movement of one wh-phrase. The analysis can rather easily capture the contrast in (12)-(13). However, it does raise some issues, for example, how to deal with Bošković's (2002) arguments that SC MWF in contexts like (12) at least does not need to involve wh-movement and the issue raised by the freezing/criterial effect (Rizzi

2006, Bošković 2003, 2008), which is standardly assumed to ban further movement from criterial positions like SpecFocP.

Another possibility would be to appeal to a filtering effect of prosody. As discussed in Bošković (2001), the constituency requirement on SC clitics is actually prosodic: what precedes them (within their intonational phrase) must be a prosodic constituent (see also Bošković in press b). It is then possible that MWF cases like (12) and MLD cases like (13) involve the same syntactic derivation, i.e. they both involve multiple focus movement. However, possibly due to prosodic peaks, or more generally prosodic properties of *wh*-phrases, the fronted *wh*-phrases here cannot be parsed into a single prosodic constituent, while the fronted non-*wh*-phrases can be. This would push the account of (12)-(13) into PF, i.e. the prosodic component.

If one of these options for analyzing the clitic cases in (12)-(13) can be developed there would be no need for the scattered deletion analysis of MLD with respect to the data discussed so far since the multiple LBE analysis would be able to handle all of them. However, if it turns out that none of the above options for analyzing the clitic cases in (12)-(13) under the multiple LBE analysis of MLD can be taken, scattered deletion may be required. In fact, in light of the above discussion, it would then be possible that while MLD in principle can involve either multiple LBE or scattered deletion, when a clitic is present only the latter would converge.

There is, however, another way of teasing apart the multiple LBE and the scattered deletion analysis of MLD. As discussed in Bošković (2012 and references therein), adjectival left-branch extraction is found only in languages without articles. However, Bošković (2013) observes an additional requirement on adjectival LBE: even in languages like SC which allow left-branch extraction only agreeing adjectives can undergo such extraction, as illustrated by (32)-(33). Both *braon* and *smedja* mean “brown”. While *braon* does not decline, hence does not agree with the noun it modifies, *smedja* does agree. *Bež* also does not decline/agree with the noun, just like *braon*. The contrast in (32)-(33) thus indicates that only agreeing adjectives undergo left-branch extraction.

- (32) ?\*Bež/braon<sub>i</sub> je on kupio t<sub>i</sub> kola.  
 Beige/brown is he bought car  
 ‘He bought a brown/beige car.’

- (33) Smedja<sub>i</sub> je on kupio t<sub>i</sub> kola.  
brown is he bought car

Observing that non-inflected adjectives must be adjacent to the noun in cases where both an inflected and a non-inflected adjective modify the same noun (34), that they cannot be used in color-combinations with inflected adjectives (35), and that, in contrast to inflected adjectives, they do not allow ellipsis of the noun they modify (36), Bošković argues that non-inflected/non-agreeing adjectives like *braon* and *bež* have a different structural status from inflected/agreeing adjectives; in particular, they are head-adjoined (i.e. they are adjoined to N), hence they cannot undergo left-branch extraction, which is a phrasal movement (the analysis also captures the facts in (34)-(36), see Bošković 2013).

- (34) a. ?\*braon/bež plastična kola  
brown/beige plastic car  
b. plastična braun/bež kola  
c. smedja plastična kola  
brown plastic car
- (35) a. ?\*plavo-braon                      b. plavo-smedja  
blue brown                              blue brown  
c. bež-braon                              d. ?\*bež-smedja  
beige brown                              beige brown
- (36) On nam je pokazao plavu kuću, a ona nam je pokazala  
he US<sub>DAT</sub> is shown blue house and she US<sub>DAT</sub> is shown  
crvenu/\*bež  
red/beige

A question now arises what happens with adjectives like *braon* and *bež* in MLD configurations. If MLD can only be derived via multiple LBE, we would expect MLD examples involving *braon* and *bež* to be as degraded as (32). On the other hand, if a scattered deletion derivation is an option for MLD we may expect (32) to improve in an MLD configuration. Although the relevant judgments are rather subtle, all the

speakers found (37b) to be better than (37a). (There is no such contrast in (38)).<sup>4</sup>

- (37) a. ?\*Bež/braon mu pokazuje kuću.  
 beige/brown him<sub>DAT</sub> is-showing house  
 ‘He is showing him a beige/brown house’  
 b. ?Onu bež/braon mu pokazuje kuću.  
 that beige/brown him<sub>DAT</sub> is-showing house  
 ‘He is showing him that beige/brown house.’
- (38) a. Smedju mu pokazuje kuću.  
 brown him<sub>DAT</sub> is-showing house  
 b. Onu smedju mu pokazuje kuću.  
 that brown him<sub>DAT</sub> is-showing house

Assuming that the LBE derivation is ruled out for both (37a) and (37b) for the reason discussed above, the data under consideration can be captured if MLD also has at its disposal the scattered deletion option. The scattered deletion derivation can then be responsible for the improved status of (37b). It should be emphasized here that the current discussion provides additional evidence against the scattered deletion derivation for simple LBE cases: the scattered deletion derivation is available in (37b), but crucially not in (37a).

The remaining issue is that while (37b) is better than (37a), (38b) is still slightly better than (37b). It is not clear why this is the case. One possibility is that the scattered deletion derivation of MLD itself is slightly dispreferred. ((37b) can only be derived via scattered deletion, while (38b) can in principle involve multiple LBE.)

At any rate, what is important for us is that (37) represents another case where LBE and MLD behave differently, which suggests that the two should be treated differently.

---

<sup>4</sup>The above reports preliminary results, with the judgments of four linguists, Aida Talić, Sandra Stjepanović, Miloje Despić, and myself. Obviously, additional data verification is needed here. Given the discussion below, one might expect the contrast in (37) to be even sharper. It is possible that the relative complexity of MLD constructions (in comparison with simple LBE constructions) interferes in a direct comparison of the two by favoring the latter (see also the point made below regarding (37b)).

### 3 Conclusion

While the situation is certainly not crystal clear, given the data discussed in this paper and the theoretical status of the relevant mechanisms, it appears that MLD can in principle involve either multiple LBE or scattered deletion (the latter is not available in simple LBE cases).

### References

- Abels, Klaus. 2003. Successive Cyclicity, Antilocality, and Adposition Stranding. Doctoral dissertation, University of Connecticut, Storrs.
- Bošković, Željko. 2001. *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Oxford: Elsevier.
- Bošković, Željko. 2002. On multiple wh-fronting. *Linguistic Inquiry* 33:351-383.
- Bošković, Željko. 2003. On wh-islands and obligatory wh-movement contexts in South Slavic. In *Multiple Wh-Fronting*, ed. by Cedric Boeckx and Kleanthes Grohmann, 27-50. Amsterdam: John Benjamins.
- Bošković, Željko. 2005. On the locality of Left Branch Extraction and the structure of NP. *Studia Linguistica* 59:1-45.
- Bošković, Željko. 2008. On the operator freezing effect. *Natural Language and Linguistic Theory* 26:249-287.
- Bošković, Željko. 2011. Rescue by PF deletion, traces as (non)-interveners, and the *That-Trace* Effect. *Linguistic Inquiry* 42:1-44.
- Bošković, Željko. 2012. On NPs and clauses. In *Discourse and Grammar*, ed. by Günther Grewendorf and Thomas Ede Zimmermann, 179-242. Berlin: De Gruyter Mouton.
- Bošković, Željko. 2013. Adjectival escapades. *Proceedings of Formal Approaches to Slavic Linguistics* 21:1-25.
- Bošković, Željko. 2014a. More on the edge of the edge. *Proceedings of Formal Approaches to Slavic Linguistics* 22:44-66.
- Bošković, Željko. 2014b. Now I'm a phase, now I'm not a phase: on the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45:27-89.
- Bošković, Željko. in press a. Getting really edgy: on the edge of the edge. *Linguistic Inquiry*.

- Bošković, Željko. in press b. On prosodic boundaries. In *Formal Studies in Slavic Linguistics. Proceedings of Formal Description of Slavic Languages* 10.
- Bošković, Željko and Jairo Nunes. 2007. The copy theory of movement: a view from PF. In *The Copy Theory of Movement*, ed. by Norbert Corver and Jairo Nunes, 13-74. Amsterdam/Philadelphia: John Benjamins.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Fanselow, Gisbert, and Damir Ćavar. 2002. Distributed deletion. *Theoretical Approaches to Universals*, ed. by Artemis Alexiadou, 65-107. Amsterdam/Philadelphia: John Benjamins.
- Franks, Steven. 1998. Clitics in Slavic. Presented at the Comparative Slavic Morphosyntax Workshop, Spencer Creek, Indiana.
- Franks, Steven, and Ljiljana Progovac. 1994. On the placement of Serbo-Croatian clitics. *Indiana Linguistic Studies* 7: 69-78.
- Koizumi, Masatoshi. 1994. Layered specifiers. In *Proceedings of the North East Linguistic Society* 24, ed. by Mercè González, 255-269. GLSA, University of Massachusetts, Amherst.
- Richards, Norvin. 2001. *Movement in Language: Interactions and Architectures*. Oxford University Press.
- Rizzi, Luigi. 2006. On the form of chains: Criterial positions and ECP effects. In *Wh-movement: Moving on*, ed. by Lisa Cheng and Norbert Corver, 97-133. Cambridge, Mass.: MIT Press.
- Rudin, Catherine. 1988. On multiple questions and multiple *wh*-fronting. *Natural Language and Linguistic Theory* 6:445-501.
- Stjepanović, Sandra. 1999. What Do Second Position Cliticization, Scrambling, and Multiple Wh-Fronting Have in Common? Doctoral dissertation, University of Connecticut, Storrs.
- Stjepanović, Sandra. 2010. Left branch extraction in multiple *wh*-questions: a surprise for question interpretation. *Proceedings of Formal Approaches to Slavic Linguistics* 18: 502-517.
- Stjepanović, Sandra. 2011. Differential object marking in Serbo-Croatian: evidence from Left Branch Extraction in negative concord constructions. *Proceedings of Formal Approaches to Slavic Linguistics* 19: 99-115.

- Talić, Aida. 2013. Extraordinary complement extraction: PP-complements and inherently case-marked nominal complements. *Studies in Polish Linguistics* 8: 127-150.
- Wurmbrand, Susi. 2013. Tagalog infinitives: consequences for the theory of phases, voice marking, and extraction. Ms., University of Connecticut, Storrs.
- Zanon, Ksenia. 2015. Russian anaphoric possessive in context. In *Proceedings of Formal Approaches to Slavic Linguistics* 23. This volume.
- Yoo, YongSuk. 2015. New definition of edge and its consequences for the PBC. Presented at Penn Linguistics Colloquium 39.

zeljko.boskovic@uconn.edu

## The Gapping that Could Δ\*

*Barbara Citko*  
University of Washington-Seattle

My main goal in this paper is to contribute to the growing body of research on types of verbal ellipsis in Slavic languages (cf. Goldberg 2005, Gribanova 2013, McShane 2000, 2005, Szczegielniak 2004, 2008, among many others), by examining a construction in Polish that looks like run-of-the-mill gapping but, interestingly, does not exhibit many of the properties typically associated with gapping.

### 1 Puzzle

The empirical puzzle I start with is very simple: what is the relationship of the gapping example in (1) to the types of ellipsis illustrated in (2-3)? In particular, is (1) closer to the verbal ellipsis types in (2a-b) (VP ellipsis and pseudogapping) or to the clausal ellipsis types given in (3a-c): sluicing, stripping (also known as bare argument ellipsis) or sentence fragments.<sup>1,2</sup>

---

\* Many thanks to the audience at the FASL 23 meeting for the feedback and to an anonymous reviewer for useful comments and suggestions.

<sup>1</sup> I will not address in this paper the larger issue of whether all, or any of them, can be analyzed in ways that do not involve ellipsis.

<sup>2</sup> Calling (2a) VPE might be a bit of an oversimplification. Note that it involves modal verbs as licensors; analogous examples with non-modal auxiliaries are out:

(i) \*Jan będzie kandydować na burmistrza a Piotr nie będzie \_\_.

Jan will.be run.INF for mayor and Piotr not will.be

'Jan will run for mayor and Peter won't.'

The term Modal Complement Ellipsis for such cases might be a better suited term (see Aelbrecht 2009 for a concrete proposal).

- (1) Jan kandyduje na burmistrza a Piotr \_\_ na gubernatora.  
 Jan runs for mayor and Piotr for governor  
 ‘Jan is running for mayor and Piotr for governor.’
- (2) a. Jan chciał kandydować na burmistrza a Piotr musiał \_\_.  
 Jan wanted run<sub>INF</sub> for mayor and Piotr had.to  
 ‘Jan wanted to run for mayor and Peter had to.’  
 b. Jan chce kandydować na burmistrza a Piotr musi \_\_  
 Jan wants run<sub>INF</sub> for mayor and Piotr must  
 na wojewodę  
 for governor.  
 ‘Jan wants to run for mayor and Peter must for governor.’
- (3) a. Jan kandyduje na jakieś stanowisko ale nie wiem na jakie.  
 Jan runs for some position but not know for what  
 ‘Jan runs for some position but I don’t know what.’  
 b. Jan kandyduje na burmistrza a nie na gubernatora.  
 Jan runs for mayor and not for governor  
 ‘Jan is running for mayor and not for governor.’  
 c. Na jakie stanowisko Jan kandyduje? Na gubernatora.  
 for what position Jan runs for governor  
 ‘What position is Jan running for? For governor.’

Existing accounts of gapping vary across two parameters, involving the size of coordination ( $vP/VP$  or  $TP$ ) and the size of the elided constituent ( $VP/vP$  or  $TP$ ) (see, among others, Lin 2000, 2002, Johnson 1996/2003, 2009, Coppock 2011 for arguments in favor of the so-called small conjunct approach and Neijt 1979 and Repp 2009 for a large conjunct approach).<sup>3</sup> In this paper, I argue that the Polish ‘gapping that could’ involves *both* clausal coordination and clausal ellipsis. This is not a novel proposal for gapping in general (see, for example, Laka 1990 for Basque, López-Carretero 1995 for Spanish and, more recently, Ai 2014 for Chinese). More specifically, I will argue that (1) involves clausal ellipsis

---

<sup>3</sup> The two are in principle independent. There is only a one way correlation between the size of coordination and the size of ellipsis. While  $vP/VP$  (I use the two terms interchangeably) coordination is incompatible with  $TP$  ellipsis,  $TP$  coordination is perfectly compatible with  $VP/vP$  ellipsis, as proposed by Coppock (2011).

of a complement of a left peripheral head.<sup>4</sup> For the sake of concreteness, I refer to this head as a Polarity head (cf.  $\Sigma$  head of Laka 1990, Kazenin 2006, Progovac 1994, among others). Furthermore, I will show that in certain respects, (1) behaves more like sluicing or sentence fragments (more ‘canonical’ cases of clausal ellipsis) than its English gapping counterpart.

## 2 The ‘Ungapping’ Like Behavior of Polish Gapping

Even though (4a) and (4b) look very much alike, they differ with respect to their compatibility with non-linguistic antecedents, grammaticality with subordinating conjunctions, grammaticality in subordinate contexts, scope of negation (relative to the two conjuncts) and polarity mismatches between the two conjuncts.

- (4) a. John runs for mayor and Peter \_\_ for governor.  
 b. Jan kandyduje na burmistrza a Piotr \_\_ na gubernatora.  
 Jan runs for mayor and Piotr for governor  
 ‘Jan is running for mayor and Piotr for governor.’

The behavior of English gapping with respect to these diagnostics is well known from the relevant literature (Johnson 1993/2004, Coppock 2001, Lin 2000, Repp 2009, Hankamer and Sag 1976, Sag 1976, among others). First, English gapping requires a linguistic antecedent (as shown by the ungrammaticality of (5a)). Second, it is impossible with subordinating conjunctions (see (5b)). Third, it cannot be embedded (see (5c)), and, fourth, it allows negation to have scope over both conjuncts (see (6a)), which, unlike its ungapped variant in (6b), allows the negated modal to scope over the two conjuncts.<sup>5,6</sup>

---

<sup>4</sup> This does not exclude the possibility that gapping can also involve small conjuncts, as I suggested in Citko 2011, 2012. All I am pointing out here is that a large conjunct analysis also needs to be available.

<sup>5</sup> See Repp (2009) for a detailed discussion of the interaction of gapping with negation, including the possibility that the type of negation that allows wide scope in English is quite different from clausal negation.

<sup>6</sup> Further evidence in favor of the availability of wide scope negation in gapping comes from the logical equivalence known as De Morgan’s Law, which takes negation of a

- (5) a. [Hankamer produces an orange, proceeds to peel it, and just as Sag produces an apple, says:]  
 # And Ivan \_\_ an apple. (Hankamer and Sag 1976: 410)
- b. \*Alfonse stole the emeralds *because/if/although/while/unless* /*whenever* Muggsy the pearls. (Hankamer 1979: 18)
- c. \*John runs for mayor and Maria said that Bill \_\_ for governor.
- (6) a. Ward can't eat caviar and Sue \_\_ beans.  
 b. Ward can't eat caviar and Sue can't eat beans. (¬ A) & (¬ B)  
 c. It is not possible (or desirable) for Ward to eat caviar and for Sue (simultaneously) to eat (merely) beans. ¬ (A & B)  
 (Siegel 1984: 524)

And lastly, gapping disallows polarity mismatches of the kind illustrated in (7), in which the first conjunct is positive but the second one is negative (otherwise the negative polarity item would not be licensed).

- (7) \*John invited someone but/and Mary \_\_ **anyone**.

Interestingly, with respect to these properties, Polish gapping is quite different. First, it can be licensed by non-linguistic antecedents:<sup>7</sup>

- (8) a. [To a receptionist in a doctor's office]:  
 My \_\_ do dr Kowalskiego  
 we \_\_ to Dr. Kowalski  
 'We're here to see Dr. Kowalski.'

---

disjunction to be equivalent to conjunction of two negations (see (id)). This is what explains why the gapped example in (ia) is interpreted as (ic) not (ib).

- (i) a. Bob can't play checkers, **or** Mary \_\_ play checkers.  
 b. Bob can't play checkers, **or** Mary can't play checkers. (Lin 2000: 277)  
 c. Bob can't play checkers **and** Mary can't play checkers.  
 d. NEG (A OR B) = (NEG A) AND (NEG B)

<sup>7</sup> These examples do appear to be contextually restricted. A direct translation of the English example in (i) is not felicitous:

- (i) [Jan produces an orange, proceeds to peel it, and Maria produces an apple]  
 # A Maria \_\_ jablko. (Citko 2011: 82)

- b. [Guests in a restaurant to the hostess]:  
 My \_\_ na obiad.  
 we for dinner  
 ‘We are here for dinner.’

The examples in (8a-b) are essentially what McShane (2000, 2005) refers to as Slavic Specific Verbal Ellipsis (SSVE). She describes it as a ‘type of verbal ellipsis that is licensed by a combination of two or more lexical licensers that “go together” both grammatically and semantically.’ McShane’s examples are given in (9a-c).

- (9) a. Ja \_\_ nienaumyślnie.  
 I unintentionally  
 ‘I didn’t mean to do it.’ (McShane 2000: 208)
- b. A Pan \_\_ do kogo?  
 and you to whom  
 ‘And you (formal) are here to see whom?’
- c. Jego świat rozsypuje się, a on \_\_ tylko o  
 his world going.to.pieces REFL and he only about  
 swoich przepisach.  
 self’s regulations  
 ‘His world is going to pieces and he keeps on about his  
 regulations.’ (McShane 2000: 215)

Second, Polish gapping is compatible with subordinating conjunctions:

- (10) Jan kandyduje na burmistrza *bo* Piotr \_\_ na  
 Jan runs for mayor *because* Piotr for  
 gubernatora.  
 governor  
 ‘Jan is running for mayor because Piotr (is running) for governor.’

Third, the gap can sometimes be embedded:<sup>8</sup>

---

<sup>8</sup> Not all embedded gaps are grammatical and the factors that determine which ones can and cannot be embedded remain somewhat unclear to me.

- (11) a. Zosia powiedziała, że Jan nadaje się na burmistrza.  
 Zosia said that Jan fit REFL for mayor  
 ‘Zosia said that Jan will make a good mayor.’  
 b. A Maria twierdzi, że Piotr \_\_ na wojewodę.  
 and Maria claims that Piotr for governor  
 ‘And Maria claims that Piotr would make a good governor.’

Fourth, it only allows narrow scope with respect to negation, as also noted by Repp (2009).<sup>9</sup>

- (12) a. Jan nie może jeść kawioru a Maria \_\_ fasoli.  
 Jan not can eat caviar and Maria beans  
 ‘Jan cannot eat caviar and Maria (cannot eat) beans.’  
 (¬A) & (¬ B)  
 b. Jan nie umie grać w szachy a Maria \_\_ w warcaby.  
 Jan not can play at chess and Maria at checkers  
 ‘Jan cannot play chess and Maria cannot play checkers.’  
 (¬A) & (¬ B)

And fifth, Polish gapping allows polarity mismatches, as pointed out by Przepiórkowski and Kupść (1999), where the antecedent clause can be positive and the gapped clause negative. This property will play a significant part in the analysis I propose in the next section, which capitalizes on the parallelism between gapping and sentence fragments in this respect. Since the negation is elided in the second (gapped) conjunct, its presence has to be diagnosed indirectly. In all the a examples in (13-15), the presence of the so-called n-word (which is a strict negative polarity item) is what provides the requisite diagnostic. What is crucial is that without ellipsis, the overt negation marker *nie* is required, as shown by the corresponding b examples.<sup>10</sup>

<sup>9</sup> De Morgan’s Laws are not testable in Polish as the typical conjunction used in ellipsis contexts is the so-called contrastive conjunction *a* (cf. Citko 2011).

(i) Jan nie może jeść kawioru a/\*i/\*albo Maria \_\_ fasoli.  
 Jan not able eat caviar and/or Maria beans  
 ‘Jan cannot eat caviar and Maria (cannot eat) beans.’

<sup>10</sup> Polish is a strict negative concord language in that its n-words require clausemate negation and are impossible in other (typical) NPI-licensing environments (yes/no

- (13) a. Jan zaprosił kogoś a Piotr **nikogo**.  
 Jan invited someone and Piotr anyone  
 ‘Jan invited someone but Piotr (invited) noone.’  
 b. Piotr **nikogo** **\*(nie)** zaprosił.  
 Piotr anyone not invited  
 ‘Piotr didn’t invite anyone.’
- (14) a. Tomek dał kwiatka Ewie, Janek Marii a  
 Tom gave flower Eve John Mary and  
**Darek** **nikomu**.  
 Darek anyone  
 ‘Tom gave a flower to Eve, John to Mary, and Darek to  
 nobody.’ (Przepiórkowski and Kupść 1999: 217)  
 b. Darek **nikomu** **\*(nie)** dał kwiatka.  
 Darek no one not gave flower  
 ‘Darek didn’t give a flower to anyone.’
- (15) a. Noama Chomskiego, ... , czyta na TT ponad  
 Noam Chomsky read on Twitter over  
 40 tys. wyznawców, a **on** **nikogo**.  
 40 thousand followers and he anyone  
 ‘Over 40 thousand readers follow Chomsky on Twitter, but he  
 (follows) no one.’  
 (<https://twitter.com/Drezyna/status/168626665688281088>)  
 b. On **nikogo** **\*(nie)** czyta.  
 he anyone not read  
 ‘He reads no one.’

---

questions, antecedents of conditionals *etc*) (as shown by, among others, Błaszczak 2001 and Przepiórkowski and Kupść 1997, 1999).

- (i) a. **Nikt** **nigdy** **\*(nie)** **nic** nie zrobił.  
 nobody never NEG nothing not did  
 ‘No one has ever done anything.’ (Błaszczak 2001: 140)  
 b. Czy **nikt** **\*(nie)** wyszedł?  
 Q no one not left  
 ‘Has no one left?’  
 c. Jeżeli **nikt** się **\*(nie)** zjawi, to odwołamy spotkanie.  
 if noone REFL not shows.up then cancel meeting  
 ‘If no one shows up, we will cancel the meeting.’

### 3 Gapping as Multiple Fragments

The parallelism that I build on, also noted by Przepiórkowski and Kupść (1999), is between the polarity reversal that we see in gapping and the polarity reversal that we see in sentence fragments. The negative fragment in (16a) parallels the gapped clause in (16b) in that the negative *nikogo* is possible even though the answer contains no overt negation.

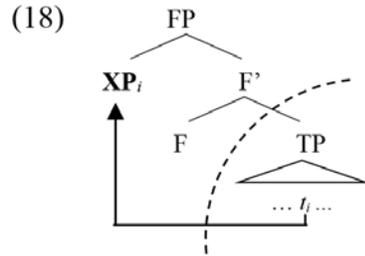
- (16) a. Kogo Piotr zaprosił? **Nikogo.**  
 who Piotr invited anyone  
 ‘Who did Peter invite? Noone.’  
 b. Jan zaprosił kogoś a Piotr **\_\_ nikogo.**  
 Jan invited someone and Piotr anyone  
 ‘Jan invited someone but Piotr (invited) noone.’

Furthermore, both require negation in the absence of ellipsis:

- (17) a. \*Jan zaprosił kogoś a Piotr zaprosił **nikogo.**  
 Jan invited someone and Piotr invited anyone  
 ‘Jan invited someone and Piotr invited noone.’  
 b. \*Piotr zaprosił **nikogo.**  
 Piotr invited anyone  
 ‘Piotr invited no one.’

The fact that that n-words like *nikogo* are allowed as sentence fragments in negative concord languages is known from the literature on negative concord, and has been taken to be indicative of a clausal ellipsis account in which the elided clause contains the licensing negation (see Giannakidou 2000 and Błaszczak 2001, among others, for such an account, and Watanabe 2004 for a different view). Extending this basic insight to gapping is fairly straightforward, as I demonstrate in the rest of this section.

The account I develop builds on two crucial sets of assumptions; one involves the syntax of fragments, and the other one the syntax of negative concord. For fragments, I assume the movement plus deletion account, following Merchant (2004), among many others. According to such an account, the structure of a fragment is the one in (18).



The evidence in favor of such an account comes from a range of connectivity and case matching effects. To illustrate briefly, the ungrammaticality of (19b) can be straightforwardly attributed to a Principle C violation if the fragment moves from the clause-internal position, where it is c-commanded by the coindexed pronoun. Similarly, the bound variable interpretation in (20b) follows naturally if the bound (reflexive) pronoun moves from the clause-internal position, where it is c-commanded by the quantified subject. These examples also show that the case of the fragment has to be the case assigned to it inside the elided clause. And (21b) shows that fragments obey Merchant's Stranding Generalization; the fragment has to include the preposition, not surprising for a language with no preposition stranding.

- (19) a. Co  $on_i$  przeczytał?  
 what he read  
 'What did he read?'  
 b. \*Artykuł  $Jana_i$  [<sub>TP</sub>  ~~$on_i$  przeczytał artykuł  $Jana_i$~~ ]  
 article Jan<sub>GEN</sub> he read article Jan<sub>GEN</sub>  
 'Jan's article.'
- (20) a. Kogo każdy student podziwia?  
 who every student admire  
 'Who does every student admire?'  
 b.  $Swojego_i$  promotora [<sub>TP</sub> ~~każdy<sub>f</sub> student podziwia~~  
 self<sub>GEN</sub> advisor<sub>ACC</sub> every student admires  
 ~~$swojego_f$  promotora ]  
 self<sub>GEN</sub> advisor<sub>ACC</sub>  
 'His advisor.'~~

- (21) a. O kim Jan rozmawiał?  
 about whom Jan talked  
 ‘About whom did Jan talk?’  
 b. *\*(O) Marii<sub>i</sub> [TP Jan rozmawiał o Marii<sub>i</sub>]*  
 about Maria Jan talked about Marii  
 ‘About Maria.’

Case matching also provides an argument for the presence of negation inside the deleted clause when the fragment is an n-word; in (22b) it is genitive, as expected of direct objects in negated clauses.<sup>11</sup>

- (22) a. Ile Jan przeczytał książek?  
 how.many Jan read books  
 ‘How many books did Jan read?’  
 b. Ani jednej. /\*Ani jedną.  
 not.even one<sub>GEN</sub> /\*not.even one<sub>ACC</sub>  
 ‘Not a single one.’

Regarding negative concord, I largely (though not completely) follow Zeijlstra (2004), who argues that n-words in negative concord languages are not negative per se but nevertheless have to be licensed by negation, which he takes to mean they have an uninterpretable negative feature (uNeg feature). I take this feature to be the more general uPol feature (an uninterpretable Polarity feature), which can be valued as either positive or negative. If it is valued as positive, the result is the positive *ktoś* ‘someone’. If it is valued as negative, the result is the negative *nikt*.<sup>12</sup> I

<sup>11</sup> There does seem to be some variation in reported judgments regarding the availability of the accusative variant. My judgments parallel Błaszczak’s.

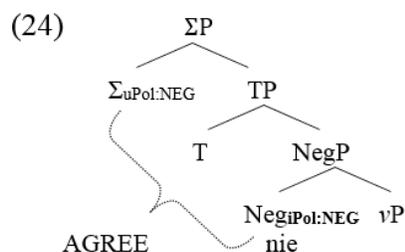
- (i) Przczytałeś jakąś książkę? Żadnej / \*Żadną.  
 read<sub>2SG</sub> some book<sub>ACC</sub> none<sub>GEN</sub> / \*none<sub>ACC</sub>  
 ‘Have you read any book?’ ‘None.’ (Błaszczak 2001: 166)
- (ii) Ile przeczytałeś książek? Żadnej /Żadną.  
 how.many read<sub>2SG</sub> books<sub>GEN</sub> none<sub>GEN</sub> /none<sub>ACC</sub>  
 ‘How many books have you read?’ ‘None.’ (Richter and Sailer 1999: 251)
- (iii) Którą spotkałeś? ?Żadną /?\*Żadnej.  
 which meet<sub>2SG</sub> none<sub>ACC</sub> /none<sub>GEN</sub>  
 ‘Which one have you met?’ ‘None.’ (Przepiórkowski and Kupść 1999: 218)

<sup>12</sup> See Merchant (2013) for a similar proposal regarding the difference between the English *someone* and *anyone* pronominal series.

differ from Zeijlstra though in that I assume that the negative marker proper (situated in Neg) has an interpretable counterpart of this negative feature.<sup>13</sup> This is illustrated schematically in (23).

- (23) Nikt<sub>uPol:\_\_\_</sub> nic<sub>uPol:\_\_\_</sub> nikomu<sub>uPol:\_\_\_</sub> nie<sub>iPol:NEG</sub> dał.  
 anyone anything anyone not gave  
 ‘Noone gave anyone anything.’

Note that the negative pronouns are not c-commanded by negation (at least not in their surface positions), which raises the question of how their uPol feature gets valued. I assume that in addition to the clause internal NegP, there is a high (i.e. left peripheral) Polarity Phrase ( $\Sigma$ P of Laka (1990), Kazenin (2006), Lopéz-Carretero (1995), Progovac (1994), among others), which also has the uPol feature. The high  $\Sigma$  head is distinct from the projection housing clausal negation (clause-internal NegP), which has an interpretable counterpart of the Polarity feature (iPol:pos or iPol:neg). The uPol feature is valued as negative via Agree with the lower Neg:



Negative concord is a result of Multiple Agree (cf. Zeijlstra 2004 and Brown 1999 for a pre-Agree variant). In the derivation of example (23), the high  $\Sigma$  head is first valued as negative via Agree with the lower Neg head (as shown in (25a)). Next, the  $\Sigma$  head undergoes Multiple Agree

<sup>13</sup> In Zeijlstra's system, negative markers in strict negative concord languages are not inherently negative either (and just like n-words, have the uNeg feature). The value for this feature is provided by the interpretable (empty) negative operator, situated in [Spec, NegP].

with all the elements with the unvalued polarity features in its scope, as shown in (25b).<sup>14,15</sup>

- (25) a.  $\Sigma_{uPol:NEG}$  nikt<sub>uPol:</sub>\_\_ nie<sub>iPol:Neg</sub> niC<sub>uPol:</sub>\_\_ nikomu<sub>uPol:</sub>\_\_  
 b.  $\Sigma_{uPol:Neg}$  nikt<sub>uPol:NEG}</sub> nie<sub>iPol:NEG}</sub> niC<sub>uPol:NEG}</sub> nikomu<sub>uPol:NEG}</sub>

Now we can be more concrete regarding the derivation of fragments. In (26a), after the uPol feature on  $\Sigma$  gets valued as negative, the fragment moves to the specifier of the  $\Sigma$  head and the complement of  $\Sigma$  is deleted, as shown in in (26b).<sup>16,17</sup>

- (26) a. Kogo Piotr zaprosił? **Nikogo.**  
 who Piotr invited anyone  
 ‘Who did Peter invite? Noone.’

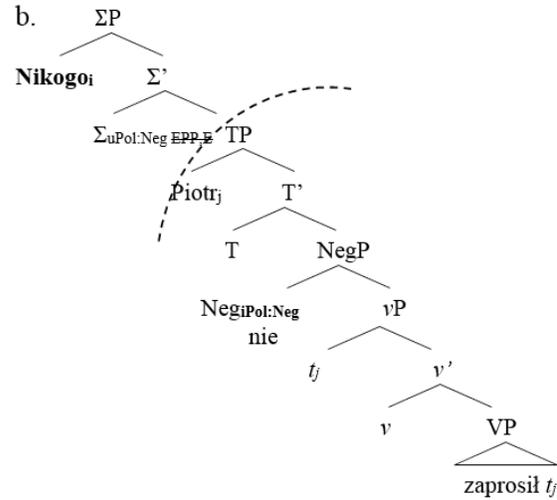
<sup>14</sup> Strictly speaking, this is a case of feature sharing (cf. Pesetsky and Torrego 2006), since both the Probe and the Goal have unvalued features.

<sup>15</sup> As pointed out by an anonymous reviewer, the Agree illustrated in (24) is the standard case of downward probing Agree, whereas the Agree needed in (25) is upward probing (of the kind proposed by Zeijlstra for negative concord, among other things).

<sup>16</sup> For the sake of concreteness, I annotate this movement as triggered by the EPP feature on the  $\Sigma$  head. Its presence is optional as movement is not required in non-elliptical contexts. In this respect, this EPP feature is similar to the EPP feature on other heads (e.g. the  $\nu$  head only has the EPP feature in object shift and/or movement contexts). I thank an anonymous reviewer for a useful discussion of this point.

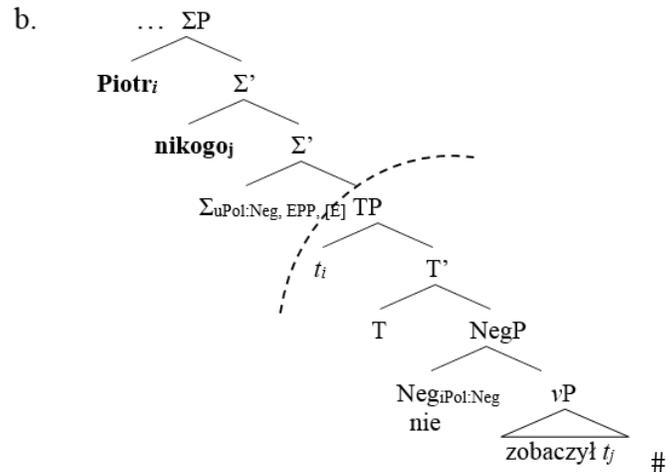
<sup>17</sup> Following Merchant (2001), I assume that the ellipsis itself is triggered by the E (ellipsis) feature on the licenser head,  $\Sigma$  head in this case. For the purpose of this paper, I abstract away from the issue of whether the only heads that can license ellipsis are phase heads (as proposed by Rouveret 2012, Gengel 2005, Citko 2014, among others) or not. The structure in (26b) is compatible with a contextual approach to phasehood of Bošković 2014, on which  $\Sigma P$  would be the phase being the highest projection in the clausal domain and its complement would get deleted. However, the fact that fragments can be embedded (as shown in (i)) is problematic for this view, since in those cases CP would be a phase and the complement of its complement would undergo deletion, something that is explicitly banned in Bošković’s system (see Citko (in preparation) for a possible solution, relying on the phase-theoretical mechanism of Feature Inheritance).

(i) Myśle, że nikogo.  
 think<sub>1SG}</sub> that anyone<sub>CGEN}</sub>  
 ‘I think no one.’



On this analysis, the only difference between fragments and ‘the gapping that could’ is that instead of one, two remnants move to the specifier of the  $\Sigma$  head:

- (27) a. Jan zobaczył kogoś a **Piotr** — **nikogo**.  
 Jan saw someone and Piotr — anyone  
 ‘Jan saw someone and Peter no one.’





Gapping also shows the same kind of case connectivity effects as sentence fragments. (30a-b) illustrate this parallelism with respect to the genitive of negation.

- (30) a. Maria przeczytała pięć książek a Ewa \_\_ ani  
 Maria read five books and Ewa not.even  
 jednej / \*ani jedną.  
 one<sub>GEN</sub> / \*not.even one<sub>ACC</sub>  
 ‘Maria read five articles and Ewa (didn’t read) a single one.’
- b. Ile Jan czytał książek? Ani jednej/\*Ani jedną.  
 how.many Jan read books not.even one<sub>GEN</sub>/\*<sub>ACC</sub>  
 ‘How many books was Jan reading?’ ‘Not a single one.’

The two also behave alike with respect to the Stranding Generalization; pied-piping is obligatory in both:

- (31) a. Jan rozmawiał o Marii a Piotr \*(o) Ewie.  
 Jan talked about Maria and Piotr about Ewa  
 ‘Jan talked about Maria and Piotr about Ewa.’
- b. O kim Jan rozmawiał? \*(O) Ewie.  
 about whom Jan talked about Ewa  
 ‘Who did Jan talk about?’ ‘About Ewa.’

The remaining mysterious properties of the Polish gapping that we started with were its grammaticality in embedded contexts (cf. example (11) above) and its compatibility with subordinating conjunctions (cf. example (10) above, repeated as (32b). Interestingly, here the parallelism breaks down; fragments can be embedded, as shown in (32) (see also Fn 17). However, fragments are *not* possible with subordinating conjunctions, as shown in (33).<sup>18</sup>

- (32) Kto kandyduje na wojewodę? Maria twierdzi, że Piotr.  
 who runs for governor Maria claims that Piotr.  
 ‘Who runs for governor? Maria claims that Piotr (does).’

---

<sup>18</sup> At present I do not have an account of this difference.

- (33) \*Jan kandyduje na burmistrza *bo* Piotr.  
 Jan runs for mayor because Piotr  
 ‘Jan runs for mayor because Piotr does.’ #

And the last area of potential parallelism (or lack thereof) I will explore involves islands. In this respect, gapping and fragments also pattern together. To illustrate briefly, (34a-c) shows that the two behave similarly with respect the Complex NP Island Constraint and (35a-c) does the same for the Coordinate Structure Constraint.<sup>19,20</sup>

- (34) a. \*Ewa poznała aktora, który dostał Oscara a Maria  
 Ewa met actor who got Oscar and Maria  
**Złota Palmę;** ~~poznała [DP-aktora, który dostał *t<sub>i</sub>*]~~  
 Golden Palm  
 ‘Ewa met an actor who got an Oscar and Maria (met an actor who got) Palme D’Or.’  
 b. Czy Maria poznała aktora, który dostał Oscara?  
 Q Maria met actor who got Oscar  
 ‘Has Maria met an actor that got an Oscar?’

<sup>19</sup>English gapping and fragments behave similarly in this respect in that both obey islands (see Neijt 1979 and Merchant 2004, respectively, for data and discussion of the two, respectively).

<sup>20</sup>Interestingly, in this respect both fragments and gapping differ from sluicing, which (as noted by Ross 1969) does not obey islands. If all involve clausal ellipsis, the contrast remains somewhat mysterious. A common solution is to attribute island violations in ellipsis contexts to the fact that violating traces remain after ellipsis (cf. Fox and Lasnik 2003 on islands in VPE). For Merchant, fragments, unlike sluicing, involve an extra movement step, resulting in an offending trace. This, I think, is somewhat speculative. A more promising alternative is to attribute the island violations in gapping and fragments to independent properties of these two constructions. Griffiths (2011) and Griffiths and Liptak (2012), islandhood correlates with contrastiveness and note that whereas contrastive remnants obey islands, non-contrastive ones do not. (34c) above becomes grammatical if the remnant is non-contrastive:

(i) Tak, za ‘Dekalog’.  
 yes for Decalogue  
 Yes, for ‘The Decalogue’.

Gapping remnants are also contrastive, whereas the remnants of sluicing are not.

- c. ?\*Nie, **Złota Palme**<sub>i</sub> ~~Maria poznała~~ [DP aktora,  
no Golden Palm Maria met actor  
~~który dostał t<sub>i</sub> ]~~  
who got  
'No, Palme D'Or.'
- (35) a. \*Jan przeczytał *Struktury Składniowe* i *Program*  
Jan read Structures Syntactic and Program  
Minimalistyczny a *Aspekty*<sub>i</sub> ~~przeczytała~~ [~~&P~~ ~~Struktury~~  
Minimalist and Maria Aspects read Structures  
~~Składniowe i t<sub>i</sub>]~~  
Syntactic  
'Jan read *Syntactic Structures* and *The Minimalist Program*  
and Maria (read Syntactic Structures and) *Aspects*.'
- b. Czy *Aspekty*<sub>i</sub> ~~przeczytała~~ [~~&P~~ ~~Struktury Składniowe~~  
Q Maria read Structures Syntactic and Program  
*Minimalistyczny*?  
Minimalist  
'Has Maria read *Syntactic Structures* and *The Minimalist*  
*Program*?'
- c. \*Nie, *Aspekty*<sub>i</sub> ~~przeczytała~~ [~~&P~~ ~~Struktury Składniowe~~  
not Aspects Maria read Structures Syntactic  
~~t<sub>i</sub> t<sub>i</sub>]~~  
and  
'No, *Aspects*.'

## 5 Conclusion

To conclude briefly, I have pointed out in this paper a number of differences between gapping in Polish and English on the one hand, and a number of parallels between Polish gapping and sentence fragments on the other. Based on these parallels, I have argued in favor of assimilating the two and treating both as involving the same kind of clausal ellipsis licensed by the same left-peripheral head, the high Polarity head.

## References

- Ai, Ruixi Ressay. 2014. Topic-comment structure, focus movement, and gapping formation. *Linguistic Inquiry* 45: 125-45.
- Aelbrecht, Lobke. 2009. *You Have the Right to Remain Silent. The Syntactic Licensing of Ellipsis*. Doctoral dissertation. Catholic University of Brussels.
- Błaszczak, Joanna. 2001. *Investigation into the Interaction between the Indefinites and Negation*. Berlin: Akademie-Verlag.
- Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: on the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45: 27-90.
- Brown, Sue. 1999. *The Syntax of Negation in Russian. A Minimalist Approach*. Stanford: CSLI Publications.
- Citko, Barbara. 2011. *Symmetry in Syntax: Merge, Move and Labels*. Cambridge: Cambridge University Press.
- Citko, Barbara. 2012. A parallel merge solution to the Johnson Merchant paradox. In *Ways of Structure Building*, ed. by V. Valmala Elguea and M. Uribe-Echevarria Goti, 119-142. Oxford: Oxford University Press.
- Citko, Barbara. 2014. *Phase Theory. An Introduction*. Cambridge: Cambridge University Press.
- Citko, Barbara. In preparation. Variation in feature inheritance and clausal ellipsis. Ms. University of Washington.
- Coppock, Elizabeth. 2001. Gapping: in defense of deletion. In *Proceedings of Chicago Linguistics Society* 37, 133–147.
- Fox, Danny Lasnik Howard. 2003. Successive-cyclic movement and island repair: the difference between sluicing and VP-ellipsis. *Linguistic Inquiry* 34: 143-54.
- Gengel, Kirsten. 2009. Phases and ellipsis. *Linguistic Analysis* 35: 21-42.
- Giannakidou, Anastasia. 2000. Negative . . . concord? *Natural Language and Linguistic Theory* 18: 457–523.
- Goldberg, Lotus. 2005. *VP Stranding and VP Ellipsis*. Doctoral dissertation. McGill University.
- Gribanova, Vera. 2013. Verb-stranding verb phrase ellipsis and the structure of the Russian verbal complex. *Natural Language and Linguistic Theory* 31: 91-136.

- Griffiths, James. 2011. An analysis of embedded declarative fragments and its consequences for the theory of sluicing in English. Manuscript, University of Leiden.
- Griffiths, James & Aniko Liptak. 2012. Contrast and island sensitivity in clausal ellipsis. Manuscript, University of Leiden.
- Hankamer, Jorge. 1979. *Deletion in Coordinate Structures*. New York/London: Garland Publishing.
- Hankamer, Jorge and Ivan Sag 1976. Deep and surface anaphora. *Linguistic Inquiry* 7: 391-428.
- Johnson, Kyle. 1996/2003. In Search of the middle field. Manuscript, UMass Amherst.
- Johnson, Kyle. 2009. Gapping is not (VP) ellipsis. *Linguistic Inquiry* 40: 289-328.
- Kazenin, Konstantin. 2006. Polarity in Russian and typology of predicate ellipsis. Manuscript, University of Tuebingen.
- Laka, Itziar. 1990. *Negation in Syntax: On the Nature of Functional Categories and Projections*. Doctoral dissertation, MIT.
- Lin, Vivian 2000. Determiner sharing. In *Proceedings of the 19<sup>th</sup> West Coast Conference on Formal Linguistics*, ed. by R. Billerey and B. D. Lillehaugen, 274-287. Somerville, MA: Cascadilla Proceedings Project.
- Lin, Vivian. 2002. *Coordination and Sharing at the Interfaces*. Doctoral dissertation, MIT.
- Lopéz-Carretero, Luis. 1995. *Polarity and Predicate Anaphora*. Doctoral dissertation, Cornell University.
- McShane, Marjorie. 1999. Verbal ellipsis in Russian, Polish and Czech. *SEEJ* 44: 195-233.
- McShane, Marjorie. 2005. *A Theory of Ellipsis*. Oxford: Oxford University Press.
- Merchant, Jason. 2001. *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford: Oxford University Press.
- Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27: 661-738.
- Merchant, Jason. 2013. Polarity items under ellipsis. 2013. In *Diagnosing Syntax*, ed. by L. Cheng and N. Corver, 441-462. Oxford: Oxford University Press.
- Neijt, Anneke. 1979. *Gapping. A Contribution to Sentence Grammar*. Dordrecht: Foris Publications.

- Pesetsky, David and Esther Torrego. 2006. Probes, goals and syntactic categories. In *Proceedings of the 7th Annual Tokyo Conference on Psycholinguistics*, ed. by Y. Otsu, 25-60. Tokyo: Hituzi Syobo.
- Progovac, Ljiljana. 1994. *Negative and Positive Polarity: A Binding Approach*. Cambridge: Cambridge University Press.
- Przepiórkowski, Adam and Anna Kupść. 1997. Negative concord in Polish. Technical Report 828, Institute of Computer Science, Polish Academy of Sciences.
- Przepiórkowski, Adam and Anna Kupść. 1999. Eventuality negation and negative concord in Polish and Italian. In *Slavic in Head-Driven Phrase Structure Grammar*, ed. by R. D. Borsley and A. Przepiórkowski, 211–246. Stanford, CA: CSLI Publications.
- Repp, Sophie. 2009. *Negation in Gapping*. Oxford: Oxford University Press.
- Richter, Frank and Manfred Sailer. 1999. Negative concord in Polish. In *Slavic in Head-Driven Phrase Structure Grammar*, ed. by R. D. Borsley and A. Przepiórkowski, 247-283. Stanford, CA: CSLI Publications.
- Ross, John R. 1969. Guess who? In *Papers from the 5<sup>th</sup> Regional Meeting of the Chicago Linguistic Society*, ed. by R. Binnick, A. Davidson, G. Green and J. Morgan, 252-286. Chicago: Chicago Linguistics Society.
- Rouveret, Alain. 2012. VP ellipsis, phases and the syntax of morphology. *Natural Language and Linguistic Theory* 30: 897–963.
- Sag, Ivan 1976. *Deletion and Logical Form*. Doctoral Dissertation, MIT.
- Siegel, Muffy 1984. Gapping and interpretation. *Linguistic Inquiry* 15: 523-530.
- Szczegielniak, Adam. 2004. *Relativization and Ellipsis*. Doctoral dissertation, Harvard University.
- Szczegielniak, Adam 2008. Islands and sluicing in Polish. In *Proceedings of the 27<sup>th</sup> West Coast Conference on Formal Linguistics*, ed. by N. Abner and J. Bishop, 404-412. Somerville, MA: Cascadilla Proceedings Project.
- Zeijlstra, Hedde. 2004. *Sentential Negation and Negative Concord*. Doctoral dissertation. LOT Publications.

## **Speculations on DP-structure: Macedonian vs Bulgarian\***

*Steven Franks*  
Indiana University

One's language is part and parcel of one's identity. While true on the individual level, on the group level it is virtually definitive. Language is essential to nation building: this is why it is so important for Croats to speak Croatian (rather than e.g. Serbian), Bosnians to speak Bosnian (rather than e.g. Croatian), and Macedonians to speak Macedonian (rather than e.g. Bulgarian). With respect to this latter, the histories of Mac and Bg as "languages" are very different. While the status of Mac as a language distinct from Bg has been highly politicized and the language was standardized only after the establishment of the Yugoslav People's Republic of Macedonia in 1946, Bg has a tradition as a literary language of well over 1000 years. However, in 1991 Macedonia became autonomous and, by 2015, any debate over Mac as a language is moot.

As formal linguists, we can investigate how distinct the grammars are for the various appellations speakers apply. We might also suspect that—given the need we humans have to belong—sometimes linguistic differences are exaggerated to strengthen national identities. And it is easier to look at an unfamiliar language and interpret its grammar in terms of the grammar of another more familiar one. It is thus tempting to regard Mac as a variant on Bg. While this may give us a leg up, it may also blind us to interesting things going on. In earlier work I have smoothed over descriptive differences between nominal expressions in Mac and Bg in order to paint a more consistent picture. The present paper is an attempt to rectify this lapse.

---

\* Special thanks to Catherine Rudin for collaboration over many years.

## 1 Background

Here I will argue for an important structural difference between these two otherwise very similar languages. In particular, it is proposed that several contrasts between them can be derived from a small distinction in the internal structures of their extended nominal projections. My point of departure is Franks (2009), which following Franco (2000) demonstrates a host of differences between Mac and Bg in the behavior of clause-level pronominal clitics. While in both languages these look like Slavic case markers, there I showed that the clitics are K(ase) heads in Bg, which subsequently move to (clausal) Agr, but in Mac K is generated directly in Agr, as a sort of incipient object agreement. The difference in the status of K in the two languages is schematized in (1) and (2):

- (1) **Bulgarian** (K moves to Agr)
- a. ... Agr ... V [<sub>KP</sub> K [<sub>DP</sub> D [<sub>NP</sub> N ... ]]] →  
... [**K+Agr**] ... V [<sub>KP</sub> K [<sub>DP</sub> D [<sub>NP</sub> N ... ]]]
  - b. [<sub>KP</sub> *ja* 'it' [<sub>DP</sub> D<sub>[+def]</sub> [<sub>NP</sub> *knigata* 'book-the']]]
- (2) **Macedonian** (K generated in Agr)
- a. ... [<sub>Agr</sub> K] ... V [<sub>DP</sub> D [<sub>NP</sub> N ... ]]
  - b. [<sub>DP</sub> D<sub>[-prox, -dist]</sub> [<sub>NP</sub> *knigata* 'book-the/this']]
  - c. [<sub>DP</sub> D<sub>[+prox, -dist]</sub> [<sub>NP</sub> *knigava* 'book-this (here)']]
  - d. [<sub>DP</sub> D<sub>[-prox, +dist]</sub> [<sub>NP</sub> *knigana* 'book-that (there)']]

(2) also shows that Mac has an enriched feature structure for D, contrasting neutral *-t-* with proximal *-v-* and distal *-n*, whereas Bg just has definite *knigata*. Later, I will exploit this contrast.

In analyzing Spanish pronominal clitics, Franco (2000) considered a variety of European languages from the perspective of a list of properties characteristic of object agreement. He showed that, of the 10 properties listed in (3), Mac is second only to Basque, which passes all the tests. My paper reviewed these properties for both Mac and Bg and reached the conclusion that, whenever a diagnostic distinguishes them, the Mac pronominal clitics behave more like agreement markers whereas the Bg ones behave more like pronouns.

- (3) a. Strict adjacency to [Aux/V]  
 b. Syntactic unit with host  
 c. Same specific host  
 d. Fixed order  
 e. Feature erosion in the forms  
 f. Different paradigm selection  
 g. Cooccurrence with accusative arguments  
 h. Unrestricted cooccurrence  
 i. Obligatoriness of cooccurrence  
 j. Cooccurrence with prepositionless NP

The examples in (4) display a phenomenon I call “adverb interpolation” and show how the languages differ with respect to (3a):

- (4) a. Az **sâm** **ti** **ja** veče dala. **Bulgarian**  
 I already you<sub>DAT</sub> it<sub>OBJ</sub> aux<sub>1SG</sub> given  
 ‘I already gave it to you.’  
 b. Tja **ti** **ja** **e** veče dala.  
 she you<sub>DAT</sub> it<sub>OBJ</sub> aux<sub>3SG</sub> already given  
 ‘She has already given it to you.’  
 c. Ti **ne** **si** **mu** **go** vse ošte dala.  
 you NEG aux<sub>2SG</sub> him<sub>DAT</sub> it<sub>OBJ</sub> still given  
 ‘You still have not given it to him.’
- (5) a. \*Jas **sum** vek’e **ti go** dala. **Macedonian**  
 b. \*Jas **sum ti go** vek’e dala.  
 c. \*Ti **ne si mu go** seušte dala.  
 [OR \*Ti ne **si mu go** seušte imala dadeno.]

While acceptable if dispreferred in Bg (4), this possibility is completely absent in Mac (5). Mac speakers consistently reject any attempt to break up the “clitic + verb” sequence.<sup>1</sup> Bg (6) versus Mac (7) show how they differ with respect to (3d):

<sup>1</sup> This is noteworthy because, as an instance of “smoothing over descriptive differences,” in Franks and King (237, fn. 9) we claimed that the phenomenon also obtained in Mac, brushing aside speaker reactions and obscuring the actual telling nature of the contrast.

- (6) a. Ti **mu gi** dade včera. **Bulgarian**  
 you him<sub>DAT</sub> them<sub>OBJ</sub> gave yesterday  
 ‘You gave them to him yesterday.’  
 b. Včera **mu gi** dade.  
 ‘Yesterday you/she/he gave them to him.’  
 c. Dade **mu gi** včera. [NOT \***Mu gi** dade včera.]
- (7) a. Ti **mu gi** dade včera. **Macedonian**  
 b. Včera **mu gi** dade.  
 c. **Mu gi** dade včera. [NOT \*Dade **mu gi** včera.]

Bg (6) shows that the pronominal clitics precede the finite verb, but follow in (6c) so as not to be initial, whereas in Mac (7) they always appear in a fixed position with respect to the verb, hence the contrast in (7c). The contrast between Bg (8) and Mac (9) shows how they differ with respect to (3g–j).

- (8) Kučeto xape kotkata. **Bulgarian**  
 dog<sub>DEF</sub> bites cat<sub>DEF</sub>  
 ‘The dog bites the cat.’
- (9) Kučeto \*(**ja**) kasa mačkata. **Macedonian**  
 dog<sub>DEF</sub> her<sub>OBJ</sub> bites cat<sub>DEF</sub>  
 ‘The dog bites the cat.’

In Mac but not Bg, if the direct object bears an article then the clitic is required.<sup>2</sup> It is, in essence, a marker on the verb indicating that the verb has a definite direct object. In sum, whereas doubling in Bg generally has semantic consequences, in Mac it serves a purely grammatical function. However, since Franks (2009) is five years old, I do not want to belabor arguments there that Mac clitics instantiate object agreement. Instead, I consider an issue I did not address at that time, namely, the fact that the object agreement analysis implies a KP-less DP for Mac, as in (2a).

<sup>2</sup> In Bg the object clitic appears only if the direct object has been topicalized, as in the following permutation of (8), where *ja* reflects the fronting of *kotkata* ‘the cat’:

(i) Kotkata **ja** xape kučeto. ‘The cat, the dog bites it.’

## 2 Some Consequences

It turns out that positing different nominal structures for Mac and Bg has additional consequences. In this section I discuss three of them

*2.1 Bulgarian has Agr in the Nominal Domain, Macedonian does Not*  
 First of all, the structure of Bg KP must be expanded to allow for dative-like clitics, which in Bg are promiscuous in the nominal domain. These elements, which Mac essentially lacks, are henceforth glossed simply as “oblique.” Most of the literature (cf. e.g. Dimitrova-Vulchanova and Giusti 1998 or Tomić 2009) treats them as possessives appearing after the highest head, as in (10):

- (10) a. majka **mi** **Bulgarian (AND Macedonian)**  
           mother I<sub>OBL</sub>  
           ‘my mother’
- b. knjigata **mi** **Bulgarian (ONLY)**  
           book<sub>DEF</sub> I<sub>OBL</sub>  
           ‘my book’
- c. interesnata **mu** kniga  
           interesting<sub>DEF</sub> he<sub>OBL</sub> book  
           ‘his interesting book’
- d. xubavoto **ni** staro selo  
           beautiful<sub>DEF</sub> we<sub>OBL</sub> old village  
           ‘our beautiful old village’
- e. večno mladata **ni** stolica  
           eternally young<sub>DEF</sub> we<sub>OBL</sub> capital  
           ‘our eternally young capital’

With the exception of (10a), the oblique pronominal clitic always appears after the articulated element, which is the highest head in the KP. This is because specifiers impose their own definiteness features on the nominals they scope over, as in English (11a) versus (11b), and—because pronouns themselves are definite, as indicated in (11c)—they necessarily impose a definite interpretation on the possessee:

- (11) a. [[**a** boy's] book] = **a** book of a boy; ≠ **the** book of a boy  
 b. [[**the** boy's] book] = **the** book of the boy; ≠ **a** book of the boy  
 c. **my** book = **the** book that is mine; ≠ **a** book that is mine

In this context, note that the type in (10a), with no article, is the only one possible in contemporary Mac. This construction is highly restricted in Mac, essentially limited to indicating the possessor with certain family relations. It is worth noting that nouns such as *majka* are intrinsically definite and this is presumably precisely what allows them to appear with the possessive clitic, which is in complementary distribution with the article (*\*majkata mi*).<sup>3</sup> I take the significance of these facts, which have not been fully appreciated (and which I glossed over in earlier work), as paramount: they suggest the availability of a functional head in Bg that is absent in Mac. This can be represented as Agr, as in (12):

- (12) a. KP allows **Bulgarian** to have AgrP:  
           [<sub>KP</sub> K [<sub>AgrP</sub> Agr [<sub>DP</sub> D [<sub>NP</sub> N ... ]]]]  
 b. No KP prohibits **Macedonian** from having AgrP:  
           [<sub>DP</sub> D [<sub>NP</sub> N ... ]]

I assume that the extended projection of a nominal can be maximally NP, DP, or KP, but not AgrP. Bg KP thus allows for an AgrP above DP, as in (12a), but in Mac (12b) there is no KP to protect AgrP from being maximal in the extended nominal projection, so there can be no AgrP above DP.<sup>4</sup>

<sup>3</sup> With respect to the *majka mi* construction in Mac, there are descriptive complications stemming from dialect variation. In the standard language, however, not only is no article possible but also no modifiers of N are either (probably for the same reason, since the modifier would be inflected for definiteness). Moreover, outside of southwestern Mac dialects, only singular clitics can appear in this construction. These restrictions lead Tomić (2012: 186) to conclude that “the Macedonian possessive nominal clitics actually seem to be on the way to become permanently attached to their antecedents, [and] ... are inserted in nominal phrases with nouns denoting family relationship, along with those nouns”; for further discussion see Tomić (2012: §3.2).

<sup>4</sup> Extending the analyses in (1b) and (12), we might expect the possibility in Mac of (i):

(i) [<sub>DP</sub> D [<sub>AgrP</sub> [K+Agr] [<sub>NP</sub> N<sub>[+def]</sub> ... ]]]

This could derive the *majki mi* forms in Mac as follows. A K head (recall from the previous note that in the standard language K must be [–plural], i.e., unmarked for number), would be directly inserted into Agr within DP and be realized—presumably by lowering, if that is what happens in general—on the N head to its right. This is however only successful if

Another underappreciated and much more interesting fact about these clitics in Bg is that they can even function as arguments of deverbal nouns, as in (13):<sup>5</sup>

- (13) a. Ivanovoto **mi** izpitvane **Bulgarian (ONLY)**  
 Ivan'S<sub>DEF</sub> I<sub>OBL</sub> examination  
 'Ivan's examination of me'
- b. interesnoto **i** pojavjavane **THEME**  
 interesting<sub>DEF</sub> she<sub>OBL</sub> appearance  
 'her interesting appearance (on the scene)'
- c. ranoto **mu** prepluvane (na kanala) **AGENT**  
 early<sub>DEF</sub> him<sub>OBL</sub> swimming (of channel-the) **OR THEME**  
 'his swimming (of) the channel early'  
 OR 'the channel's early swimming'
- d. pisaneto **mu** **AGENT**  
 writing<sub>DEF</sub> he/it<sub>OBL</sub> **OR THEME**  
 'his writing' OR 'the writing of it' **NOT GOAL**  
 [BUT NOT '\*the writing to him']
- e. predostavjaneto **i/\*mu** (na vâzmožnostta) na Ivo **THEME**  
 offering<sub>DEF</sub> her-it<sub>OBL</sub>Of opportunity<sub>DEF</sub> to Ivo **BUT NOT**  
 'the offering of it (the opportunity) to Ivo' **GOAL**

The important point is that Mac of course allows nothing like (13), since it has no way of introducing these arguments of the deverbal nouns as clitics (they can however be full DPs preceded by the preposition *na* 'of'). To summarize at this juncture, my claim is that Bg has an extended nominal structure as in (14a), whereas Mac has one as in (14b):

- (14) a. **Bulgarian:** [<sub>KP</sub> (*ja*) [<sub>AgrP</sub> (*mi*) [<sub>DP</sub> D [<sub>NP</sub> *knigata*]]]]  
 b. **Macedonian:** [<sub>DP</sub> D [<sub>NP</sub> *knigata/knigava/knigana*]]

---

N is definite and, given the structure in (i), N will only be definite if it is intrinsically so. This limits the construction in Mac to a small set of nouns.

<sup>5</sup> (13a) is from Tasseva-Kurktchieva (2004) and the rest of the examples are drawn from Franks (2001). The semantic restrictions on how the pronominal clitic may be interpreted are treated in that paper. A minimality account is presented: there is only *one* clitic position in each KP (i.e., only one AgrP within KP), and this can only correspond to the highest argument of the deverbal N (Theme becoming highest under passivization).

## 2.2 Colloquial Bulgarian *NEGO SI*

The standard reflexive in Bg is *sebe si*, where obligatory *si* is part of the general paradigm for oblique clitics. *Sebe si* works like a normal long distance (LD) and subject-oriented anaphor, for example, Russian *sebja*, and can be opposed to the Bg colloquial reflexive *nego si* in (15):

- (15) a. Ivan govori na **nego si**. **Bulgarian (ONLY)**  
 ‘Ivan talks to him SI.’<sup>6</sup>  
 b. Marija vinagi misli za **neja si**.  
 ‘Maria always thinks about her SI.’  
 c. Momčeta razčitat na **tjax si**.  
 ‘The boys rely on them SI.’

Note that *nego si* is a non-standard form, with considerable disagreement among speakers about interpretation, usage, and acceptability.<sup>7</sup> For present purposes, however, the relevant observation is that no such form exists in Mac. This makes sense if Bg *nego si* is parasitic on the generalization of the oblique clitic in that language, which in turn implicates more complex nominal structure in Bg than in Mac. Hence there is no source for *si* with pronouns in Mac (although, as discussed at the end of this subsection, Mac does use *si* to augment reflexives).

While these forms seem to parallel English reflexives in their morphology, there are some striking differences. First of all, the reflexive piece is nominal and shows number in English, whereas in Bg it is an oblique reflexive clitic. Second, the head of the phrase is the second *self/selves* piece in English, with the  $\phi$ -feature part a possessive-like modifier, while the head in Bg is the first pronominal *nego/neja/tjax* ‘him/her/them’ piece, with the clitic formally identical to the possessive clitic. Alongside these structural differences come remarkable interpretative differences: whereas English reflexives respect the Binding Theory, their apparent Bg counterparts do not. Schürcks offers some very surprising examples, of which those in (16)–(18) are representative:

<sup>6</sup> Here and below I gloss Bg *si* as SI because English lacks any comparable morpheme; whereas in (15) *self/selves* works, there is no appropriate counterpart in (16)–(18).

<sup>7</sup> All examples and judgments in this section should be taken with the caveat that they are either drawn from the work of Schürcks (2006 and elsewhere) or have been corroborated in personal correspondence with her.

- (16) a. [Ivanovijat bašta] kritikuva **nego si**. **Bulgarian**  
 ‘Ivan’s father criticizes him SI.’ [= Ivan or bašta]  
 b. [Sinūt na [Ivanovija brat]] kritikuva **nego si**.  
 ‘The son of Ivan’s brother criticizes him SI.’  
 [= sinūt OR Ivan OR brat]  
 c. [Da razkažes šegi za **nego si**] ne učudva Javor.  
 ‘That you tell stories about him SI does not surprise Javor.’  
 [= Javor]
- (17) a. Ivan kazva, če doktorât mrazi **nego si**.  
 ‘Ivan says that the doctor hates him SI.’ [= Ivan OR doktorât]  
 b. Marija kaza, če kralicata pokani Ivan i **neja si** na partito.  
 ‘Marija said that the queen invited Ivan and her SI to the party’  
 [= Marija OR kralicata]  
 c. Ivan se poxvali, če statijata e napisana ot Marija i **nego si**.  
 ‘Ivan boasted that the paper was written by Marija and him SI.’  
 [= Ivan]
- (18) Ivan popita bašta si za [Petrovata statija za **nego si**].  
 ‘Ivan asked his father about [Petâr’s article about him SI].’  
 [= Ivan OR Petūr OR bašta]

The examples in (16) show that c-command appears not to be required for antecedency and (17) reveals the absence of domain/locality restrictions. (18) shows both and, given the possibility of coreference with matrix object *bašta* ‘father’, demonstrates that the relationship between *nego si* and its antecedent cannot be one of LD binding.<sup>8</sup>

This construction is treated in more detail in Franks (2013), where it is shown that roughly comparable “logophoric” entities can be found in other Balkan languages, including Albanian, Greek, and Turkish. Consider for example Turkish (19a), from Kornfilt (2001):

<sup>8</sup> Even LD anaphors that have an alternate local life—as evidenced by their ability to be locally bound by a non-subject—never allow for LD binding by anything but a subject. True LD binding thus implies subject-orientation. Some Icelandic speakers accept binding by the local object as in (i), but none ever allow the attempted LD object binding in (ii), from Hyams and Sigurjonsdottir (1990):

(i) Ég sendi Haraldi föt á sig. ‘I sent Harald clothes for himself.’

(ii) \*Ég sagði Jóni að María hefði boðið sér. ‘I told Jon that Maria had invited himself.’

- (19) a. Ali<sub>i</sub> Ahmed -e<sub>j</sub> [ Selim-in<sub>k</sub> **kendi-sin** -i<sub>i/j/k</sub>çok  
 Ali Ahmet -Dat Selim -Gen self -3.sg -Acc very  
 beğen -diğ -i] -i söyle -di.  
 admire -Ger -3.sg -Acc say -Past  
 ‘Ali told Ahmet that Selim admires him(self) very much.’  
 [= Ali OR Ahmet OR Selim]
- b. [*pro*<sub>i</sub> kendi-sin-i<sub>j</sub>]

Like Bg *nego si*, Turkish inflected reflexives such as *kendisi* cannot be analyzed as some kind of LD anaphor, but rather seem to have the same potential latitude of interpretation as *nego si*. Kornfilt’s account says that Turkish inflected reflexives are phrasal and that there is a *pro* subject in SpecAgrP with which *kendi* agrees in person and number, as in (19b). This *pro* refers back to some antecedent, exactly as any other pronoun does, and *pro* itself—rather than the antecedent—locally binds the reflexive. My analysis of Bg *nego si* extends Kornfilt’s account by exploiting the fact that there is enough space within the extended nominal structure to accommodate a pronominal, either *pro* as in Turkish or overt as in Bg. It also develops Schürcks’s insight that these forms reflect the grammaticalization of “point of view,” in that the pronominal refers back to a discourse prominent established topic. Details aside, the crucial point here is simply that no such structure is available in Mac, hence there is nothing comparable to Bg *nego si* in that language.

I have admittedly glossed over one apparent inconsistency: although Mac lacks *nego si*, where *si* occurs with the pronoun, it does have the well-behaved anaphor *sebe(si)* ‘self self<sub>OBL</sub>’, with *si* an optional augment in objective case environments; (20a) is an example from Tomić (2012: 171). While I know of no literature addressing the optionality of *si* after *sebe* in Mac, I suspect that *si* plays some kind of discourse role here (such as marking focus or point of view). Indeed, although this proposal requires more careful investigation, speakers do seem to sense a subtle pragmatic contrast, reporting a preference for *sebesi* when fronted and focused, as in (20b):

- (20) a. Go zede so **sebe(si)**. **Macedonian**  
 ‘(She/He) took it with (her/him) self SI.’
- b. **Sebesi** Ivan se razbira.  
 ‘Ivan understands (him) self SI.’

Be that as it may, I take the possibility of this use of *si* to be a specific consequence of the fact that it cooccurs with *sebe*. That is, I analyze Mac *sebe(si)* as involving optional morphological fission of reflexive features off of *sebe*, perhaps under focus. Crucially, this option is unavailable for *nego*, since it lacks reflexive features, and—since there is no independent Agr in the nominal projection—Mac has no way to derive *nego si*. A final suggestive fact is orthographic: in Mac, unlike Bg, *si* is written together with *sebe*, as if inflectional and intrinsic to its reflexive host.

### 2.3 *The Bulgarian Particle –to*

A third place in which the more complex Bg nominal structure may be seen concerns the obligatory use of the particle *–to* on non-interrogative *wh*-words, as in (21):

- |      |    |   |                  |
|------|----|---|------------------|
| (21) | a. | čovekât koj*(to) govori ...<br>‘the man who-TO is talking ...’  | <b>Bulgarian</b> |
|      | b. | Vzemi kakvo*(to) iskaš.<br>‘Take what(ever) you want.’  |                  |
|      | c. | kâde*(to) i da otideš, ...<br>where-TO and to go <sub>2SG</sub><br>‘wherever you go, ...’   |                  |
|      | d. | Toj e po-goljam, otkolko*(to) ni trjabva.<br>it is more-big than-how-much-TO us is-necessary<br>‘It’s bigger than we need.’       |                  |
|      | e. | brâmbar, goljam kolko*(to) dlanta vi ...<br>beetle big how-much-TO palm <sub>DEF</sub> your ...<br>‘a beetle as big as your palm’ |                  |

(21a) is an ordinary relative clause, (21b) is a free relative, (21c) is a universal concessive conditional (UCC) with *i* a focus marker, (21d) is a comparative, and (21e) is an equative. Note that in (21d) the comparison is clausal whereas in (21e) it is nominal, but *–to* appears regardless.

Because the relative clauses in (21) require *–to*, this element is traditionally taken to be a morphological mark of relative pronouns, as in Hauge (1999). Assuming X-bar theoretic clausal structure it would then seem to make sense to regard *–to* as some kind of C head, as argued for example in Rudin (2009). As such, it would be a dedicated relative complementizer. This is tempting, because, as Rudin (2013) points out,

not only does it seem comparable to English *that* but in closely related Mac the complementizer *što* is used instead with relative pronouns:

- (22) a. čovekot koj(**što**) zboruva ... **Macedonian**  
 ‘the man who is talking ...’  
 b. Jovan, kogo(**što**) vie ne go sakate, ...  
 ‘Jovan, who you do not like, ...’  
 c. mestoto kade(**što**) se sretnavme ...  
 ‘the place where we met ...’

Interestingly, the parallelism between Mac *što* and English *that* is particularly striking in the optionality of the complementizer; see Tomić (2012: ch. 14) for details on Mac relative clauses. This optionality does not extend to Bg *-to* however and, in addition, in the other Bg constructions which require *-to*, such as the UCC in (21c), Mac does not replace it with *što*, but rather uses no special marker. Compare Mac (23) with Bg (21c):

- (23) kade(**\*što**) i da odeš, ... **Macedonian**  
 ‘wherever you go, ...’

In short, Mac has nothing comparable to Bg *-to* in the functions in (21); instead Mac resembles English in using the complementizer *što* ‘that’ when called for—i.e., when introducing a CP—and nothing otherwise. This is not especially surprising if, as I have argued, Mac has evolved into a DP-language like English (and, unlike Bg, both lack KP or AgrP).

The question remains however of what is going on with Bg *-to*. An obvious approach, as for example in Izvorski (2000), is to treat it as some kind of definiteness marker, i.e., a D element, which it clearly looks like. Although Rudin (2009) argues against this, here I attempt to develop and extend this insight. Rudin’s objections basically boil down to the fact that *-to* is an invariant form: it (i) does not change depending on its host; (ii) cannot be replaced by any other determiner; (iii) cannot attach to non-*wh* pronouns. These properties are illustrated in (24)–(26):

- (24) a. koj**to** / koj**ato** / ko**eto** / ko**ito** **Bulgarian**  
 ‘who<sub>MASC-TO</sub>’ / ‘who<sub>FEM-TO</sub>’ / ‘who<sub>NEUT-TO</sub>’ / ‘who<sub>PL-TO</sub>’  
 b. xljab**ât** / kol**ata** / kaf**eto** / knig**ite**  
 ‘bread<sub>MASC.DEF</sub>’ / ‘car<sub>FEM.DEF</sub>’ / ‘coffee<sub>NEUT.DEF</sub>’ / ‘book<sub>PL.DEF</sub>’

- (25) a. **deteto** ‘the child’  
       [AND **tova** dete ‘that child’, **edno** dete ‘a child’]  
       b. **kâdeto** ‘where-TO’  
       [BUT NOT \***tova** kâde ‘\*that where’, \***edno** kâde ‘\*a where’]
- (26) a. koj ‘who’ ⇒ **kojto** ‘who-to’  
       b. toj ‘he’ ≠ \***tojto** or \***tojât**

I propose that Bg invariant *-to* in this function is a kind of bleached D element, an amalgam of D plus a [-agreement] Agr, as in (27):

- (27) [<sub>AgrP</sub> Agr<sub>[-agreement]</sub> [<sub>DP</sub> D + Agr ... ]]

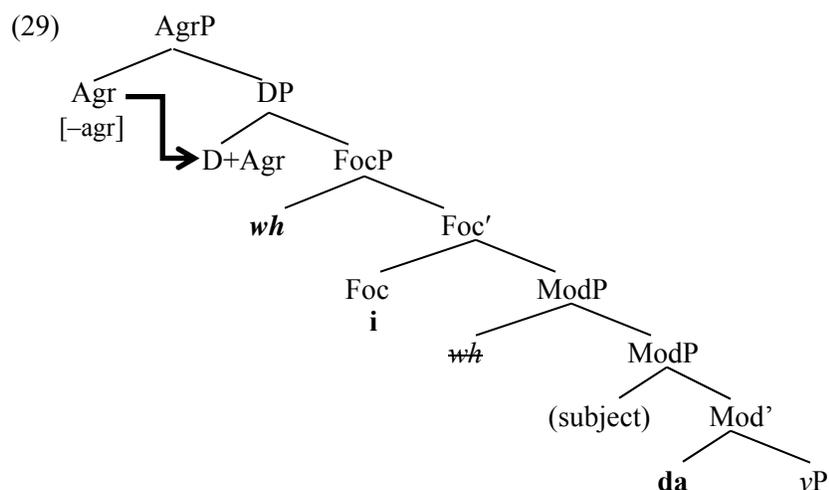
This is not available in Mac, which lacks AgrP in its extended nominal projection and, moreover, has a richer feature structure for D. Recall from (2) that Mac contrasts *-t-* with proximal *-n-* and distal *-v-*:

- (28) a. **kniga** ‘book’ versus **Macedonian**  
       b. **knigata** ‘this book’  
           OR **knigava** ‘this here book’  
           OR **knigana** ‘that there book’

Bg on the other hand just has *knigata* ‘the book’. In sum, whereas Mac has a simpler overall nominal structure than Bg, with just DP, Mac D has a richer feature structure than its Bg counterpart.

With this in mind, let me quickly sketch the analysis of Bg *-to* offered in Franks and Rudin (2014). First of all, we need to assimilate all instances of invariant *-to* to something like (27), with AgrP above DP. This is important because Agr selects a definite DP and the oblique clitic licensed by Agr necessarily follows a head inflected for definiteness. Agr might lower to D, as indicated, or D raise to Agr; however technically implemented, there is at some point a [D+Agr] amalgam. And, more importantly, invariant *-to* has nothing to do with there being a full CP, which is why it extends to constructions beyond the relative clause, as was shown in (21). Recall also in this light the equative (21e), where there isn’t even any apparent clausal structure to accommodate *kolkoto* ‘how-much+to’. Turning now to UCCs, since there is never any agreement (*-to* is invariant) we might imagine that in the UCC *-to* construction Agr too has no agreement features. In Bg DPs there is no proximity marking. And since “neuter” is actually the absence of any specified gender feature, *-to*

has no gender features. That means *-to* is devoid of agreement features, deictic features, and gender features. It is thus so greatly bleached of features that it can serve in UCCs and the other constructions in Bg (21). I thus suggest that Bulgarian UCCs, as in (21c), have a rough structure as in (29), combining (27) with the account of UCCs given in Rudin and Franks (2014):



AgrP thus connects adjunct UCC clauses to the main clause, with invariant *-to* instantiating D+Agr and being realized in the clausal domain just as paradigmatic *-to* is in the nominal domain. Let us explore this parallelism and its consequences.

Bošković (2009) argues for a structure in which AP is contained within NP, either in SpecNP or adjoined to it, as in (30a). A related alternative is for AP to be introduced in SpecFP, i.e., in the specifier of some functional projection above NP, as in (30b):

- (30) a.  $[_{DP} D [_{NP} AP [_{N(P)} N ]]]$       b.  $[_{DP} D [_{FP} AP [F [_{NP} N ]]]]$

Either way, what this means is that definiteness is marked on the head of the specifier of D's complement (in the absence of modification it is marked directly on D's complement, i.e., NP). This could be regarded as

the closest goal, in a probe-goal system.<sup>9</sup> Let us assume (30b) and embed it into the additional structure I have proposed, as in (31):

- (31) **Bulgarian KPs:**  
 [KP K [AgrP Agr [DP D [FP AP [F [NP N ]]]]]]]

This should be compared to the structure in (29) for UCCs:

- (32) **Bulgarian UCCs:**  
 [AgrP Agr [DP D [FocP *wh*-phrase [Foc *i* [ModP *wh*-phrase [Mod *da* ...

In both, *-to* is a realization of D+Agr features on the head of the closest goal, i.e., the modifier in SpecFP (or the NP) in the nominal domain and the *wh*-phrase in SpecCP in the clausal domain. In the former, *-to* is an inflection of the substantive to which it suffixes, hence varies in form, in the latter it does not reflect agreement, hence is invariant. In Mac, on the other hand, no *-to* is available and UCCs are maximally FocPs, not CPs. This is why no *što* was possible in Mac (23).

Returning finally to relative clauses, recall that Mac as a DP-language behaves very much just like English, which is not surprising given that the Bg instantiation is not available. But it also means we should say something about how and why Bg analyzes relative clauses as it does. We have seen that *-to* is clearly not a version of *što* ‘that’. It is perhaps also worth noting that modern standard Bg lacks the lexical item *što* in all its uses, unlike other Slavic languages. The relative complementizer in Bg is *deto*, not *što*, and the *wh* word ‘what’ is *kakvo* rather than *što*, in both interrogatives and free relatives.<sup>10</sup> *Kakvo* can be regarded as the invariant form of *kakâv* ‘what kind of’, which (not coincidentally) is its neuter form, like *-to*. The reasons behind this lexical idiosyncrasy and its relevance to the status of *-to* are unclear; however, it does underscore the existence of significant differences between Bg and Mac in this area of the grammar. In particular, the *-to* forces us to treat Bg relative clauses as nominalizations of some kind, that is, to assimilate relative clause *-to* to UCC *-to*,

<sup>9</sup> It could also be implemented through Spec-head agreement. The point here is simply to highlight the parallelism between nominal and clausal *-to*.

<sup>10</sup> As noted by Zaliznjak (1981: 93) *štoto* is used as a relative pronoun in some Bg dialects and in the literary language exists with the meaning ‘in order that’. Interrogative pronoun *što* ‘what’ is common in dialects.

which involves embedding a FocP into a higher nominal structure. Extending this parallelism, we could analyze Bg relative clauses as CPs embedded in the same higher structure, as in (33):

(33) **Bulgarian RCs:** [<sub>AgrP</sub> Agr [<sub>DP</sub> D [<sub>CP</sub> *wh*-phrase [<sub>C</sub> Ø [<sub>TP</sub> ...

Nominalization of relative clauses is found cross-linguistically, e.g., in various languages of the Americas or Turkish. The Bg strategy does not involve conversion of the verb to a noun or the VP to a participle, but rather nominalizes the entire CP.<sup>11</sup> The “D + Agr” element is, as before, realized as invariant *-to* and once again I suggest the phrase is maximally a functional category above DP, which I have labelled AgrP.<sup>12</sup>

### 3 Postscript: A Puzzle for the Future

This section treats a related set of interesting differences between Mac and Bg in their deverbal nominalizations. I plan in future work to assimilate these differences to the proposed contrast in DP-structure.

#### 3.1 Bare Accommodation Data

At first glance, verbal nouns in Mac seem comparable to their Bg counterparts. Compare Mac (34a) with Bg (34b):

- |      |   |                   |
|------|---|-------------------|
| (34) | a. negovoto baranje na isčeznatite                              | <b>Macedonian</b> |
|      | b. negovoto târsene na izčeznalite                              | <b>Bulgarian</b>  |
|      | ‘his <sub>DEF</sub> searching for lost <sub>DEF</sub> (things)’ |                   |

<sup>11</sup> As is in fact the case in at least some languages of the Americas, for instance, according to C. Rudin (p.c.), Omaha-Ponca.

<sup>12</sup> Zaliznjak (1981) argues that *-to* historically had a relativizing function (*reljativizator*), which it gradually lost. *-to* occurred unambiguously in this function in the earliest Old Russian manuscripts, but could, he writes, have been misinterpreted as the demonstrative particle *-to* by later scribes, just as it would be by the naïve modern Russian. His materials show that *-to* as a *reljativizator* was characteristic of Russian texts from the 11<sup>th</sup>–13<sup>th</sup> centuries, but by the 13<sup>th</sup>–14<sup>th</sup> centuries had already begun to die out noticeably. It is just possible that in Bg, which was developing a DP with a postpositive article, the *reljativizator -to* was preserved through reanalysis as the article *-to*. In Mac, on the other hand, the availability of the alternative *što* strategy for forming relative clauses enabled relativizing *-to* to become extinct, as in Russian (with both full *kotoryj* ‘which’ and reduced participial relative clauses).

These are verbs converted to nouns with the suffix *-nje/-ne*, so that bare complements are not tolerated and the case-marking preposition *na* ‘of, to, for’ is required instead. But that is where the similarity ends. Bg (35) reveals the following pattern: when the deverbal noun is indefinite a bare object is possible. A nominalization such as *očakvane* ‘expecting’ can be transitive, just as *expecting* is in English. However, when the nominalization is definite, as in (35b), then it cannot be transitive and the preposition *na* is required instead:

- (35) a. očakvane(\*to) velika(ta) promjana **Bulgarian**  
       ‘expecting<sub>(\*DEF)</sub> great<sub>(DEF)</sub> change’  
       b. očakvaneto **na** velika(ta) promjana  
       ‘expecting<sub>DEF</sub> of great<sub>(DEF)</sub> change’

Bg forms in *-ne* thus behave like their English counterparts in *-ing*, which presumably involve nominalization of VPs rather than just Vs (cf. *(\*the) expecting the/a great change*) when they take bare objects. Mac (36), on the other hand, displays a completely different pattern:<sup>13</sup>

- (36) a. očekuvanje(to) golema promena **Macedonian**  
       ‘expecting<sub>(DEF)</sub> great change’  
       b. očekuvanje(to) **na** golemata promena  
       ‘expecting<sub>(DEF)</sub> of great<sub>DEF</sub> change’

Here the obligatoriness of *na* depends on the semantics of its *object* rather than of the deverbal noun itself. It is only when the object is definite, as in (36b), that *na* is required. This is true regardless of the form of the deverbal noun.

### 3.2 A Possible Explanation

Building on arguments in this paper that nominal phrases in Mac and Bg are only superficially similar, I claim that *-nje* and *-ne* nominalizations must be very different. The puzzle presented by Mac (36) has two sides: (i) Why is the presence of *na* obligatory when the object of the deverbal noun is definite? And (ii) Why is the transitivity of the deverbal noun

---

<sup>13</sup> Thanks to O. Vangelov (p.c.) for helpful discussion of the workings of Mac. Further aspects of Mac nominalizations are explored in Franks (2014).

independent of the definiteness inflection of that noun? One idea,<sup>14</sup> suggested by Dimitrova-Vulchanova and Mitkovska (2009: 167), is that (36a) and its ilk (also called “direct accommodation”) involve some kind of “object incorporation” (in Baker’s classic sense) and Čašule (1989: 288) points out that “particularly characteristic ... is the possibility of preposing the object which may lexicalize in a compound word,” citing (37a). This is in accord with Koneski’s (1976: 452) intuition that without *na* these sometimes bear a “lexical character,” confirmed by “the fact that the object ‘not rarely’ precedes the verbal noun,” as in his (37b, c).

- (37) a. luđe ubivanje    b. son vidanje    c. oro igranje **Macedonian**  
       ‘people-killing’    ‘dream-seeing’    ‘oro-dancing’

Also noteworthy is that fact that any argument of the verb can incorporate, not just its direct object:<sup>15</sup>

- (38) a. doma odenje                    b. svadba kanenje                    **Macedonian**  
       ‘home-going’                      ‘wedding-inviting’

Interestingly, Bg cannot express compounds in this way. When asked, Bg speakers regarded incorporation as in (37) and (38) “hypothetical” or “totally impossible.” My idea is thus that transitive nominalizations in Bg are unrelated to direct accommodation in Mac, which involves incorporation. This would require an account of why the “V-*nje* N(P)” order still remains the more common one. Note that it is generally single words which precede the deverbal noun (although the PP also can in *po voda odenje* ‘for water going’), in keeping with the traditional analysis of incorporation as head movement. I believe, however, that the restriction is a morphological one, i.e., that there is no syntactic problem with incorporating a phrase, it is just that a phrase cannot be pronounced in a

<sup>14</sup> Another tempting idea is that the correlation between *na* and definiteness is somehow connected to the grammaticalization of clitic doubling in Mac. Any such approach would however need to explain why *na* has this function, and why only in this construction.

<sup>15</sup> The prohibition against definite incorporation may be semantic in that, so far as I can tell, only generic objects participate in direct accommodation. Consider the following examples, from Wikipedija Makedonija:

- (i) a. možnosti za baranje muzika                    b. superiornosta vo davanje usluzi  
       ‘possibilities for searching music’                    ‘the superiority in providing services’

head position. (36a) thus involves incorporation of the object *golema promena* ‘great change’ into the V *očekuva-* ‘expect’, which is then nominalized by adding *-nje*, but *golema promena* cannot be pronounced inside the verb because it is phrasal, hence the lower copy is. This scenario, which is tantamount to incorporation applying in LF rather than overtly, is sketched in (39):

(39) [<sub>N</sub> [<sub>V</sub> [<sub>NP</sub> ~~golema promena~~] [<sub>V</sub> očekuva-]]] –nje] [<sub>DP</sub> D [<sub>NP</sub> **golema promena**]]

Not only does this incorporation analysis address the relationship between the “N V-nje” and “V-nje N(P)” orders, it might also resolve the prohibition against “V-nje DP.” Antilocality—the prohibition against too local movements—is avoided if NP embedded within DP moves rather than DP itself. Such movement would only be possible if D does not impart any definiteness or deictic features to the NP, and the entire DP cannot incorporate because of antilocality.

## References

- Bošković, Željko. 2009. More on the No-DP analysis of article-less languages. *Studia Linguistica* 63: 187–203.
- Čašule, Ilija. 1989. *Sintaksa na makedonskata glagolska imenka*. Skopje: NIO “Studentski zbor”.
- Dimitrova-Vulchanova, Mila and Liljana Mitkovska. 2009. Nominalizations in Bulgarian and Macedonian. In *Investigations in the Bulgarian and Macedonian Nominal Expression*, ed. by Mila Dimitrova-Vulchanova and Olga Mišeska Tomić. Trondheim: Tapir Academic Press. 148–174.
- Franco, Jon. 2000. Agreement as a continuum: the case of Spanish pronominal clitics. In *Clitic Systems in European Languages*, ed. by Frits Beukema and Marcel den Dikken. Amsterdam: Benjamins. 147–89.
- Franks, Steven. 2001. The internal structure of Slavic NPs, with special reference to Bulgarian. In *Generative Linguistics in Poland: Syntax and Morphosyntax*, ed. by Adam Przepiórkowski and Piotr Bański, 53–69.

- Franks, Steven. 2009. Macedonian pronominal clitics as object agreement markers. In *A Linguist's Linguist: Studies in South Slavic Linguistics in honor of E. Wayles Browne*, ed. by Steven Franks, Vrinda Chidambaram and Brian Joseph. Bloomington, IN: Slavica, 189–221.
- Franks, Steven. 2013. Bulgarian NEGO SI is a Balkan anaphor. *Linguistique balkanique* 52:2/3, 3–28.
- Franks, Steven. 2014. Macedonian nominalization puzzles. Talk given at Slavic Linguistics Society annual meeting, Seattle.
- Franks, Steven and Tracy Holloway King. 2000. *A Handbook of Slavic Clitics*. New York: Oxford University Press.
- Franks, Steven and Catherine Rudin. 2014. Invariant –TO in Bulgarian. Unpublished ms, Indiana University and Wayne State College.
- Hauge, Kjetil Rå. 1999. *A Short Grammar of Contemporary Bulgarian*. Bloomington, IN: Slavica.
- Hyams, Nina and Sigríður Sigurjónsdóttir. 1990. The development of “long-distance anaphora”: A cross-linguistic comparison with special reference to Icelandic. *Language Acquisition* 1, 57–93.
- Izvorski, Roumyana. 2000. Free adjunct free relatives. In *Proceedings of WCCFL 19*, ed. by Roger Billerey and Brook Danielle Lillehaugen. Somerville, MA: Cascadilla Press, 232–245.
- Koneski, Blaže. 1976. *Gramatika na makedonskiot literaturni jazik*. Skopje: Kultura.
- Rudin, Catherine. 2009. The Bulgarian relative marker –to. In *A linguist's Linguist: Studies in South Slavic linguistics in honor of E. Wayles Browne*, ed. by Steven Franks, Vrinda Chidambaram and Brian Joseph. Bloomington, IN: Slavica, 403–422.
- Rudin, Catherine. 2013. Sorting out –to and –što: Bulgarian and Macedonian relative markers. Unpublished paper presented at SLS8, Szczecin, Poland.
- Rudin, Catherine and Steven Franks. 2014. Focusing on irrealis concessions. In *Formal Approaches to Slavic Linguistics: The Hamilton Meeting*, ed. by Ivona Kučerová et al. Ann Arbor, MI: Michigan Slavic Publications, 266–286.
- Schürcks, Lilia. 2006. Binding and point of view in Bulgarian. *Zeitschrift für Slawistik* 51: 386–405.
- Tasseva-Kurkchieva, Mila. 2004. Possessives, theta roles, and the internal structure of Bulgarian DPs. In *Proceedings of FASL 12: The Ottawa*

*Meeting*, ed. by Olga Arnaudova, Wayles Browne, Maria Luisa Rivero, and Danijela Stojanović. Ann Arbor, MI: Michigan Slavic Publications, 251–268.

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

Zaliznjak, Andrej Anatolevič. 1981. Protivopostavlenie odnositel'nyx i voprositel'nyx mestoimenij v drevnerusskom. *Balto-slavjanske issledovanija* 1980: 89–107.

franks@indiana.edu

**Nullifying Null Expletives:  
Accounting for EPP in Russian Impersonal and  
Nominative *in situ* Constructions\***

*Allison Germain*  
University of Washington

The aim of this paper is to account for the satisfaction of the EPP in Russian constructions in which it is not obviously satisfied by a Nominative DP, including impersonal constructions and nominative *in situ* constructions. I propose that the subject-like properties and non subject-like properties of preverbal DPs in these constructions can be derived by assuming Split Feature Inheritance. In this analysis, [ $\text{uPhi}$ ] can be inherited by T from the Finite head in the CP layer separately from EPP, allowing for movement of any XP to Spec FinP to satisfy EPP without causing discourse effects. Because T can optionally not bear EPP, no null expletive need be proposed to merge into Spec TP.

**1 The Problem with the Russian EPP**

One of the goals of syntactic research is to explain the relationship between form and meaning. This means that the lack of form in the presence of meaning (e.g. ellipsis, null complementizers) and the lack of meaning in the presence of form (e.g. expletives) are also in the purview

---

\* Many profound thanks to Barbara Citko for all her guidance, Julia Herschensohn and Edith Aldridge for their comments, Marina Oganyan and Olga Trichtchenko for their judgments, members of the UW Syntax Roundtable for comments and support, and the audience and reviewers of FASL 23 for helpful and interesting comments. Any mistakes are my own.

of syntactic investigation. This paper is concerned with another logical possibility: elements with no form and no meaning, referred to in the literature as null expletives.<sup>1</sup> Focusing on Russian, I show that null expletives can be eliminated from the grammar, while maintaining the universality of the EPP. I do so by appealing to the concept of Feature Inheritance, which requires for uninterpretable features to be inherited from phase heads (Chomsky 2008, Richards 2007, among others) and modification of Split Feature Inheritance (Germain 2014).

The empirical focus of this paper is on impersonal constructions in Russian, which have been argued to involve null expletives, as well as personal constructions in which the Nominative (NOM) case bearing DP has been argued to not be in Spec TP. I focus mainly on mono-clausal constructions in this paper, with the exception of Section 4.4. The first two constructions in (1) and (2) are impersonal with 3rd person singular neuter agreement on the verb. For researchers who propose a null expletive to satisfy EPP here, the source of the agreement morphology is the phi-features of the expletive itself as it is third person singular neuter. Thus, if there is no external argument at all or, for some reason, it does not bear NOM, a null expletive is merged in Spec TP.

(1) Adversity Impersonal

Lodku  $\emptyset$  vybrosilo na skaly (volnoj).

boat<sub>ACC</sub> expl<sub>NOM</sub> threw<sub>3.SG.NEUT</sub> on rocks<sub>ACC</sub> wave<sub>INST</sub>

‘The boat was thrown on the rocks (by a wave)’ (Babby 1994: 25)

The psych verb construction below is a personal construction as the verb agrees in person and number with the NOM DP. However, as I will discuss below, this construction shares several similarities to the impersonals above with respect to the subject-likeness of the preverbal DP.

---

<sup>1</sup> A reviewer notes that viewing null expletives as “null subjects of Nature” clears up the issue of semantic content. It is unclear, though, how to constrain the existence or scope of such elements, especially in impersonal clauses with verbs like *kazat’sja* ‘to seem’.

- (2) Nominative *in situ*  
 Saše nraivsja Boris.  
 Sasha<sub>DAT</sub> liked<sub>3.SG.SJA</sub> Boris<sub>NOM</sub>  
 ‘Sasha liked Boris.’

As I discuss in the next section, if it’s not the case that a null expletive is merged into Spec TP to satisfy EPP, then either the EPP does not exist in Russian or it is satisfied by some other means (i.e. some other XP moves to Spec TP). The problem that these constructions present is that, by some diagnostics, the non-NOM preverbal DP has undergone A-movement, and by other diagnostics it has not. In the next section I review the data that has lead some researchers to propose that the EPP on T in Russian is less selective and others to propose that there is a null expletive.

## 2 Previous Accounts

If, as the Extended Projection Principle (EPP) (Chomsky 1986) states, each clause must contain a subject in the specifier of TP, there are three possible analyses for (1) through (3):

1. The EPP is optional in Russian and nothing needs to occupy Spec TP, as proposed by Babby (1989).
2. The EPP in Russian need not be associated with Nominative case assignment and any XP may satisfy it (Babyonyshev 1996, Lavine 2000, Lavine and Friedin 2002, Bailyn 2003, among others).
3. Nothing is special about the EPP(T) in Russian and Spec TP in the data above is occupied by a null element bearing third person singular neuter features (Slioussar 2011, among others).

As my goal is to preserve the universality of the EPP, in the next section I will review only the second two analyses and the data that supports both of them.

### 2.1 Loose EPP

Researchers who have argued that the EPP is “looser” in Russian point to the subject-like properties of the non-NOM preverbal DP. They argue, that like NOM DPs, these DPs have undergone A-movement to Spec TP.

2.1.1 Argument From Reconstruction. One analysis that invokes this is Bailyn's (2003, 2004) Generalized Inversion (GI) proposal. The following are examples of scrambling wherein a non-SVO word order is discourse neutral.

- (3) a. Locative inversion  
 PP-V-S  
 V klasse pojavilsja noven'kij.  
 in class appeared<sub>3.SG.MASC</sub> new(one)<sub>NOM</sub>  
 'A new boy appeared in the class.'
- b. Dative experiencer  
 DAT-V-S  
 Soldatam vidna doroga.  
 soldiers<sub>DAT</sub> visible<sub>FEM.SG</sub> road<sub>NOM.FEM.SG</sub>  
 'The soldiers could see the road.' (Bailyn 2003: 3)

Generalized Inversion derives the sentences in (3) in two steps: the movement of an XP to Spec IP and the movement of v+V to T. Bailyn's evidence that (3) involves A-movement comes from the lack of reconstruction effects. Because a constituent will be interpreted in its new position after A-movement, the binding relationships are affected. In (4), movement of the PP containing the pronoun such that the c-command relationship between it and its antecedent is changed bleeds the Principle B violation in (4).

- (4) a. \*Staršij brati pojavilsja v egoi dome  
 [older brother]<sub>NOM</sub> appeared<sub>3.SG-REFL</sub> in his house  
 'The older brother appeared in his house.'
- b. ? V egoi dome pojavilsja staršij brati  
 in his house appeared<sub>3.SG-REFL</sub> [older brother]<sub>NOM</sub>  
 'The older brother appeared in his house.'

2.1.2 Arguments From WCO. While Bailyn's Generalized Inversion proposal addresses cases in which the EPP(T) is satisfied by some non-NOM XP in the presence of a DP that bears NOM (i.e. agrees with T), Lavine and Freidin (2002) (hereafter L&F) examine how the EPP(T) is satisfied when no DP bears NOM, as in the Adversity Impersonal in (5).

- (5) Soldata ranilo pulej.  
 soldier<sub>ACC</sub> wounded<sub>3.SG.NEUT</sub> bullet<sub>INST</sub>  
 ‘The soldier was wounded by a bullet.’  
 (Lavine & Freidin 2002: 258)

The authors posit a T in Russian that is defective in its phi-features, but maintains an EPP feature. Thus for them, the lack of agreement morphology follows from the fact that T does not enter into an Agree relation with any DP (i.e. it has no need to have its [uPhi] features valued.) Along with Bailyn, they conclude that any XP may move to the Spec TP, and, in the case of (5) above, the internal argument valued Accusative (ACC) moves to this position. The derivation of (5) is shown in Figure 1 below:

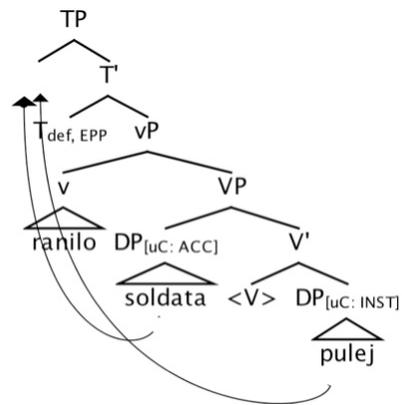


Fig. 1: Movement of a non-NOM DP to Spec TP

The authors appeal to the lack of WCO data in Russian to argue that the ACC DP has actually undergone A-movement to Spec TP. In (6), QR of the Instrumental (INST) DP ‘every gun’ from its base position within the VP causes WCO as *ego* ‘his’ can no longer be indexed to *vladel’ca* ‘owner’. In (7) overt scrambling of the INST DP, no WCO effects arise, and the authors conclude that this must be A-movement as the INST DP does not reconstruct to a lower position.

- (6) a. \*(Rano ili pozdno) [ ego<sub>i</sub> vladel'ca]<sub>k</sub> ub'et *t<sub>k</sub>*  
 soon or late its owner<sub>ACC</sub> will.kill<sub>3.SG</sub>  
 každym pistol'etom<sub>i</sub>.  
 [each gun]<sub>INST</sub>  
 'Sooner or later [its<sub>i</sub> owner]<sub>k</sub> will be killed by every gun<sub>i</sub>.'
- b. LF interpretation after QR (A'-movement)  
 [every gun<sub>i</sub> [[its<sub>i</sub> owner]<sub>k</sub> will kill *t<sub>k</sub>* *t<sub>i</sub>*]] (L&F 2002: 275)
- (7) a. (Rano ili pozdno) [ každym pistol'etom<sub>i</sub>]<sub>k</sub> ub'et ego<sub>i</sub>  
 soon or late [ each gun]<sub>INST</sub> will.kill<sub>AGR</sub> its  
 vladel'ca *t<sub>k</sub>*.  
 owner<sub>ACC</sub>  
 'Sooner or later every gun<sub>i</sub> will kill its<sub>i</sub> owner.'
- b. LF interpretation after scrambling (A-movement) (and QR)  
 [[every gun<sub>i</sub>]<sub>k</sub> [*t<sub>k</sub>* [ will kill its<sub>i</sub> owner *t<sub>i</sub>*]]] (L&F 2002: 275)<sup>2</sup>

2.1.3 Arguments From Focus Projection. Another piece of evidence for A-movement that L&F (2002) cite is the discourse neutral status of (5) above. The whole sentence bears wide focus, despite the fact that the object *soldata* 'soldier' is preverbal. L&F (2002) compare AIs to the data in (8), pointing out that cases of A-scrambling do have an effect on information structure. In (8) movement of an object in a double object construction destroys the discourse neutral reading that the sentence bears in (8), as 'boy' is now narrowly focused.

- (8) a. Odna ženščina podarila mal'čiku jabloko  
 one woman<sub>NOM</sub> gave<sub>3.SG.FEM</sub> boy<sub>DAT</sub> apple<sub>ACC</sub>  
 (i) 'A woman gave a boy an apple'  
 (ii) 'A woman gave the boy an apple'
- b. Odna ženščina podarila jabloko<sub>i</sub> mal'čiku *t<sub>i</sub>*  
 one woman<sub>NOM</sub> gave<sub>3.SG.FEM</sub> apple<sub>ACC</sub> boy<sub>DAT</sub>  
 (i) \*'A woman gave an apple to a boy.'  
 (ii) 'A woman gave the apple to a boy.'  
 (Junghanns and Zybatow 1997: 295)

<sup>2</sup> Nikolaeva (2014) points out that these data are problematic because the control sentence with *ego vladel'ca* 'his owner' preverbal is ungrammatical due to what she terms, "anti-cataphora effects". I leave these data because she ultimately concludes that L&F (2002) are correct in that these constructions do not show WCO. I refer the reader to Nikolaeva (2014: 99-103) for discussion.

Because L&F (2002) assume that both internal arguments in an Adversity Impersonal are equidistant from T, their analysis predicts that either may raise to Spec TP to bear wide focus. As observed by Nikolaeva (2014), this is not born out, and *soldata* ‘soldier’ bears narrow focus in (9) below.

- (9) Pulej ranilo soldata.  
 bullet<sub>INST</sub> wounded<sub>3.SG.NEUT</sub> soldier<sub>ACC</sub>  
 ‘It’s the soldier that the bullet wounded.’  
 \*Wide focus: ‘A soldier was wounded by a bullet.’ (Nikolaeva 2014)

Note that this is also the case with the nominative *in situ* construction. Only when the thematically higher experiencer argument ‘Sasha’ appears pre-verbally in (10) does the sentence bear discourse neutral status.

- (10) a. Saše ne nraivitsja Boris  
 Sasha<sub>DAT</sub> not likes<sub>3.SG</sub> Boris<sub>NOM</sub>  
 ‘Sasha does not like Boris.’  
 Answer to: Do you foresee any problems with our group trip?  
 b. Boris ne nraivitsja Saše  
 Boris<sub>NOM</sub> not likes<sub>3.SG</sub> Sasha<sub>DAT</sub>  
 ‘Sasha does not like Boris.’  
 Not answer to: Do you foresee any problems with our group trip?  
 Answer to: Who likes Boris? (Slioussar 2011: 2059)

## 2.2 Arguments for Null Expletives

While the preverbal non-NOM DP in these data seem to be undergoing A-movement, the following discussion suggests that it is not moving to Spec TP. The most obvious piece of evidence for this is that these preverbal non-NOM DPs are unable to bind anaphors. If binding of anaphors in Russian is from the subject position, following Rappaport (1986), then it must be the case that these preverbal XPs are not in Spec TP. Binding and co-reference data collected as part of a grammaticality judgment survey in Slioussar (2011) leads Slioussar to conclude that it is never the case that a non-NOM XP occupies Spec TP. In (11), the internal argument *soldata* ‘soldier’ cannot bind the reflexive anaphor *svoej* ‘self’s’ and only the pronominal *ego* ‘his’ is grammatical.

- (11) a. Soldata ranilo ego / \*svoej pulej.  
 soldier<sub>ACC</sub> wounded<sub>3.SG.NEUT</sub> his / self's bullet<sub>INST</sub>  
 'A/the soldier was wounded by his bullet.'
- b. Každogo soldata ranilo ego / \*svoej pulej.  
 [every soldier]<sub>ACC</sub> wounded<sub>3.SG.NEUT</sub> his / self's bullet<sub>INST</sub>  
 'Every soldier was wounded by his bullet.'
- (Slioussar 2011: 2060)

Another instance of a non-NOM DP failing to bind an anaphor is the case of preverbal dative (DAT) experiencers of psych verbs (see Section 4.3 below for more discussion). Slioussar (2011) observes that, despite the fact that the sentence has wide focus including the preverbal DAT experiencer, the DAT experiencer cannot bind an anaphor. This is shown in (12) below.

- (12) \*Maše nraivitsja svoja rabota  
 Masha<sub>DAT</sub> likes<sub>3SG</sub> self'<sub>SNOM</sub> work<sub>NOM</sub>  
 'Masha likes her work.'
- (Slioussar 2011: 2065)

Viewing these data, Slioussar proposes that in both impersonal sentences and in the Nominative *in situ* constructions, a null expletive is merged in Spec TP to satisfy the EPP. For impersonal constructions, the lack of verbal agreement morphology is due to the presence of this null element.

An additional argument in favor of null expletives comes from Szucsich's (2007) response to L&F's (2002) proposal, wherein he argues that it cannot be possible for both the internal arguments of AIs to be equidistant from T and at the same time for Adversity Impersonals and their personal counterparts to have the same argument structure, as Babby (1994) proposes. He points out that if the arguments were equidistant, this could allow the wrong DP (the instrument) to be raised to Spec TP in personal sentences. Therefore,  $v$  of AIs simply selects for a "semantically bleached" null D element. As pointed out with (9) above, this doesn't actually occur. If we restrict movement to Spec TP to the most local argument, however, we are still left with the same set of facts to account for. As I will show in the next section, proposing a null expletive does not have to be the answer to this puzzle.

Finally, as Nikolaeva (2014) discusses, the data cited by Bailyn (2003) and L&F (2002) may be best explained by positing A-movement

of the non-NOM XP, but they do not necessarily point to the landing site of this movement being Spec TP.<sup>3</sup> She adopts Kučerova's (2012) proposal that a Topic Phrase can act as an A-position and that movement there is driven by a G-(iveness) operator Agreeing with a [+Given] feature on a DP. This, however, cannot account for neutral word orders as in (5) and (10) above where nothing is marked as given. The next section, I propose that we can account for the lack of information structure effects and the lack of binding via EPP driven movement by tweaking our understanding of how the EPP is related to the movement to Spec TP.

### 3 The Proposal

#### 3.1 *Split Feature Inheritance*

In order to model this movement of a non-NOM DP to a preverbal position that is associated with no discourse effects and no binding capability, I argue that this movement and canonical EPP-driven movement to Spec TP are both a possible outcome of one process: feature inheritance (Chomsky 2008, Richards 2007). Essentially, Split Feature Inheritance, as proposed in Germain (2014), is the idea that this operation applies to [uPhi] and the EPP feature separately. I build on the idea that EPP and [uPhi] are in some sense independent of one another, as is proposed for morphologically ergative languages in which [uPhi] on T is (sometimes) not valued by the DP or clitic which moves to satisfy EPP (see Nevins and Arnand (2003) for such an analysis of Hindi and Rezac (2008) for an analysis of Basque.) The notion that these can act separately from each other dates back also to Ura's (1996) Grammatical Function Splitting proposal. Adger and Svenonius (2011: 25) refer to the EPP as a "second order feature" which in most cases attracts "specific features". This explicit account of the EPP's second order status, inspired by the theory of Autosegmental Phonology (Goldsmith 1976), posits that the EPP in Russian is only loosely attached to [uPhi], much like a floating tone over a tone bearing unit. When C merges with T and feature inheritance of [uPhi] applies, EPP is "stranded" and can land on either of

---

<sup>3</sup> Livitz (2006) and Wood and Livitz (2012) also note this and propose, as I do in this paper, that a subject may move to a position higher than Spec TP.

two possible “tone” bearing units, C or [uPhi] on T. This is pictured in Figure 2 below.

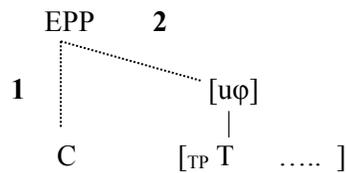


Fig. 2: EPP “floating” after feature inheritance of [uPhi]

Finally, like tonal spreading, the movement of EPP to a different head must be triggered by some other process, namely feature inheritance of [uPhi]. If the EPP, has a unit to bear it, it will not “drift” on its own.

3.1.1 Valuing [uPhi] and Crash. The proposal that [uPhi] and EPP can be inherited or donated individually will only produce the correct results if we reconsider what exactly causes a derivation to crash. Consequently, the second part of my proposal is that a derivation can converge even if [uPhi] on a head remains unvalued.<sup>4</sup> If T inherits only [uPhi] and not EPP, there are two possibilities that I predict. One, [uPhi] will probe and find an active Goal DP, which will provide it with value following Pesetsky and Torrego’s (2007) feature valuation system. Or, two, [uPhi] will probe and find no DP that is active for Agree, will remain unvalued, and will subsequently be pronounced as a default agreement marker, following the notion of elsewhere cases developed in Distributed Morphology (Halle and Marantz 1993).

3.1.2 Optionality in Feature Inheritance and Expanded CP. Optionality is explicitly developed in Ouali’s (2006) feature inheritance system. Here, C does not always pass along its [uPhi] feature and EPP to T, and in fact, may Donate (i.e. in the case of simple declaratives), Keep, or Share its uninterpretable features. In this feature inheritance system (as well as in the ones discussed in Chomsky (2005) and Richards (2007)), T inherits

<sup>4</sup> This idea is explored in Preminger (2011), where the failure of “phi-agreement” only results in the Probe remaining unvalued. Failed “phi-agreement” in this system only leads to ungrammaticality in the case in which an operation which depends on phi-agreement takes place.

all its features from C as a complete bundle. This refers to the [uPhi] and EPP pair traditionally assumed to drive A-movement in a clause. In my proposal I restate Ouali's options in terms of Rizzi's (1997) split CP hypothesis and argue that feature inheritance occurs from Fin to T. As I will show, an expanded CP opens up the possibility of movement to the left periphery that is not discourse driven. Additionally, multiple functional heads allow for the possibility that there are multiple EPP features that could induce A'-movement of more than one DP (e.g. multiple wh-movement). This allows us to avoid an operation such as Ouali's (2006) Share wherein features can be copied and inherited rather than simply inherited (i.e. Donate). Dyakonova (2009) proposes the following structure as an expanded CP field for Russian:

- (13) [<sub>ForceP</sub> Force [<sub>FrameP</sub> Frame [<sub>InterP</sub> Inter [<sub>TopP</sub> Top [<sub>topP</sub> top [<sub>FocP</sub> FOC [<sub>topP</sub> top [<sub>FinP</sub> Fin]]]]]]]]]]  
 (Dyakonova 2009: 145)

As my proposal is concerned with case valuation on DPs, I will ignore FrameP which typically hosts modifying PPs or adverbs, InterP which Dyakonova (2009) takes to be related to interrogatives, and the two optional TopP projections.

- (14) Russian Left Periphery  
 [<sub>ForceP</sub> Force [<sub>TopP</sub> Top [<sub>FocP</sub> Foc [<sub>FinP</sub> Fin]]]]]]

3.1.3 Predicted Outcomes. An expanded CP field, given in (14) above, combined with the proposal that [uPhi] and EPP can be inherited from Fin by T separately via the mechanism outlined above for Split Feature Inheritance gives rise to three possible scenarios, shown in the following three figures. Recall that in this model EPP will not "drift" to T on its own to leave [uPhi] on Fin.

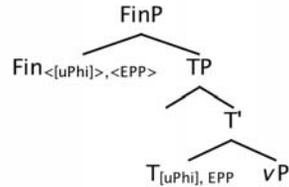


Fig. 3: EPP and [uPhi] on T

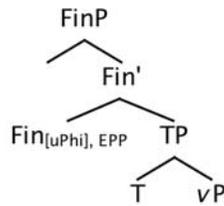


Fig. 4: EPP and [uPhi] on Fin

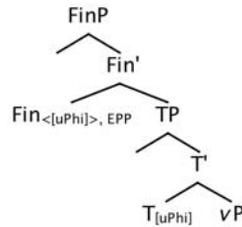


Fig. 5: EPP on Fin and [uPhi] on T

The configuration in Figure 3 leads to canonical SVO personal constructions in which whatever DP that Agrees with T and is valued NOM also moves to Spec TP. The configuration in Figure 4, I propose, is the source of the structural DAT case that is proposed in Landau (2008) and adopted in Livitz (2012). In the next subsection I will show that the configuration in Figure 5 can derive the two constructions discussed above, Adversity Impersonals and Nominative *in situ* constructions.

### 3.2 *The Proposal in Action*

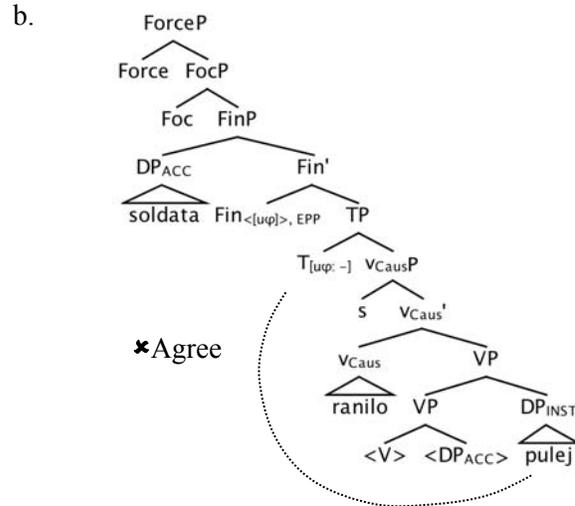
3.2.1 Adversity Impersonals. The configuration in Figure 5, coupled with the proposal that non-Agree or unvalued [uPhi] is the source of the impersonal agreement morphology on the verb, is what allows an impersonal construction in Russian to converge. Essentially, any configuration in which a DP is not available for Agree with [uPhi] on T will result in the derivation of an impersonal sentence. In mono-clausal constructions like the Adversity Impersonals, the only elements bearing phi-features are the direct argument, which bears structural ACC case, and the instrument, which bears inherent INST case.<sup>5</sup> In the following structure in (15), T inherits [uPhi] and probes down the tree. No DP is available as the ACC DP *soldata* ‘soldier’ is already in an Agree relation with  $\nu_{\text{Caus}}$ , and [uPhi] remains unvalued.<sup>6</sup> As the higher internal argument, the ACC DP then moves to Spec FinP to satisfy the EPP on that head. This, then, is left-ward movement without discourse effects, to a position from which the DP cannot bind an anaphor. If we propose that Spec FinP can act as an A-position, then the WCO data is accounted for as well.

- (15) a. Soldata ranilo pulej.  
 soldier<sub>ACC</sub> wounded<sub>AGR</sub> bullet<sub>INST</sub>  
 ‘A soldier was wounded by a bullet.’

---

<sup>5</sup> In the derivation in (16), the instrument is simply adjoined to VP for ease of illustration, but I assume that it is also an argument.

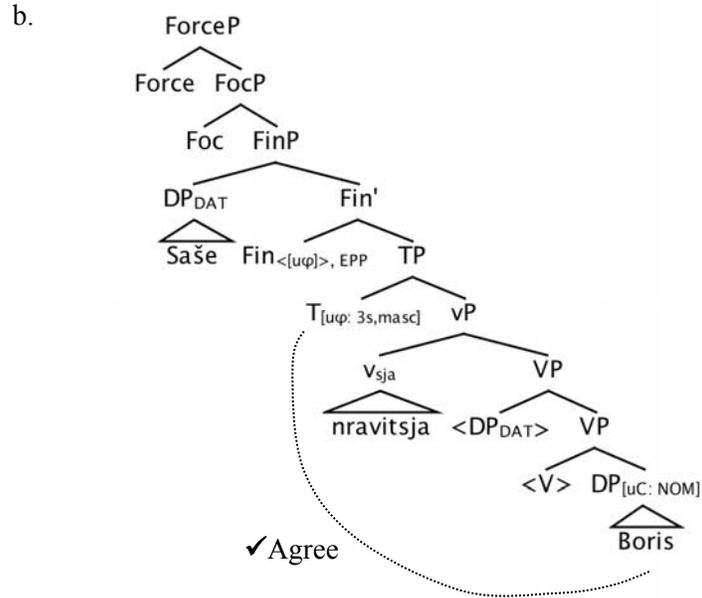
<sup>6</sup> That  $\nu_{\text{Caus}}$  is what heads the  $\nu\text{P}$  of an Adversity Impersonal is due to Markman (2004). In her proposal,  $\nu_{\text{Caus}}$  is “unbundled” from  $\nu_{\text{Voice}}$  in Russian. The  $\nu_{\text{Caus}}$  introduces a causing event *s* in its specifier, which is normally associated with the external agent argument in the specifier of  $\nu_{\text{Voice}}$ . Adversity Impersonals, in her account lack the  $\nu_{\text{Voice}}$  head and therefore an agent argument.



3.2.2. Nominative *In situ*. This distribution of [uPhi] and EPP can also account for structures in which a lower internal argument bears NOM. Under the proposed system, the lack of discourse effects noted by Slioussar (2011) for (10) above can be accounted for if T does not bear an EPP feature. As with the Adversity Impersonal above, the T bears [uPhi] and probes down the structure. In this case, the internal argument has not been valued with structural case<sup>7</sup> and enters into an Agree relation with T. The experiencer *Saše* ‘Sasha’ bears inherent DAT and is unavailable to Agree with T, even though it is more local. As the EPP on Fin is not associated with uninterpretable phi-features, this DAT DP moves to Spec FinP to satisfy it.

- (16) a. Saše nraivitsja Boris.  
 Sasha<sub>DAT</sub> likes<sub>3.SG.SJA</sub> Boris<sub>NOM</sub>  
 ‘Sasha likes Boris.’

<sup>7</sup> Here the v head is the Accusative absorbing *-sja* also found in middle passives (see Fowler (1993) for discussion).



#### 4 Discussion

The proposal in the previous section accounts for the subject-like and non subject-like properties for non-NOM DPs, which are preverbal in Adversity Impersonals and Nominative *in situ* constructions. In this section, I will address a few remaining issues.

##### 4.1 Zero-Place Predicates

Perhaps the most obvious construction for which to propose a null expletive is the construction in which there are no arguments at all. Called “Zero-place predicates” in Babby (2009), these predicates are often weather verbs, as in the example in (17) Zero-place predicate.

- (17) Zero-place predicate  
 Temneet  
 become.dark<sub>3.SG.NEUT</sub>  
 ‘It gets dark.’

Because the EPP in this system is (optionally) not associated with [uPhi] features, any XP will satisfy it. I propose that a  $\nu$ P or VP is what moves to Spec FinP in Zero-place predicates. This is a stipulation and raises the issue of why a  $\nu$ P doesn't satisfy EPP on Fin every time, and I leave that for future research.

#### 4.2 *When High Arguments Stay Low*

According to Nikolaeva (2014), any proposal that relies on EPP-driven movement to derive the kind of movement that fronts internal arguments around verbs runs into problems with the kind of data in (18) and (19) below. In these examples, the thematically lower argument is given and moves to some higher position, while the thematically higher argument stays low. If there were both a Givenness-probe in the left periphery and also EPP on T (or someplace below), the word order would be predicted to be OOV for (18) and OSV for (19).

- (18) Pulej ranilo soldata.  
 bullet<sub>INST</sub> wounded<sub>3.SG.NEUT</sub> soldier<sub>ACC</sub>  
 'Bullet wounded SOLDIER.' (Nikolaeva 2014: 108)
- (19) Šampanskoe budut pit' ljudi  
 champagne<sub>ACC</sub> will drink<sub>3.PL</sub> people<sub>NOM</sub>  
 'PEOPLE will drink champagne.' (Nikolaeva 2014: 108)

Here, I would like to propose that the thematically higher argument stays low because it is in a low Focus position of the kind advocated for in Dyakonova (2009). Dyakonova (2009) argues that the verb moves to Asp of AspectP via head movement and that a low FocP occupies a projection between AspP and  $\nu$ P. If the higher argument *soldata* 'soldier' moves to Spec FocP to check a [+Info] feature, then it is unavailable to move further to satisfy the EPP.<sup>8</sup> If, as I have proposed, the EPP may be split from [uPhi] and remain on Fin, the lower argument *pulej* 'bullet' can move to satisfy it. The derivation for (18) above is given in Figure 6 below.

---

<sup>8</sup> Dyakonova (2009) relies on Rizzi's (2007) notion of criterial freezing to account for fixed positions of nominals in her analysis.

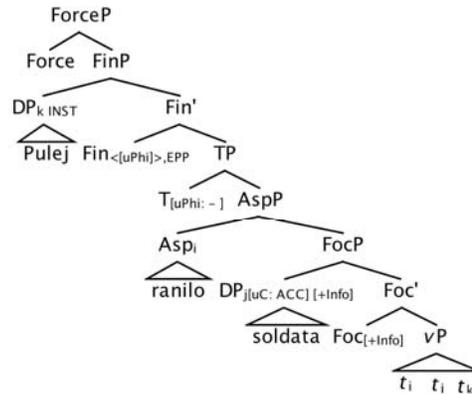


Fig. 6: Higher argument *soldata* remains in low Foc position

#### 4.3 A Note on Binding by Dative Experiencers

For experiencers that bear lexical DAT,<sup>9</sup> there seems to be a split between DAT experiencers that are arguments of psych verbs and DAT experiencers that are part of other predicates. As noted in Franks (1995), these DAT experiencers can bind anaphors.

- (20) Mne žal' sebja  
 me<sub>DAT</sub> sorry self<sub>ACC</sub>  
 'I feel sorry for myself' (Franks 1995: 253)

Preverbal DAT experiencers that are arguments of psych verbs, however, cannot. The following in (21) are more examples of this.

- (21) a. \*Mne nadoedaet svoj učebnik  
 me<sub>DAT</sub> bore<sub>3.SG</sub> self's textbook<sub>NOM</sub>  
 'My textbook bores me.'  
 b. \*Mne dosaždaet svoj brat'  
 me<sub>DAT</sub> vex<sub>3.SG</sub> self's brother<sub>NOM</sub>  
 'My brother vexes me.'

<sup>9</sup> Lexical Dative should be considered in contrast to the functional Dative of Moore and Perlmutter (2000) or structural Dative of Landau (2008) and Livitz (2012).

At this point, I assume that the difference between (20) and (21) stems from the fact that the experiencers in (21) are the arguments of psych verbs and that the experiencer in (20) is the experiencer of a different kind of predicate, one consisting of a copula and a nominal psych predicate.<sup>10</sup> Why the predicate might matter here is a point for future research, but I conclude this discussion section with one more remark on binding.

#### 4.4 *Anaphor Binding and Agree: A Suggestion*

One reviewer points out that the analysis given above raises the question as to why it should be that DPs that Agree with T as opposed to Fin are able to bind anaphors. This is the question of how to derive the subject-oriented nature of binding in Russian (cf. Rappaport 1986). Indeed, c-command is not sufficient to license a DP as a binder, as shown in (22).

- (22) *Militsioner<sub>i</sub> rassprašival arestovannogo<sub>j</sub> o sebe<sub>i/\*j</sub>*  
*policeman<sub>NOM</sub> questioned suspect<sub>ACC</sub> about self<sub>PREP</sub>*  
 ‘The policeman<sub>i</sub> questioned the suspect<sub>j</sub> about himself<sub>i/j</sub>.’  
 (Rappaport 1986: 101)

Following a suggestion made by Edith Aldridge (p.c.), I’d like to put forward the idea that it’s not Agree with T necessarily that allows for binding, but rather Agree with [uPhi]. In Germain (2014), I argue that the following infinitival construction with a DAT subject is bi-clausal, following Fleisher (2006), and that the subject raises to the matrix clause after having been assigned structural DAT as a reflex of Agree with [uPhi] (see Jung 2008 for another raising analysis where a null “prepositional complementizer” assigns lexical DAT case in an ECM scenario).

- (23) *Emu<sub>i</sub> ∅<sub>BYT</sub> [ Fin<sub>[uPhi:3SG]</sub> ne t<sub>i</sub> opulikovat’ svoe<sub>i</sub> stat’ji ]*  
*him<sub>DAT</sub> is not publish<sub>INF</sub> self’s article<sub>GEN</sub>*  
 ‘It’s not in the cards for him to publish his article’  
 (Nikolaeva 2014: 62)<sup>11</sup>

<sup>10</sup> I credit a reviewer with the labeling of *žal’* as a “nominal psych predicate”.

<sup>11</sup> Brackets, Fin, and the addition of the copula *byt’* (null in the present tense) are my own. The intent is to show the structure proposed in Fleisher (2006) alongside the Fin<sub>[uPhi]</sub>.

That Agree with [uPhi] is the critical factor gives rise to the appearance that subjecthood is the criterion for binding in Russian. Of course, the picture is much muddier than this, and I leave this point for future research, referring the reader to Nikolaeva (2014) for a discussion of and proposal regarding argument status and binding facts.

## 5 Conclusions

In this paper, I have attempted to draw a link between two kinds of constructions in Russian that challenge the traditional notion of the EPP as the requirement that a subject occupy Spec TP: Adversity Impersonals and Nominative *in situ* constructions. Researchers who have noticed that preverbal arguments in these constructions seem to undergo A-movement (i.e. they bear wide focus and do not trigger WCO effects), propose that these non-NOM XPs move to Spec TP to satisfy EPP on a “phi-defective” T. These DPs aren’t completely subject-like, however, in that they cannot bind anaphors. This fact has led researchers to propose that the EPP is satisfied by a null expletive in these cases. The proposal offered here accounts for these facts without recourse to a null expletive. Via Split Feature Inheritance, T can optionally inherit only [uPhi] because EPP is only loosely linked to [uPhi] in Russian, leaving EPP on Fin to be satisfied by any XP. Thus, the only difference between impersonal constructions and constructions in which NOM remains *in situ* is the availability for a Goal to Agree with [uPhi] on T. In this way, I hope to preserve the universality of the useful theoretical construct that is the EPP without proposing stop-gap measures like null expletives.

## References

- Adger, David and Peter Svenonius. 2011. Features in minimalist syntax. *The Oxford Handbook of Linguistic Minimalism*, 27-51.
- Babby, Leonard. 1989/2004. Subjectlessness, external subcategorization, and the Projection Principle. *Journal of Slavic Linguistics* 10: 341-88.
- Babby, Leonard. 1994. A Theta-Theoretic analysis of adversity impersonal sentences in Russian. In *FASL 2*, eds. Sergei Avrutin et al., 25-67. Ann Arbor: Michigan Slavic Publications.
- Babby, Leonard. 2009. *The Syntax of Argument Structure*. Cambridge: Cambridge University Press.
- Babyonyshev, Maria. 1996. *Structural Connections in Syntax and Processing: Studies in Russian and Japanese*. Dissertation. Massachusetts Institute of Technology.
- Bailyn, John Frederick. 2003. Does Russian scrambling exist? In: S. Karimi (ed.), *Word Order and Scrambling*. Oxford: Blackwell, 156-176.
- Bailyn, John Frederick. 2004. Generalized inversion. *Natural Language and Linguistic Theory* 22:1-50.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- Chomsky, Noam. 2007. Approaching UG from below. *Interfaces + Recursion = Language*, 1-29.
- Chomsky, Noam. 2008. On phases. In: Freidin, R., Otero, C.P., Zubizarreta, M.-L. (Eds.), *Foundational Issues in Linguistic Theory*. MIT Press, Cambridge, MA, 133-166.
- Dyakonova, Marina. 2009. *A Phase-Based Approach to Russian Free Word Order*. Dissertation. LOT: Utrecht
- Fleisher, Nicolas. 2006. Russian Dative subjects, case and control. Manuscript, University of California, Berkeley.
- Fowler, George. 1993. A syntactic account of derivational -sja in Russian. In *American Contributions to the Eleventh International Congress of Slavists*, eds. Robert Maguire and Alan Timberlake. Columbus: Slavica Publishers, 270-284.
- Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. Oxford University Press, Oxford.

- Germain, Allison. 2014. A reanalysis of Russian impersonal constructions based on Split Feature Inheritance. Unpublished manuscript. University of Washington.
- Goldsmith, John. A. 1976. *Autosegmental Phonology* (Vol. 159). Bloomington: Indiana University Linguistics Club.
- Halle, Morris and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. *The View from Building 20*: 111-176.
- Jung, H. 2008. Licensing of the Dative in infinitival modal sentences: A Case of ECM. *Proceedings from the Annual Meeting of the Chicago Linguistic Society*. Vol. 44. No. 1. Chicago Linguistic Society.
- Junghanns, Uwe and Gerhild Zybatow. 1997. Syntax and information structure of Russian clauses. Wayles Browne et al., eds. *FASL 4: The Cornell Meeting*. Ann Arbor, MI Michigan Slavic Publications, 289–319.
- Kučerová, Ivona. 2012. Grammatical marking of givenness. *Natural Language Semantics* 20:1-30
- Landau, Idan. 2008. Two routes of control: Evidence from case transmission in Russian. *Natural Language & Linguistic Theory* 26: 877–924.
- Lavine, James. 2000. *Topics in the Syntax of Nonagreeing Predicates in Slavic*. Dissertation. Princeton University.
- Lavine, James and Robert Freidin. 2002. The subject of defective T(ense) in Slavic. *Journal of Slavic Linguistics* 10(1-2): 253-289.
- Livitz, Inna. 2006. *What's in a Nominative? Implications of Russian Non-Nominative Subjects for a Crosslinguistic Approach to Subjecthood*. Undergraduate Senior Thesis. Harvard University.
- Livitz, Inna. 2012. Modal possessive constructions: Evidence from Russian. *Lingua* 122. 714-747.
- Livitz, Inna. 2014. *Deriving Silence through Dependent Reference: Focus on Pronouns* Doctoral dissertation, New York University.
- Markman, Vita. 2004. Causatives without causers and Burzio's Generalization. *NELS* 34: 425-440.
- Moore, John and David M. Perlmutter. 2000. What does it take to be a Dative subject? *Natural Language and Linguistic Theory* 18, 373-416.
- Nevins, Andrew and Pranav Anand. 2003. Some agreement matters. *G. Garding and M. Tsujimura, eds*, 370-383.
- Nikolaeva, Liudmila. 2014. *The Secret Life of Pronouns*. Dissertation.

- Massachusetts Institute of Technology.
- Ouali, Hamid. *Unifying Agreement Relations: A Minimalist Analysis of Berber*. ProQuest, 2006.
- Pesetsky, David, and Esther Torrego. 2007. The Syntax of valuation and the interpretability of features. In *Phrasal and Clausal Architecture: Syntactic Derivation and Interpretation*, 262-294.
- Preminger, Omar. 2011. *Agreement as a Fallible Operation*. Dissertation. Massachusetts Institute of Technology.
- Rappaport, Gilbert. 1986. On anaphor binding in Russian. *Natural Language & Linguistic Theory* 4: 97-120
- Rezac, Milan. 2008. Phi-Agree and theta-related Case. *Phi Theory: Phi-features across interfaces and modules*, 83-129.
- Richards, Marc. 2007. On feature inheritance: an argument from the phase Impenetrability Condition. *Linguistic Inquiry* 38: 563-572.
- Rizzi, Luigi. 1997. The fine structure of the Left Periphery. In L. Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer Academic Publishers, 287-296.
- Rizzi, Luigi. 2007. On some properties of criterial freezing. *Studies in Linguistics* 1.145-158.
- Slioussar, Natalia. 2011. Russian and the EPP requirement in the Tense domain. *Lingua* 121 2048–2068.
- Szucsich, Luka. 2007. Nothing wrong with finite T: non-agreeing Accusative impersonal sentences. In *FASL 15*, eds. Richard Compton, et al., 401-419. Ann Arbor: Michigan Slavic Publications.
- Ura, H. 1996. *Multiple Feature-Checking: A Theory of Grammatical Function Splitting*. Dissertation. Massachusetts Institute of Technology.
- Wood, Jim, and Inna Livitz. 2012. What isn't an oblique subject in Icelandic and Russian. *Non-Canonically Case-Marked Subjects within and across Languages and Language Families: Stability, Variation and Change*, University of Iceland.



- (2) a. cardinal *many*:  $|\text{men} \cap \text{in the garden}| \geq n$   $n = \text{large number}$   
 b. proportional *many*:  $|\text{men} \cap \text{in the garden}|/|\text{men}| \geq k$   $k = \text{large \%}$

*Many* can appear in canonical adjectival positions – attributive, see (3a), and predicative, see (3b), as well as in quantificational, see (3c), and differential, see (3d), positions, usually associated with adverbs. In the adjectival position, *many* has only the cardinal reading (Partee 1989: 9).

- (3) a. The many students who attended enjoyed the lecture. Solt (2014)  
 b. John’s friends are many.  
 c. Many students attended the lecture.  
 d. Many more than 100 students attended the lecture.

Taking into account both the syntax and the semantics<sup>1</sup> of *many*, we arrive at the asymmetric picture of *many*, in which the quantificational (or adverbial) *many* is ambiguous between the cardinal and the proportional readings and the adjectival *many* is unambiguously cardinal, see (4). The above discussion of *many* and the picture in (4) can be extended to *few* and with some qualifications to *much* and *little*.

- (4) *many* in English
- |              | Quantificational | Adjectival |
|--------------|------------------|------------|
| proportional | many             |            |
| cardinal     | many             | the many   |

In this paper, I look at the distribution of Russian *many*, which has two morphological forms – an adverbial uninflected *mnog-o* ‘many-adv’ and an agreeing adjectival *mnog-ie* ‘many-agr’, and show that the emerging

---

<sup>1</sup> The semantic discussion here is oversimplified. It does not reflect the difficulty of defining weak vs. strong determiners (see Barwise and Cooper 1981, Keenan 1987). In addition, it takes as a basis the quantificational approach to quantity expressions and does not make justice to other approaches that treat quantity expressions as predicates of sets of individuals or degrees and/or derive proportional reading from cardinal (Hackl 2009, Krasikova 2011, Solt 2014). My primary focus is on the syntax of *many*.

picture of Russian *many* is also asymmetric, but it is strikingly different from English. As shown in (5), Russian adverbial *many* is also ambiguous<sup>2</sup> and, like in English, Russian adjectival *many* is unambiguous. However, unlike English, Russian adjectival *many* is unambiguously proportional (and not cardinal).

(5)	<i>many</i> in Russian		
		Adverbial	Adjectival
	proportional		mnogie
	cardinal Ind	mnogo	
	non-Ind	mnogo	

The generalizations above, if correct, raise a number of interesting questions. First of all, why do the gaps in (4) and (5) exist in the first place? In other words, why does English not have an adjectival proportional *many* and why an adjectival cardinal *many* is missing in Russian? Secondly, if we compare these two languages and put the emphasis on the languages, the question we can ask is why does English lack a proportional adjectival *many* and Russian a cardinal adjectival *many*? I.e. why not the other way round? Although the goal of this paper is to arrive at a better understanding of quantity expressions in Russian and I will not be able to provide full answers to the questions above, the results of the present investigation will allow us to answer these interesting questions partially and show new ways of thinking about the structure and nature of quantification.

The rest of this paper is organized as follows: In section 2, I show that Russian has three types of *many*: two adverbial *manys* and one adjectival *many*. The difference between the two adverbial *manys* is that one of them has a group interpretation and the other the individuated (or referential) interpretation. The adjectival *many*, in addition to being referential, has a familiarity interpretation. Section 3 provides an analysis of *many* in Russian. The main idea of the analysis is that the three *manys* have different structures. The difference between the two adverbial *manys* is in the maximal level of projection being either a QP or a full DP (parallel to the similar distinction in nominals with numerals, e.g. Franks and Pereltsvaig 2004, Pereltsvaig 2006).

---

<sup>2</sup> However, this ambiguity is of a different kind, see below.

## 2 Constructing the Three-Partite Division of Russian *many*

In this section, I argue that Russian has three kinds of quantity expressions, as shown in (5) above. I construct this partitioning in three steps: First, based on the morphology and the syntactic distribution of Russian *many*, I show that it has an adverbial and an adjectival form. Second, I demonstrate that the adjectival *many* and the adverbial *many* in its most salient use pattern similarly to strong and weak determiners respectively. Third, I provide evidence for the ambiguity of the adverbial *many*, thus arriving at the three-partite distinction.

### 2.1 Adverbial vs. Adjectival *many*

In this section, I present evidence that Russian has two forms of *many* – an adverbial and an adjectival. There are two sets of evidence: one comes from the morphology of two forms of *many* in Russian and the other from their syntactic distribution. On the morphological side, *mnog-o* ‘many’ has the derivational adverbial suffix *-o*, whereas *mnog-ie* ‘many’ has an adjectival inflection, see (6-7). This inflection encodes gender, number and case.

- |     |    |                |                               |    |               |                         |
|-----|----|----------------|-------------------------------|----|---------------|-------------------------|
| (6) | a. | <i>mnog-o</i>  | <i>many</i> <sub>ADV</sub>    | b. | <i>tix-o</i>  | quietly                 |
| (7) | a. | <i>mnog-ie</i> | <i>many</i> <sub>PL.NOM</sub> | b. | <i>tix-ie</i> | quiet <sub>PL.NOM</sub> |

On the syntactic side, the adverbial *many*, like other quantifiers, numerals and measure phrases and unlike the adjectival *many* and adjectives, assigns the so-called genitive of quantification to the following noun. However, this evidence is slightly blurred by the fact that there are quantifiers in Russian that do not assign the genitive of quantification and pattern with *mnogie*, see (8a,b). In addition, neither *mnogo* nor *mnogie* can appear in the predicative position, see (8c).<sup>3</sup> Although we do not expect the adverbial *many* to be in the predicative position, the fact that the

---

<sup>3</sup> There are constructions, in which *mnogo* occurs after the verb ‘be’, as in (i). I take such constructions to involve fronting of the complement of *mnogo*, as evidenced by the fact that the noun *druzej* ‘friends-m.gen’ is in genitive (not nominative) and the plural agreement on the predicate is excluded.

(i) *Druz-ej*      *u* *Ivana* *bylo*      *mnogo*.  
 friends<sub>SM.GEN</sub>    at Ivan    be<sub>PAST.NEUT</sub> many  
 ‘Ivan had many friends.’

adjectival *many* is impossible in this position is puzzling. Note also that the verb in (8a) can agree with the subject noun or appear in the neuter form. I will discuss this question in section 2.3.

- (8) a. *Mnogo/neskol'ko/pjat'* studentov *prišl-o/-i* na lekciju.  
 many/some/five students<sub>GEN</sub> came<sub>NEUT/-PL</sub> on lecture  
 'Many/some/five students came to the lecture.'
- b. *Mnogie/vse/dobrosovestnye* studenty *prišli* na lekciju.  
 many<sub>PL.NOM</sub>/all<sub>PL.NOM</sub>/diligent<sub>PL.NOM</sub> students<sub>NOM</sub> came on lecture  
 'Many/all/diligent students came to the lecture.'
- c. *Druzja Ivana byli \*mnogo /\*mnogie / veselye.*  
 friends Ivan were many / many<sub>PL.NOM</sub> / cheerful<sub>PL.NOM</sub>  
 'Ivan's friends were many/cheerful.'

However, there are data that strongly suggest that *mnogo* is adverbial, whereas *mnogie* is adjectival. As illustrated in (9), *mnogo*, like other adverbs and unlike *mnogie* and adjectives, can modify a verb, see (9a), and be a differential modifier in comparatives, see (9b):

- (9) a. *Ivan mnogo / dolgo / \*mnogie spal.*  
 Ivan many / long / many slept  
 'Ivan slept a lot / for a long time.'
- b. *Ivan (na)mnogo /na dva santimetra /\*(na)mnogie vyše Mishy.*  
 Ivan by-many /by two centimetres / by-many taller Misha  
 'Ivan much/two centimetres taller than Misha.'

Finally, there is a piece of morpho-syntactic evidence: the non-inflected *mnogo* is restricted to environments in which only the structural cases are assigned. In other words, it is excluded from the environments where an inherent case is assigned, because inherent cases have to be spelled out and *mnogo* does not inflect, nor does it allow the inherent case to be assigned 'through' to the embedded nominal as in other languages (see Alexiadou et al. 2007). The examples in (10) illustrate this fact.

- (10) a. *Ja rabotaju so mnogimi pianistami.*  
 I work with many<sub>PL.INS</sub> pianists<sub>SPL.INS</sub>  
 'I work with many (of the) pianists.'

- b. \*Ja rabotaju so mnogo pianistov /pianistami.  
 I work with many pianists<sub>PL.GE</sub> /pianists<sub>PL.INS</sub>  
 I work with many pianists.’

These observations allow us to categorize *mnogo* as an adverb and *mnogie* as an adjective. There is no disagreement on this classification in the generative literature (Pereltsvaig 2006, Krasikova 2011).

## 2.2 Weak vs. Strong many

In this section, I review the tests that are presented in the literature (Babko-Malaya 1998, Krasikova 2011) to argue that the adverbial *many* in Russian patterns with weak determiners, whereas the adjectival *many* with strong determiners. I will add some other tests that demonstrate weak vs. strong distinction in Russian. The general conclusion of this section is that ‘many-adv’ in its most salient use is similar to weak determiners, whereas ‘many-agr’ is similar to strong determiners.

I start with the observation that the objects in (11) are interpreted differently. More precisely, the adjectival *mnogix* in (11b) gives rise to the familiarity interpretation of the noun phrase.

- (11) a. Ja znaju mnogo pianistov.  
 I know many pianists<sub>GEN</sub>  
 ‘I know many pianists.’  
 b. Ja znaju mnogix pianistov.  
 I know many<sub>ACC.PL</sub> pianists<sub>ACC</sub>  
 ‘I know many (of the) pianists.’=from a familiar set of pianists

(11b) is infelicitous in the out-of-the-blue context. It is natural in the context in which particular pianists are discussed or during a reception after a piano recital where many pianists are present.<sup>4</sup> The conclusion that

---

<sup>4</sup> A reviewer points out that in case of *mnogie* the notion of familiarity as being present in the previous discourse is too strong and provides the following example in which *mnogie* is felicitous with no previous mentioning of *pianists*:

- (i) - Čem ty uvlekaeš’sja?  
 ‘What are you interested in?’  
 - Ja ljublju klassičeskuju muzyku i, kstati, znaju mnogix pianistov.  
 ‘I like classical music and even know many pianists’

it is, indeed, the agreement that is responsible for the familiarity interpretation is supported by the fact that cases, in which the agreement is obligatory, have only the familiarity interpretation, as we saw in (10a) above. To express the indefinite meaning, an amount noun phrase, in which both the adjective and the noun can inflect, is used, see (12):

- (12) Ja rabotaju s bol'sim količestvom pianistov.  
 I work with large<sub>INS</sub> amount<sub>INS</sub> pianists<sub>GEN</sub>  
 'I work with a large number of pianists.'

In the paragraphs that follow, I show that this difference in interpretation is supported by standard tests for weak vs. strong determiners.

Krasikova (2011) uses three tests to show that Russian *mnogo* is a weak determiner, whereas *mnogie* is a strong determiner. The first test is standard: whether or not a determiner can appear in existential constructions (Milsark 1977). In Russian, *mnogo*, but not *mnogie*, can be used in existential sentences, see (13).

- (13) a. V lesu bylo mnogo razbojnikov. Krasikova (2011: 95)  
 in wood was<sub>NEUT</sub> many outlaws  
 'There were many outlaws.'  
 b. \*V lesu byli mnogie razbojniki.  
 in wood were many outlaws  
 'There were many outlaws in the wood.'

The second test uses individual-level predicates. *Mnogo*, unlike *mnogie*, cannot be a subject of an individual-level NP predicate, see (14) with the judgements provided in Krasikova (2011):

---

I assume that *mnogie* has *weak* familiarity as in Roberts (2000). Weak familiarity does not require that an antecedent discourse referent of NP be introduced via the utterance of NP. Weak familiarity can be satisfied by the perceptual accessibility of the referent to interlocutors, bridging, contextual existence entailments or global familiarity in the general or local culture (Roberts 2000: 14-5). (i) is an illustration of the global familiarity in the general culture. Compare (i) with (ii) in which *classical music* is substituted by *motorcycles*, which makes the use of *mnogie* much less acceptable:

- (ii) - Čem ty uvlekaeš'sja?  
 'What are you interested in?'  
 - Ja ljublju motocikly i, kstati, #znaju mnogix pianistov.  
 'I like motorcycles and even know many pianists'

- (14) a. *Mnogie razbojniki byli vorami.* Krasikova (2011): 95  
 many outlaws were burglars  
 ‘Many outlaws were burglars.’  
 b. \**Mnogo razbojnikov byli vorami.*  
 many outlaws were burglars  
 ‘Many outlaws were burglars.’

It is notoriously difficult to find clear tests for weak vs. strong distinction in Russian. This can be observed on the example of this test. Although at the first glance there is a distinction in acceptability of (14a) vs. (14b), it is not clear that the deviance of (14b) is due to the fact that *mnogo* is weak. First, observe that if we add more context, constructions parallel to (14b) become marginally (and for some speakers fully) acceptable:<sup>5</sup>

- (15) ?*V90-e v SSSR esše mnogo škol’nikov byli pionerami.*  
 in90 in USSR still many pupils were pioneers  
 ‘In 90s in the USSR, many pupils still were ‘pioneers’.’

Second, note that in both (14b) and (15) the copular ‘be’ agrees with the plural subject. We saw in (8a) that *mnogo* ‘many-adv’ can appear with agreeing verbs, as well as with verbs in the neuter form. I will discuss the differences between these two constructions in detail in section 2.3. Anticipating this discussion, I mention the generalization that emerges from the data in section 2.3: *mnogo* with the agreeing verb has an individuated interpretation, whereas *mnogo* with the verb in neuter has a non-individuated (quantificational) interpretation, which is expected to be semantically odd (but not ungrammatical) with individual level-predicates. The quantificational non-individuated reading of *mnogo* is more salient and more easily accessible for Russian speakers. This may explain the fact that adding more context, as in (15), reduces the deviance of the sentence.

The third test builds on the observation made by Herburger (1997) for English and Babko-Malaya (1998) for Russian, that cardinal quantifiers are focus-sensitive, i.e. the truth-conditions of the sentence depend on the placement of focus. The sentence in (16a), in which *flu* is focused, is true

---

<sup>5</sup> The addition of the frame adverbs facilitates the individuated interpretation of *mnogo* making the use of individual-level predicate more acceptable.

if the number of children who have the flu is large compared to the number of children who have other illnesses. The sentence in (16b) with the focus on *children* has different truth-conditions: it is true if the number of children having the flu is large relative to other age groups. *Mnogie*, on the other hand, does not give rise to focus-sensitive readings. The only difference between (17a) and (17b) is contrastive and not truth-conditional.

- (16) a. *Mnogo detej boleet [gripom]<sub>F</sub>*. Krasikova (2011: 106)  
 many children be.ill flu-m  
 ‘Many children have the flu.’  
 b. *Gripom boleet mnogo [detej]<sub>F</sub>*.  
 flu be.ill many children  
 ‘Many children have the flu.’
- (17) a. *Mnogie deti bolejut [gripom]<sub>F</sub> (a ne vetrjankoj)*.  
 many children be.ill flu but not chickenpox  
 ‘Many children have the flu (not chickenpox).’  
 b. *Gripom bolejut mnogie [deti]<sub>F</sub> (a ne vzroslye)*.  
 flu be.ill many children but not adults  
 ‘Many children (not adults) have the flu.’  
 Krasikova (2011: 106)

For Krasikova (2011), this last difference between *mnogo* and *mnogie* is the most important. She uses it to explain the existence of two *manys* in Russian. In her view, *mnogo* and *mnogie* have different lexical entries: The adverbial *mnogo* has a denotation close to a measure phrase (it denotes a predicate of degree sets) and a wired-in property of being focus-sensitive. As a result, *mnogo* can never give rise to a proportional reading. The adjectival *mnogie* has a denotation similar to a gradable non-intersective adjective (it denotes a predicate of individuals), whose comparison class argument is saturated by the noun phrase that it modifies. As a result, *mnogie* has an unambiguous proportional reading. Krasikova (2011) suggests that the ambiguity of Russian *many* is a strategy the language chooses in order to have both proportional and cardinal readings in its inventory: ‘If the adverb-type *many* happens to be focus-sensitive it fails to express a proportional reading and the language has to make available an adjectival *many* which combines with a noun non-intersectively and triggers a proportional reading’ (Krasikova 2011: 111).

In this paper, I will take another direction. Based primarily on syntactic arguments, I will suggest that we do not have to postulate a lexical ambiguity between *mnogo* and *mnogie* in Russian. As Krasikova herself points out, the line of reasoning she proposes does not explain why this particular strategy is chosen, i.e. why with the absence of a proportional *many*, Russian uses an adjectival *many* and not another mechanism that transforms *mnogo* into a proportional quantifier. Nor does it explain the English facts. In addition, as I will show in the next section, Russian adverbial *many* is ambiguous, although this ambiguity does not align neatly with the weak-strong distinction.

However, before concluding this section, I would like to present three other tests that show that *mnogo* patterns with weak determiners, whereas *mnogie* with strong determiners. The first test is mentioned in Barbara Partee's lecture on March 25th, 2004 at the RGGU7 in Moscow. She attributes this observation to one of her students - Yura Lander. The observation is that strong determiners are disallowed as arguments of the verb *imet'sja* 'have-refl'. This test seems to draw a line between the adverbial and adjectival *manys* in Russian, see (18):

- (18) a. V muzee imeetsja mnogo kartin.  
           in museum have many paintings  
           'There are many paintings in the museum.'
- b. \*V muzee imejutsja mnogie kartiny.  
           in museum have many paintings  
           'There are many paintings in the museum.'

The second one is the test for specificity which involves placing the adjective *opredelennye* 'certain' before the quantity expression (see Pereltsvaig 2006: 442). Unfortunately, neither *mnogo* nor *mnogie* can be used in this test because there is a syntactic restriction disallowing adjective preposing, see (19).<sup>6</sup>

---

<sup>6</sup> I thank an anonymous reviewer for pointing this out to me.

- (19) \*Talantlivye /\*opredelennye mnogie/mnogo balerin  
 talented / certain many ballerinas  
 tanceval-i v Mariinskom teatre.  
 danced in Mariinsky theatre  
 ‘Many talented/certain ballerinas danced in the M. Theatre.’

The third test builds on the hypothesis that the complement of a verb with the cumulative prefix *na-* is a QP (and never a DP), see Bailyn 2004, Franks and Pereltsvaig 2004, Pereltsvaig 2006.<sup>7</sup> With respect to this test, *mnogie* patterns with DPs, whereas *mnogo* with QPs, see (20).

- (20) a. Ivan na-kupil [<sub>QP</sub> mnogo /džužinu /stol’ko knig].  
 Ivan <sub>NA</sub>bought many /dozen<sub>ACC</sub> /so-many books<sub>GEN</sub>  
 ‘Ivan bought (so) many/a dozen of books.’  
 b. \*Ivan na-kupil [<sub>DP</sub> mnogix knig].  
 Ivan <sub>NA</sub>bought many<sub>GEN</sub> books<sub>GEN</sub>  
 ‘Ivan bought many (of the) books.’

To summarize, in this section, we saw six tests that suggest that the adverbial *many* in Russian patterns with weak determiners, whereas the adjectival *many* patterns with strong determiners. One test that uses individual-level predicates seems to test not for weak vs. strong distinction, but rather for individuated vs. non-individuated readings. In the next section, I turn to these readings. I show that it is not only the case that ‘many-agr’ has an individuated reading and ‘many-adv’ non-individuated, but also that ‘many-adv’ is ambiguous with respect to these two readings. We will arrive at a three-partite distinction of Russian *many*.

### 2.3 *The Ambiguity of many-adv*

The discussion in this section builds on the work done by Pereltsvaig (2006) (see also Franks and Pereltsvaig 2004). Pereltsvaig (2006) starts with the observation that quantity expressions with numerals can trigger either a plural or neuter agreement, see (21):

---

<sup>7</sup> This hypothesis is based on the observation that such complements bear an obligatory genitive case (the so-called genitive of quantification), presumably assigned by a null Q-head. (26b) shows that when the Q-head is spelled out by a nominal, it appears in the accusative, whereas its complement still bears the genitive of quantification, see (20a).

- (21) V etom fil'me igral-i /-o [pjat' izvestnyx akterov].  
 in this film played<sub>PL/-NEUT</sub> five famous actors  
 'Five famous actors played in this film.'

Based mainly on data from Russian, Pereltsvaig (2006) argues for the existence of Small Nominals, as parallel to Small Clauses. She hypothesizes that Small Nominals (QPs and bare NPs) can appear in the argument position (cf. Longobardi 1994 and subsequent literature). However, unlike full DPs, they have a non-individuated<sup>8</sup> interpretation. As a result, Small Nominals cannot control PROs, be antecedents of reflexives or reciprocals, and trigger the plural agreement on predicates. However, unlike full DPs, they can be arguments of verbs with the cumulative prefix *na-*. To account for the different behavior of Small Nominals, Pereltsvaig (2006) proposes to divide referentiality into two concepts: the criterion of identity (see Baker 2003) and 'the ability to pick out an individual referent' (Pereltsvaig 2006: 483). Syntactically, this division is captured by associating the criterion of identity with a set of unvalued  $\phi$ -features present in Small Nominals and the ability to pick out an individual referent with a set of  $\phi$ -features valued by a D-head, thus, present only in full DPs, as illustrated in (22).<sup>9</sup>

- (22) a. [DP D [QP pjat' [NP banditov]]] full DP  
 b. [QP pjat' [NP banditov]] Small Nominal

In what follows, I employ some of the tests proposed by Pereltsvaig (2006) to argue that Russian *mnogo* is ambiguous and although its ambiguity resembles that of numerals, it is different in that the individuated meaning is provided by the constituent smaller than DP (I will call it Ind(ividuated Reference)P).

I begin with the observation that subjects with quantity expressions in Russian can trigger either plural or neuter agreement on the predicate. Traditionally, subjects with the plural agreement are associated with an

<sup>8</sup> Pereltsvaig (2006) uses 'non-individuated' as equivalent to 'non-referential'.

<sup>9</sup> Pereltsvaig (2006) proposes to distinguish between structural  $\phi$ -features and grammatical gender and number. The latter are specified lexically and responsible for the concord-agreement within NP. The mismatch between these two sets can account for cases like *madame le directeur...* and conjunctions like *the boy and the girl...*

individuated reading and subjects with the neuter with a group reading, see (23-24).<sup>10</sup> The examples in *a* show verb agreement, the examples in *b* show the agreement with short-form adjectives. (25) shows that *mnogie* can appear only with agreeing predicates.<sup>11</sup>

- (23) a. Mnogo ljudej daval-i v sude položitel'nye xarakteristiki..  
 many people gave<sub>PL</sub> in court positive comments  
 'Many people gave positive comments in the court...'  
 b. Mnogo ljudej obespokoen-y tem čto etot Krepkij vyjdet  
 many people concerned<sub>PL</sub> that what this Krepkij will.go  
 na svobodu...  
 to freedom  
 'Many people worried that this K. will be released...'
- (24) a. Mnogo volkov pal-o žertvami volč'ego terrora.  
 many wolves fell<sub>NEUT</sub> victims wolves slaughter  
 'Many wolves became victims of wolves' slaughter.'  
 b. Mnogo trevog svjazan-o s caricej cvetov - rosoj.  
 many worries connected<sub>NEUT</sub> with queen flowers rose  
 'Many worries are associated with the queen of flowers, the rose.'
- (25) Mnogie momenty vyzyval-i / \*-o ulybku.  
 many moments triggered<sub>PL</sub> /-NEUT smile  
 'Many moments caused a smile.'

It is important to note that not all quantity expressions are equally

---

<sup>10</sup> A reviewer correctly points out that the distinction between 'individuated' and 'non-individuated' interpretations are very difficult to pinpoint in terms of truth-conditions. Pereltsvaig (2006: 440) offers the following example to illustrate the difference between these two interpretations:

- (i) Rol' Dz'ejmsa Bonda ispolnjali /#ispolnjalo  
 role James Bond performed<sub>PL</sub> /#performed<sub>NEUT</sub>  
 [pjat' izvestnyx akte'rov].  
 five famous actors  
 'Five famous actors performed the role of James Bond.'

In (i), only the individuated interpretation is possible because the role of James Bond can be performed only individually by each actor on different occasions. See the discussion in Pereltsvaig (2006) and references cited there.

<sup>11</sup> The examples in (23-25) are from the Russian National Corpus online (<http://www.ruscorpora.ru/index.html>) searched on June 14, 2013.

frequent with the plural agreement (many factors, such as animacy of the subject and choice of the predicate influence its availability). *Mnogo* with a predicate in neuter is a more salient option, see Table 1 from Corbett (1981).

quantity expression:	% of plural agreement:
2-4	83%
5-10	50%
<i>neskol'ko</i> 'several'	36%
<i>mnogo</i> 'many', <i>skol'ko</i> 'how much', <i>stol'ko</i> 'so much'	3%

Table 1: Plural agreement with quantified subjects, from Corbett (1981)

Given these agreement facts, it would be natural to expect that *mnogo*, like numerals, can be either DP or Small Nominal, i.e. QP. However, this does not seem to be the case. I will use two tests from Pereltsvaig (2006) to demonstrate this point.

Like Small Nominals with numerals, *mnogo* with non-agreeing predicate cannot be used in constructions that require the subject to be referential, such as control and anaphora-binding. (26a) and (27a) show that the adverbial *mnogo* with non-agreeing predicate cannot license PRO and be an antecedent of a reciprocal, as expected.<sup>12</sup> The unexpected fact is that the adverbial *mnogo* with agreeing predicate is equally infelicitous in these constructions, unlike numerals, see (26b) and (27b).

- (26) a. [\*Pjat' /??mnogo soldat]<sub>i</sub> ležal-o na zemle [PRO<sub>i</sub> ranenye].  
           five /many soldiers lay<sub>NEUT</sub> on ground wounded  
           'Five/many soldiers lay on the ground wounded.'
- b. [Pjat' /??mnogo soldat]<sub>i</sub> ležal-i na zemle [PRO<sub>i</sub> ranenye].  
           five /many soldiers lay<sub>PL</sub> on ground wounded  
           'Five/many soldiers lay on the ground wounded.'

<sup>12</sup> The examples in (26)-(27) are modelled on the examples from Pereltsvaig 2006.

- (27) a. [\*Pjat' /??mnogo soldat] prikryval-o drug druga ot vetra.  
 five /many soldiers shielded<sub>NEUT</sub> each other from wind  
 'Five/many soldiers shielded each other from the wind.'
- b. [Pjat' /??mnogosoldat] prikryval-i drug druga ot vetra.  
 five /many soldiers shielded<sub>PL</sub> each other from wind  
 'Five/many soldiers shielded each other from the wind.'

These data show that nominals with the adverbial *many* in Russian are ambiguous, however their ambiguity is not that between Small Nominals and full DPs, as is the case with nominals with numerals.<sup>13</sup> Similarly to Pereltsvaig (2006), I will propose to capture this ambiguity structurally by adding another layer to the D-domain, see section 3.

To summarize, the goal of this section was to construct a three-partite division of Russian *many*. According to this classification, *mnogie* 'many-agr' is referential and restricted to entities familiar from the discourse, which is equivalent to saying that it is proportional. And there are two adverbial *mnogo* in Russian, one of which has an individuated (referential) interpretation and triggers plural agreement on V. The other one has a non-individuated (non-referential) interpretation and neuter agreement.

### 3 Analysis: Decomposing the D-Head

The core idea of my proposal is that three *manys* in Russian are structurally different. To explain the ambiguity of 'many-adv', I will build on the analyses of nominals with numerals advanced in Franks and Pereltsvaig (2004) and Pereltsvaig (2006). I will treat the non-referential adverbial *many* as a Small Nominal, whose maximal projection is QP, see (28a), and the referential adverbial *many* as projecting into the D-domain. However, I depart from Franks and Pereltsvaig (2004) and Pereltsvaig (2006) in proposing that the functional projection above QP is not a DP, but what I will call an Individuated Reference Phrase - IndP, which is a locus of the

---

<sup>13</sup> A reviewer correctly points out that as the diagnostics that distinguish Small Nominals from full DPs do not work for *mnogos*, the empirical generalization that *mnogo* is ambiguous is weakened. Unfortunately, at this point, I was not able to find further empirical evidence to support the generalization. However, the recent work on semantics of definites that distinguishes between definiteness and determinacy extended to plurals (see esp. Coppock and Beaver 2014) can provide new insights that will help to locate new empirical data and capture the generalization in semantic terms.

individuated referential interpretation and is a head from the DP-zone, see (28b). Another head from the DP-zone – a Fam(iliarity) head – is present in the structure of the adjectival *many*, see (28c). Such a decomposition of DP into (at least) two phrases – IndP and FamP – has a familiar flavor of decomposing a CP into the variety of phrases (Rizzi 1997).

- (28) a. [<sub>QP</sub> mnogo [<sub>NP</sub> studentov]] ‘many-adv’ (non-ind)  
 b. [<sub>IndP</sub> Ind<sub>φ</sub> [<sub>QP</sub> mnogo [<sub>NP</sub> studentov<sub>φ</sub> ]]] ‘many-adv’ (ind)  
 c. [<sub>FamP</sub> mnogie [<sub>IndP</sub> Ind<sub>φ</sub> [<sub>NP</sub> studentov<sub>φ</sub> ]]] ‘many-agr’

In the following sections, I will first look at the difference between the individuated/referential and non-individuated/referential ‘many-adv’ and explain what the IndP is. It will turn out that having an Ind as a separate head in the structure is useful not only to capture Russian data, but some cross-linguistic data as well. In section 3.2, I will discuss the adjectival *many*, which I propose should be treated as a ‘definite’ adjective in Slavic.

### 3.1 *The Analysis of ‘many-ADV’: IndP*

The idea that the DP-zone is richer than just a projection of a single head is not entirely novel (see Alexiadou et al. 2007 and Kyriakaki 2011 from whom I borrow the label FamP). It is particularly useful when in work on languages without articles and the contribution that is usually assigned to articles should be divided between some syntactic and non-syntactic mechanisms.

Consider the proposal outlined in Pereltsvaig (2006) for Small Nominals: they can be arguments, they have a criterion of identity represented as a set of unvalued  $\phi$ -features. However, they are ‘reduced’ arguments: they cannot be part of constructions that require individual reference, e.g. control PRO, bind anaphora, and have a specific or a partitive interpretation, as we saw above. On the other hand, if DP is projected the Small Nominal becomes fully referential.

Imagine now a slight variation on this picture. Small Nominals are QPs, which have a group interpretation. They are transformed into arguments via some last resort semantic operation, which explains their ‘reduced’ properties. If QP needs to receive an individuated interpretation, an Individuated-Reference head (Ind-head) is merged. This head is from a cluster of heads conventionally referred to by a cover-term Determiner-head. Using Pereltsvaig’s phrasing cited above, it adds ‘the ability to pick

out an individual referent'. In other words, Ind-head transforms a property-like NP into an argument. In a sense, it marks the boundary between the lexical NP-zone and the functional NP-zone by signaling that the structure in its c-command domain is the description of the referent, functional layers that are added above can only specify the referent with respect to some discourse related properties, e.g. familiarity, definiteness, etc. In this respect Ind-head is similar to Infl-head in clauses and represents the nominal argument for the purposes of external agreement. Assume that Ind-head has a set of unvalued  $\phi$ -features; it can probe down and value them using the closest nominal in its c-command domain, then it can value the features of the main predicate, which takes this nominal as an argument. Small Nominal does not project IndP, thus although accepted as arguments, they do not have agreeing predicates, and the predicate surfaces in the default neuter form, see (29):

- (29) a. Mnogo volkov pal-o žertvami volč'ego terrora.  
 many wolves fell<sub>NEUT</sub> victims wolves slaughter  
 b. [[QP mnogo volkov] <sub>$\phi$ : 3pl</sub>] [Predicate-NEUT <sub>$\phi$ : \_</sub> ]]

QPs with individuated interpretation, on the other hand, project IndP and their predicates agree with the embedded noun, see (30):

- (30) a. Mnogo ljudej daval-i v sude položitel'nye xarakteristiki...  
 many people gave<sub>PL</sub> in court positive comments  
 b. [[IndP Ind <sub>$\phi$ : 3pl</sub>] [QP mnogo ljudej] <sub>$\phi$ : 3pl</sub>]] [Predicate-PL <sub>$\phi$ : 3pl</sub> ]]

Now, consider a construction, in which there are two nominals in a noun phrase with mismatching  $\phi$ -features. Ind-head should probe down and agree with the closest nominal. This scenario is illustrated with a pseudo-partitive construction in (31), in which the Q-head is filled with the nominal *oxapka* 'bunch-f.sg.nom'. That this nominal merges in the Q-head is suggested by the fact that it assigns the genitive of quantification to the noun. The noun *cvetov* 'flowers-f.pl.gen' has different  $\phi$ -features. However, the predicate can agree only with the nominal in Q-head, which is predicted by the proposal outlined above.

- (31) Oxapka        cvetov        ležal-a /\*-i    na stole.  
 bunch<sub>F.SG.NOM</sub> flowers<sub>F.PL.GEN</sub> lay<sub>SG</sub> /\*\_PL    on table  
 ‘A bunch of flowers was laying on the table.’

An interesting question is what would happen if Ind-head merged below the nominal Q-head or the two nominals were equidistant. The prediction is that the main predicate would be able to agree with the embedded nominal or with either one. The reason is that it is the Ind-head that determines the referent. This situation is, indeed, attested in Greek pseudo-partitive constructions. The Greek example in (32) from Alexiadou et al. (2007:424) shows that the predicate can bear plural and thus agree with *flowers*, as well as *a bunch*.

- (32) Ena buketo luludja itan pesmen-o/-a sto        patoma.  
 a    bunch flowers was/were thrown on.the floor

To summarize, I proposed in this section that the difference between the individuated and non-individuated ‘many-adv’ can be accounted for by the fact that the individuated ‘many-adv’ has an additional functional layer – IndP.

### 3.2 *The Analysis of ‘many-AGR: FamP*

With respect to the adjectival *many* in Russian, I propose to treat it on a par with ‘definite’ adjectives in Slavic. The agreement on *mnogie* ‘many-agr’ is the same agreement that distinguishes long-form (LF) adjectives from short-form (SF) adjectives in Russian. It is a shared belief among linguists that LF-adjectives in Slavic languages were derived by addition of the 3rd person singular pronoun *-ji* to a corresponding short form in prehistoric Slavic: *dobri* ‘kind-m.sf’ + *ji* → *dobriji* ‘kind-m.lf’, *dobra* ‘kind-f’ + *ja* → *dobraja* ‘kind-f.lf’ (e.g. Kramsky 1972, Schmalstieg 1976, Larsen 2007, Mladenova 2007, among others). At that time, if a noun was modified with an LF-adjective, it was interpreted as definite (or more precisely, familiar; see Larsen (2007)), SF-adjectives produced an indefinite interpretation. For some time, all Slavic languages enjoyed the definite/indefinite distinction in modified noun phrases. Then, some Slavic languages, including Russian, Czech and (standard) Bulgarian, lost this distinction. In Russian, SF-adjectives became specialized to occur only in the predicative position, whereas LF-adjective – only in the attributive

position (see Siegel 1976, Matushansky 2008, and Babby 2010, among many others).

This is a standard explanation. I would like to propose that in some contrastive cases, Russian retained ‘definiteness’ marking on adjectives. *Mnogie* is an example of such a case. We saw in section 2.2 that *mnogie*, unlike *mnogo*, has a familiar interpretation, is unacceptable in existential constructions and cannot be the argument of the verb with the cumulative prefix *na-*. To capture these facts, I propose that *mnogie* is a ‘definite’ adjective in Russian. It merges as a specifier of the Fam-head in the DP-zone. This explains the adjectival agreement on *mnogie* and its interpretative properties. This also explains why *mnogie* does not assign the genitive of quantification to the noun – it is not a Q-head. The structure is illustrated in (33). This treatment of *mnogie* is very similar to the proposal for ‘definite’ adjectives in Serbo-Croatian in Aljovich (2002).

- (33) a. mnog-ie studenty  
       many students  
       b. [<sub>FamP</sub> mnogie [<sub>Fam</sub> Fam [<sub>IndP</sub> Ind [<sub>NP</sub> studenty ]]]]

To summarize, I proposed in this section that the three *manys* in Russian have different structures. The difference between the two adverbial *manys* is the presence or absence of the Individuated Reference head. The head from the DP-zone provides the individuated interpretation to the quantity expression and is responsible for the agreement with the main predicate. The adjectival *many* is treated as a ‘definite’ adjective.

#### 4 Conclusion

Now, I would like to return to the questions raised in the introduction section and provide some partial answers. To remind the questions, I repeat the asymmetric division of different types of *many* in English and Russian in (34) (note that the schema to Russian is slightly more complex than presented in the introduction). The questions were: i) Why do the gaps in these systems exist? and ii) Why do these particular gaps exist in English and Russian? I.e. why does English lack a proportional adjectival *many* and Russian a cardinal adjectival *many* (not the other way around)?

(34)	a. <i>many</i> in English		b. <i>many</i> in Russian	
		Quant. Adj.		Adv. Adj.
	prop. many		prop.	mnogie
	card. many	the many	card. (ind)	mnogo
			(non-ind)	mnogo

As already mentioned above the two systems do not match perfectly. The two Russian adverbial *manys* are distinguished along the line of referentiality (individuation), rather than existential presupposition as in English and it is the adjectival *many* in Russian that has a familiar interpretation. It remains to be determined, what is the relation between the ability to have an individuated referent and proportionality (i.e. existential presupposition). It seems that referentiality is weaker than the existential presupposition. What I would like to propose is that the difference between *many* in English and Russian stems from the difference in the determiner, rather than quantificational, system of these two languages.

Let us consider Russian first. If the analysis in this paper is on the right track, the relevant part of Russian NP has the structure in (35). The ‘top’ part of the structure (FamP > IndP) is the DP-zone, which hosts ‘definite’ APs in Slavic. If FamP is not projected or more information on definiteness is needed, these functions are done by the discourse. Adjectives are merged low as adjuncts to NumP.

(35) [<sub>FamP</sub> def-AP [<sub>Fam'</sub> Fam [<sub>IndP</sub> Ind [<sub>QP</sub> Q [<sub>NumP</sub> AP [<sub>Num'</sub> Num [<sub>NP</sub> ]]]]]]]]]

Suppose that Russian once had two adjectival *manys*: definite (a long-form which we see today as *mnogie*) and indefinite (a short-form without the agreement which merged as an adjunct to NumP). There is no problem to keep the long-form (proportional) *many*; however, the short-form *many* was lost when short-form adjectives became specialized for the predicative position only. As we saw above, Russian *many* cannot be used in the predicative position. Thus, Russian does not have the adjectival *many* with cardinal (non-familiar) interpretation.

The role of the lacking short-form cardinal *many* in Russian is played by the adverbial *many* with the individuated reading. As we saw above, the most salient reading of the adverbial *mnogo* is a group-reading, but when the Ind-head is merged, it receives individuated reading and triggers

agreement on the main predicate, like nominals with the adjectival *many*, but assigns the genitive of quantification like the adverbial *many*. This is a strategy that the language uses to compensate for the lack of indefinite adjectival *many* (the two readings are equivalent to my knowledge).

Russian does not have a strong proportional adverbial *many* simply because it does not have (and never had) means to add definiteness apart from as inflection on adjectives. This account makes two predictions: i) Slavic languages that keep the distinction between definite and indefinite adjectives should have both strong and weak adjectival *many* (unless there are additional language-specific restrictions); ii) if a Slavic language has an overt definiteness marker, it should allow for both strong and weak adverbial *many*. The second prediction is born out. As shown in (36), Bulgarian definiteness marker *to* can attach to the adverbial *many*:

- (36) mnogo(to) knigi                      Bulgarian, Tasseva-Kurktchieva 2006  
       many<sub>(DEF)</sub> books

Finally, let us consider English. Suppose that English *the* spells out at least two heads FamP and *t*P as in Kyriakaki (2011), see (37). The adjectival *many* will merge low in NumP and combine with NP intersectively. Combined with *the*, it would be interpreted as unique, definite and cardinal. English does not have definite adjectives that can combine with the nominal, nor can it move an adjective to Spec-FamP to acquire familiarity reading with the exclusion of uniqueness (*t*P).

- (37) ([<sub>UP</sub> *t* [<sub>FamP</sub> Fam) [<sub>IndP</sub> Ind [<sub>QP</sub> Q [<sub>NumP</sub> AP [<sub>Num</sub> Num [<sub>NP</sub> ]]]]]]]  
       = *the*

This is also in line with the proposal in Giusti (1991) and subsequent work. To account for the data in English and Italian, Giusti (1991) argues that *many* can merge either as an adjunct to NP like adjectives, or as a functional head that selects either DP or NP depending on its semantics.

At the present moment I do not have anything to add to this picture, however, I believe that looking for the source of the differences between Russian and English *many* in the DP-domain is a promising approach.

## References

- Alexiadou, Artemis, Liliane M. V. Haegeman, and Melita Stavrou. 2007. *Noun Phrase in the Generative Perspective*. Berlin: Mouton de Gruyter.
- Aljovich, Nadira. 2002. Long adjectival inflection and specificity in Serbo-Croatian. In *Recherches linguistiques de Vincennes*, 31, 27–42.
- Babby, Leonard H. 2010. The syntactic differences between long and short forms of Russian adjectives. In *Adjectives: Formal Analyses in Syntax and Semantics*, ed. Ora Matushansky Patricia Cabredo Hofherr, 53–84. Amsterdam, Philadelphia: John Benjamins.
- Babko-Malaya, Olga. 1998. Context-dependent quantifiers restricted by focus. In *Proceedings of the Workshop on Focus*, ed. E. Benedicto, M. Romero, and S. Tomioka, 1–18. Amherst: GLSA.
- Bailyn, John F. 2004. The case of Q. In *Proceedings of FASL 12: The Ottawa Meeting 2003*, ed. G. Arnaudove et al., 1–36. Ann Arbor, MI: The University of Michigan Press.
- Baker, Mark C. 2003. *Lexical Categories. Verbs, Nouns, and Adjectives*. Cambridge: Cambridge University Press.
- Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4:159–219.
- Bowers, John. 1991. The syntax and semantics of nominals. *Proceedings from Semantics and Linguistic Theory* 1:1–30.
- Coppock, Elizabeth and David Beaver. 2014. Definiteness and determinacy. Ms.
- Franks, Steven, and Asya Pereltsvaig. 2004. Functional categories in the nominal domain. In *Proceedings of FASL 12. The Ottawa Meeting 2003*, ed. O. Arnaudova et al., 109–128. Ann Arbor, MI: University of Michigan Press.
- Giusti, Giuliana. 1991. The categorial status of quantified nominals. *Linguistische Berichte* 136:438–452.
- Hackl, Martin. 2009. On the grammar and processing of proportional quantifiers: ‘most’ versus ‘more than half’. *Natural Language Semantics* 17:63–98.
- Herburger, E. 1997. Focus and weak noun phrases. *Natural Language Semantics* 5:53–78.
- Keenan, Edward. 1987. A semantic definition of “indefinite NP”. In *The Representation of (In)definiteness*, ed. Eric Reuland and Aliceter

- Meulen, 286–317. Cambridge, MA: MIT Press.
- Kramsky, Jury. 1972. *The Article and the Concept of Definiteness in Language*. The Hague: Mouton de Gruyter.
- Krasikova, Sveta. 2011. On proportional and cardinal ‘many’. *Generative Grammar in Geneva* 7:93–114.
- Kyriakaki, Maria. 2011. DETs in the Functional Syntax of Greek Nominals. Doctoral Dissertation, University of Toronto.
- Larsen, Karin. 2007. The loss of the attributive short form in Old Russian as a process of grammaticalization. *Scando-Slavica* 53:33–57.
- Longobardi, Giuseppe. 1994. Reference and proper names: a theory of N-movement in syntax and Logical Form. *Linguistic Inquiry* 25:609–665.
- Matushansky, Ora. 2008. On the attributive nature of superlatives. *Syntax* 11:26–90.
- Milsark, Gary. 1977. Toward an explanation of certain peculiarities of the existential construction in English. *Linguistic Analysis* 3:1–29.
- Mladenova, Olga M. 2007. *Definiteness in Bulgarian: Modelling the Processes of Language Change*. Berlin, N. Y.: Mouton de Gruyter.
- Partee, Barbara H. 1989. Many quantifiers. *ESCOL 89: Proceedings of the Eastern States Conference on Linguistics*: 383–402.
- Pereltsvaig, Asya. 2006. Small nominals. *Natural Language and Linguistic Theory* 24:433–500.
- Rizzi, Luigi. 1997. The fine structure of the Left Periphery. In *Elements of Grammar*, ed. Liliane M. V. Haegeman, 281–337. Dordrecht: Kluwer.
- Roberts, Craige. 2000. Uniqueness in definite noun phrases. Ms. The Ohio State University.
- Schmalstieg, W. R. 1976. *An Introduction to Old Church Slavic*. Cambridge, Mass.
- Siegel, Muffy E.A. 1976. Capturing the Russian adjective. In *Montague Grammar*, ed. Barbara Partee, 293–309. New York: Academic Press.
- Solt, Stephanie. 2014 Q-adjectives and the semantics of quantity. *Journal of Semantics*, published online doi:10.1093/jos/fft018.
- Tasseva-Kurktchieva, Mila. 2006. The categorial status of quantifiers in Bulgarian: evidence for DP over QP, 378–393. Michigan Slavic Publications.

## **On Tensed Modals in Polish\***

*Łukasz Jędrzejowski*  
Universität Potsdam, Germany

This paper examines modal verbs in Polish in different temporal environments and illustrates that scope relationships between modal verbs and synthetic as well as analytic tense forms cannot be deduced from the external syntax, contrary to what has been commonly assumed. I will define the class of modal verbs in Polish based on the availability of two distinct modal bases, demonstrate to what extent they can combine with tense forms by looking more closely at the universal quantifier *musieć* ‘must’, and, finally, propose a new analysis.

### **1 Introduction**

Cross-linguistically, modal verbs (henceforth: MVs) are assumed to occupy two distinct syntactic positions. If they receive a non-epistemic

---

\* Some parts of this paper were presented at the conference *Formal Approaches to Slavic Linguistics 23* (FASL 23) in Berkeley (May 2014) and at the workshop *How Categorical are Categories?* in Wrocław (January 2013). I would like to thank heartily the participants for their questions, valuable comments and insightful discussions, in particular (in alphabetical order): Joanna Błaszczak, Wayles Browne, Bridget Copley, Kristin Melum Eide, Steven Franks, Valentine Hacquard, Dorota Klimek-Jankowska, Krzysztof Migdalski, Adam Przepiórkowski and Vesela Simeonova. Furthermore, I am indebted to Werner Abraham, Elly van Gelderen, Pierre-Yves Modicom, Małgorzata Szajbel-Keck, two anonymous reviewers and three anonymous abstract reviewers for their thought-provoking comments and interesting remarks. Obviously, all errors and inadequacies are my own. This work was supported in part by Bundesministerium für Bildung und Forschung (BMBF) (Grant Nr. 01UG0711).



- b. Hij zal moeten afwassen. T(Future) > Mod<sub>necessity</sub>  
 he will<sub>3SG</sub> must<sub>INF</sub> do.the.dishes<sub>INF</sub>  
 ‘He will have to do the dishes.’  
 (<sup>OK</sup>non-epistemic/\*epistemic)
- c. Hij moet zijn kamer hebben opgeruimd.<sup>3</sup> Mod<sub>epistemic</sub> > T(Past)  
 he must his room have<sub>INF</sub> clean.up<sub>PTCP</sub>  
 ‘He must have cleaned up his room.’  
 (\*non-epistemic/<sup>OK</sup>epistemic)

We can infer from (3a) and (3b) that syntactically neither the past perfect auxiliary *hebben* ‘have’ nor the future auxiliary *zullen* ‘will’ can take scope over *moeten*. If the temporal operators co-occur with MVs, they disambiguate their interpretation and only a non-epistemic interpretation is available. If a MV takes an epistemic modal base, as in (3c), the speech act time and the epistemic evaluation time coincide, even if the event time itself is rooted in the past. Accordingly, the embedded proposition falls under the scope of the epistemic MV (see also Hacquard 2006, 2010 who provides semantic arguments for the hierarchy given in 1).

Polish (and probably other West-Slavic) MVs behave differently. The scope relationships between MVs and different temporal operators cannot be deduced from the temporal syntax, let alone from the external syntax. See (4) for the universal quantifier *musieć* ‘must’ occurring with the future tense auxiliary *będzie* ‘will’ and, simultaneously, taking an epistemic modal base:

- (4) PO będzie musiała w końcu zacząć popełniać błędy.  
 PO will<sub>3SG</sub> must<sub>L-PTCP.3SG.F</sub> finally begin<sub>INF</sub> make<sub>INF</sub> mistakes  
 ‘≈ I suppose that PO (= a political party) will finally start making mistakes.’ (NKJP, *Dziennik Zachodni*, 9/10/2008)  
<sup>OK</sup>T(Future) > Mod<sub>epistemic</sub> / <sup>?</sup>Mod<sub>epistemic</sub> > T(Future)

According to the rigid hierarchy of functional projections and based on the Dutch data presented above, we would expect *musiała* to be interpreted non-epistemically. Note, however, that such a reading is very hard to

<sup>3</sup> (3c) also allows a reading according to which the propositional event is rooted in the future and the modal *moet* is evaluated against a deontic conversational background, even though it is accompanied by the past participle. I am not concerned with such cases in this paper.

obtain in (4).<sup>4</sup> Only an epistemic reading appears to be appropriate in this context. But what is more intriguing about (4) is that T(Future) outscopes Mod<sub>epistemic</sub>. In addition to that, a relatively free word order in Polish allows us to reverse the order of the future auxiliary *będzie* and the MV *musiała*:

- (4') PO *musiała będzie* w końcu zacząć popełniać błędy.  
<sup>OK</sup>epistemic/\*non-epistemic  
 Mod<sub>epistemic</sub> > T(Future) / \*T(Future) > Mod<sub>epistemic</sub>

The situation changes radically. In (4') *musiała* precedes *będzie* and gains scope over it. If MVs merging above TP are assumed to be interpreted epistemically, we expect *musiała* to be evaluated against an epistemic modal base. This prediction is borne out, as a non-epistemic reading of (4') is ruled out. Remarkably, though, (4') ought not to be taken as a representative example, if we want to draw far-reaching theoretical conclusions. Compare (5) mirroring the same word order and (6) with a topicalized infinitive in the front of *musiała*:

- (5) Po przyjęciu spadku *musiała będzie*  
 after receiving inheritance must<sub>L-PTCP.3SG.F</sub> will<sub>3SG</sub>  
 spłacić połowę długów.  
 pay.off<sub>INF</sub> half<sub>GEN</sub> debts<sub>GEN</sub>  
 'Having received the inheritance, she will have to pay a half of  
 the debts.' (NKJP, *Magazyn Puls Studenta*, 1/2001)  
<sup>OK</sup>epistemic/<sup>OK</sup>non-epistemic

<sup>4</sup> We can analyze *będzie* in (4) as an epistemic MV and claim that there exists an epistemic concord relationship between *będzie* and *musiała*. In this case, the scope mismatch would not occur. Note that this scenario cannot be maintained though. If it were the case, we would also expect other epistemic MVs to co-occur with *będzie* and to behave as *musiała* in (4) does. However, if we replace *musiała*, for instance, by the existential quantifier *mogła* 'may', which usually also allows epistemic readings (see section 2 below), it is disambiguated to the extent that only a non-epistemic reading occurs:

- (i) PO *będzie mogła* w końcu zacząć popełniać błędy.  
<sup>OK</sup>non-epistemic /\*epistemic

- (6) Każda z dziewcząt przebiec musiała będzie  
 each from girls run<sub>INF</sub> must<sub>L-PTCP.3SG.F</sub> will<sub>3SG</sub>  
 trasę jednego kilometra.  
 route<sub>ACC</sub> one<sub>GEN</sub> kilometre<sub>GEN</sub>  
 ‘Every girl will have to run 1 km.’<sup>OK</sup>epistemic/<sup>OK</sup>non-epistemic  
 (NKJP, *Dziennik Polski*, 14/9/2001)

Both (5) and (6) allow an epistemic as well as a non-epistemic interpretation. In the light of the data presented in this paper, I outline a new account of MVs in Polish and claim that they are base-generated as V-heads and move to a higher functional projection, ModP, either above or below TP, where they are semantically narrowed down by a modal base and a conversational background. In what follows, I will briefly define the class of MVs in Polish based on the availability of two distinct modal bases that a MV can take (section 2). Section 3 focuses on the universal quantifier *musieć* ‘must’ and demonstrates to what extent Polish MVs can combine with synthetic as well as with analytic tense forms. As it will turn out, no syntactic restrictions can be observed. Section 4 provides first steps of my own analysis. Finally, section 5 concludes the paper.

## 2 Modal Verbs in Polish

Polish MVs do not differ from lexical verbal heads merging within the verbal phrase. Members of both groups, for instance, undergo a V-to-T movement to check some formal features within TP and, to the best of my knowledge, there are no observable syntactic differences making them belong to one or the other class. Therefore, I put aside all syntactic criteria and adopt the following semantic definition of MVs:

- (7) A verb is a modal verb iff it is evaluated against a non-epistemic and against an epistemic modal base.

According to this definition, we can identify five MVs in Polish. I illustrate their use based on the pattern [MV<sub>present tense</sub> + infinitive].

(i) *móc* 'can, may, be allowed':

- (8) a. Teraz możesz grzeszyć. [non-epistemic]  
 now can<sub>2SG</sub> sin<sub>INF</sub>  
 'Now you can/may sin.' (BKR, p. 69)
- b. W czwartek może padać śnieg. [epistemic]  
 in Thursday can<sub>3SG</sub> snow<sub>INF</sub>  
 'It may be snowing on Thursday.'  
 (NKJP, *Polski Głos Wielkopolski*, 10/1/2005)

(ii) *mieć* (lit. 'have') 'have to, must, be said, be claimed':

- (9) a. Masz wyjść i zastrzelić ją! [non-epistemic]  
 have<sub>2SG</sub> go.out<sub>INF</sub> and shoot<sub>INF</sub> her<sub>ACC</sub>  
 'You have to go out and shoot her!'  
 (NKJP, *Dziennik Zachodni*, 30/10/2009)
- b. W okolicy ma powstać parking. [epistemic]  
 in vicinity have<sub>3SG</sub> be.built<sub>INF</sub> car.park  
 'A car park is supposed to be built nearby.'  
 (NKJP, *Mazowieckie To i Owo*, 23/4/2009)

If the subject is equipped with the feature [+human], two evidential interpretations occur:

- (10) Migalski ma mieć jakiś program w telewizji.  
 M. have<sub>3SG</sub> have<sub>INF</sub> a program in television
- a. 'Migalski is supposed to get a program on television.'  
 (information source of *p* = foreign (unknown) source)
- b. 'Migalski claims to get a program on television.'  
 (information source of *p* = clause subject)  
 (UwRz 7/(54), p. 7)

In (10), the information source of the embedded proposition can be attributed to two different individuals. It can be a person who is not included in the discourse, i.e. neither the speaker nor the hearer. In this case, the source can remain unknown or be specified by additional means, for instance, by the phrase *zgodnie z* 'according to'. It can also refer to the

clause subject, meaning that *Migalski* himself argues that he will get a program on TV.

(iii) *musieć* 'must, have to':

- (11) a. Musimy już iść. [non-epistemic]  
 must<sub>1PL</sub> already go<sub>INF</sub>  
 'We must go now.' (NKJP, *Chaszcze* 2009)
- b. Musi się czuć jak szejk. [epistemic]  
 must<sub>3SG</sub> REFL feel<sub>INF</sub> like sheik  
 'He must be feeling like a sheik.' (UwRz 22/(69), p. 17)

(iv) *powinien* 'should, be supposed':

- (12) a. Kościół powinien wyciągnąć wnioski. [non-epistemic]  
 Church should<sub>3SG</sub> draw<sub>INF</sub> conclusions  
 'The Church should draw conclusions (from that).' (BKR, p. 38)
- b. Nowy sprzęt powinien się pojawić  
 new equipment should<sub>3SG</sub> REFL appear<sub>INF</sub>  
 u nas za kilka miesięcy. [epistemic]  
 at us in few months  
 'The new equipment is supposed to be delivered in a couple of months.' (NKJP, *Dziennik Zachodni*, 18/6/2008)

(v) *winien* 'should, be supposed':<sup>5,6</sup>

---

<sup>5</sup> In comparison to *móc*, *mieć* and *musieć*, *powinien* and *winien* are defective. First, they do not possess an infinitive form. Second, they do not inflect for the synthetic past tense. If they are used in past contexts, the past tense auxiliary *być* 'be' is required:

(i) Powinien był pomyśleć o innych.  
 should<sub>3SG</sub> be<sub>3SG.M</sub> think<sub>INF</sub> about others  
 'He should have think about the others.' (NKJP, *Dziennik Zachodni*, 14/12/2001)

It is also worth mentioning that the presence of the past tense auxiliary *być* automatically gives rise to a counterfactual reading in the past.

<sup>6</sup> Due to the fact that *powinien* and *winien* do not inflect for the synthetic past tense, they occur with the past tense auxiliary *być* 'be', if the embedded proposition is rooted in the past. However, this auxiliary is often dropped in spoken Polish giving rise to two different temporal interpretations:

- (13) a. Debata publiczna winna być rzeczowa. [non-epistemic]  
 debate public should<sub>3SG</sub> be<sub>INF</sub> argumentative  
 'A public debate should be argumentative.'  
 (UwRz 13(60), p. 6)
- b. Bez oleju zaczyn winien się udać. [epistemic]  
 without oil sourdough should<sub>3SG</sub> REFL succeed<sub>INF</sub>  
 'Without oil the sourdough is supposed to come off, too.'  
 (Abraham et al. 2011: 163)

Additionally, Błaszczak et al. (2010: 10) assume *będzie* 'will', the perfective verb form of *być* 'be', to function as an epistemic modal verb:

- (14) A: Somebody is knocking at the door. Who do you think is this?  
 B: To będzie Ewa.  
 this will<sub>3SG</sub> E.  
 'This will be Ewa.'

However, one classification problem arises with ranking *będzie* as a MV. If we assume all future-oriented readings to be epistemic, there is no possibility to treat *będzie* as a MV based on the semantic definition given in (7) above. If *będzie* cannot be evaluated against a non-epistemic modal base, it does not meet the criterion of the availability of two distinct modal bases. Therefore, I would classify *będzie* as a modal predicate taking only an epistemic modal base. Kissine (2008), however, illustrates that all modal flavors of future auxiliaries do not really come from the semantics of the auxiliary itself, but from a pragmatic mechanism restricting the domain of a covert epistemic operator scoping over the entire embedded proposition. Following this line of reasoning, English *will* and Polish *będzie* are inherent tense operators. I do not elaborate on *będzie* in more detail, leave it aside here and concentrate on the five MVs listed above. Table 1 gives an overview:

- 
- (i) Rodzice (po-)winni zaopiekować się dzieckiem.  
 parents should<sub>3PL</sub> take.care<sub>INF</sub> REFL child<sub>INS</sub>  
 a. 'The parents should take care of the child.'  
 b. 'The parents should have taken care of the child.'

The MV in (i) expresses a weak necessity, either projecting into future or having a future-in-the-past reading. I have no explanation for why *powinien* and *winien* behave this way and I leave it open here.

		non-epistemic	epistemic
1.	<i>móc</i>	✓	✓
2.	<i>mieć</i>	✓	✓
3.	<i>musieć</i>	✓	✓
4.	<i>powinien</i>	✓	✓
5.	<i>winien</i>	✓	✓

Table 1: Modal verbs in Modern Polish

In the next section, I demonstrate to what extent Polish MVs can combine with different tense forms.

### 3 Polish Modal Verbs and Tense Forms

As it has been shown in the literature, non-epistemic MVs can combine with all kinds of synthetic and analytic tense forms, whereas their epistemic counterparts are usually resistant to most analytic tense forms (cf. 3a and 3b above for Dutch and Wurmbrand 2001 for German or Picallo 1990 for Catalan, among many others). What appears to be different about Polish is that MVs taking an epistemic modal base are compatible with all synthetic and analytic tense forms. Contrary to what we would be expecting from Germanic and Romance data, no syntactic restrictions occur. In order to demonstrate this, I focus on the universal quantifier *musieć* 'must' and its co-occurrence possibilities with various temporal operators. I will show that a particular tense form - regardless of whether synthetic or analytic - does not disambiguate the reading of the modal.

*The Present Tense.* Similar to its Germanic and Romance counterparts, *musieć* 'must' can be interpreted both non-epistemically and epistemically:<sup>7</sup>

- (15) a. Robotnicy muszą opuścić plac budowy.  
workers must<sub>3PL</sub> leave<sub>INF</sub> building site  
'The workers must leave the building site.'  
(NKJP, *Mazowieckie To i Owo*, 30/4/2009)

<sup>7</sup> MVs occurring in a-examples are non-epistemic, in b-examples epistemic.

- b. To musi być pomyłka.  
 this must<sub>3SG</sub> be<sub>INF</sub> mistake  
 'It must be a mistake.'  
 (NKJP, *Dziennik Zachodni*, 13/3/2007)

*The Synthetic Past Tense.*<sup>8</sup> Polish MVs - except for *powinien* and *winien* (see footnote 5 above) - can bear the synthetic past tense morphology. The past morphology has no impact on the modal interpretation:

- (16) a. Wszystkiego musiała nauczyć się sama.  
 all must<sub>L-PTCP.SG.M</sub> learn<sub>INF</sub> REFL alone  
 'She had to learn everything alone.'  
 (NKJP, *Mazowieckie To i Owo*, 7/8/2008)
- b. Nieopodal musiała istnieć większa osada.<sup>9</sup>  
 nearby must<sub>L-PTCP.SG.M</sub> exist<sub>INF</sub> big<sub>COMPAR</sub> settlement  
 'There must have been a bigger settlement.'  
 (NKJP, *Gazeta Wroclawska*, 24/10/2003)

<sup>8</sup> Strictly speaking, there is no synthetic past tense in Modern Polish. What we have instead is a compound tense form consisting of an *l*-participle and a clitic attached to the *l*-participle (cf. Sussex 1980, Booij & Rubach 1987, Spencer 1991 and Migdalski 2006: 223-285). The clitics, in turn, are treated as perfect auxiliaries (for their emergence see in particular Migdalski 2013). For the sake of simplicity, I label this tense form as a synthetic past tense in order to distinguish it from the analytic pluperfect (see below). We observe a similar situation in German and Dutch. It has been assumed that the past tense of weak verbs formed with the dental suffixes *-t-* or *-d-* emerged out of the verb *tun/doen* 'do'. The only difference between Polish clitics and the West-Germanic dental suffixes is that the former can float (cf. Embick 1995, Kupść 2005), whereas the latter cannot. I would like to thank Krzysztof Migdalski who brought this issue to my attention.

<sup>9</sup> One of the anonymous reviewers suggests a preliminary analysis of (16a): "the modal can just head-move to its Mod-epist position via T-past, picking the tense features surfacing as past morphology on its way." At first sight, this solution appears to be very attractive. However, it does not seem to be what we have observed so far based on the Dutch data above and what we know from the cross-linguistic literature. If epistemic MVs are base-generated above TP, there is no technical possibility for them to pick tense features on their way. It would contradict the hierarchies given in (1) and (2) above. Of course, we can assume Polish epistemic MVs to tense-lower, but then two additional problems arise. First, it remains unclear how to analyze epistemic MVs occurring with analytic tense forms, if a tense auxiliary occupies a T-head (cf. e.g. 4' above). Second, epistemic MVs need not outscope TP (cf. Homer 2010 and Mari & Schweitzer 2010). If they do not take the scope over TP, they should go one more layer down, below TP.

Borgonovo & Cummins (2007) illustrate that if Spanish MVs bear the past tense morphology, three different modal readings appear to be appropriate (see also Laca 2012). An epistemic reading is also available:<sup>10</sup>

- (17) Pedro debió ganar la carrera.  
 Pedro must<sub>3SG.PAST</sub> win<sub>INF</sub> the race  
 a. 'Pedro must have won the race.' [epistemic]  
 b. 'Pedro was forced to win the race.' [actuality entailment]  
 c. 'Pedro should have won the race.' [counterfactual]  
 (Borgonovo & Cummins 2007: 6)

Polish epistemic MVs pattern with their Spanish counterparts allowing an epistemic reading with the past morphology.

*The -no/-to Suffixation.* Polish possesses an impersonal suffix with two allomorphs, *-no* and *-to*, attaching to a verbal head:<sup>11</sup>

- (18) a. Grano, śpiewano, tańczono.  
 play-no sing-no dance-no  
 'They played, sang, danced.'  
 (NKJP, *Gazeta Poznańska*, 1/12/2005)  
 b. Przebito mu oponę w samochodzie.  
 puncture-to him<sub>DAT</sub> tire<sub>ACC</sub> in car<sub>LOC</sub>  
 'They punctured a tire in his car.'  
 (NKJP, *Dziennik Łódzki*, 26/8/2005)

<sup>10</sup> Note that epistemic MVs in Germanic languages do not usually bear the past tense morphology:

- (i) Nach dem Elfmeter musste das Spiel kippen.  
 after the penalty must<sub>3SG.PAST</sub> the game change<sub>INF</sub>  
 'After the penalty the game had to change.'  
 (<sup>OK</sup>non-epistemic/\*epistemic) (Reis 2007: 13)

There are special cases, however, in which epistemic MVs can occur in past environments (e.g. in free indirect discourse):

- (ii) Ich wusste, dass er da sein musste.  
 I know<sub>3SG.PAST</sub> that he there be<sub>INF</sub> must<sub>3SG.PAST</sub>  
 'I knew that he must have been there.' (Klein 2009: 320)

Due to the lack of space, I do not elaborate on special cases in this paper.

<sup>11</sup> We can identify the EPP features of the *pro* subject in the *-no/-to* constructions: [+plural], [+virile], [+human] (for more details see Dziwirek 1994, Kibort 2004, 2008 and Krzek 2010).

The *-no/-to* suffix anchors the embedded event time prior to the speech time ( $t_1 > t_{\text{SPEECH}}$ ):<sup>12</sup>

- (19) a. Twierdzi, że przeczytano ten list.  
 claim<sub>3SG</sub> that read-no this letter<sub>ACC</sub>  
 'He claims that this letter has been read.'  
 b. Twierdzi, że przebito mu oponę.  
 claim<sub>3SG</sub> that puncture-to him<sub>DAT</sub> tire<sub>ACC</sub>  
 'He claims that they punctured a tire in his care.'

Similar to the synthetic past tense, the *-no* morpheme can merge both with non-epistemic and with epistemic MVs:<sup>13</sup>

- (20) a. Musiano jej założyć 89 szwów.  
 must-no her<sub>DAT</sub> set<sub>INF</sub> 89 stitches<sub>GEN</sub>  
 'She had to get 89 stitches.'  
 (NKJP, *Cosmopolitan*, 7/2000)  
 b. O tych wydarzeniach musiano wiedzieć w Polsce.  
 about these events must-no know<sub>INF</sub> in Poland  
 'They must have known about these events in Poland.'  
 (NKJP, *Dynastia Piastów w Polsce*, 2005)

*The Analytic Pluperfect.* As far as the past tense forms are concerned, Polish also possesses an analytic pluperfect. The verb complex consists of three elements: (i) a modal verb in the form of an *l*-participle, (ii) the

---

<sup>12</sup> Notice, however, that if *-no* and *-to* morphemes are embedded under a desiderative predicate, a future-oriented reading of the embedded proposition is forced:

- (i) Pragnie, żeby przeczytano ten list.  
 wish<sub>3SG</sub> that read-no this letter<sub>ACC</sub>  
 'He wants us to read this letter.'  
 (ii) Pragnie, żeby przebito mu oponę.  
 wish<sub>3SG</sub> that puncture-to him<sub>DAT</sub> tire<sub>ACC</sub>  
 'He wants them to puncture a tire in his car.'

The speech time coincides with the matrix event time (= wishing) and its volitional illocutionary force shifts the embedded event time (= reading) into future.

<sup>13</sup> The suffix *-no* cannot attach to the MV *móc* 'can' in Modern Polish, though. At this moment I have no explanation for why this is so. I would like to thank Wayles Browne who pointed this out to me.

auxiliary verb *być* 'be' and (iii) an infinitive. Again, no semantic restrictions follow from the temporal syntax:

- (21) a. Poeta musiał był wyjechać do Londynu.  
 poet must<sub>L-PTCP.3SG.M</sub> be<sub>L-PTCP.3SG.M</sub> move<sub>INF</sub> to London  
 'The poet had to move to London.'  
 (NKJP, *Przestrzeń dzieł wiecznych*, 1993)
- b. Ulewa musiała była przejść.  
 downpour must<sub>L-PTCP.3SG.F</sub> be<sub>L-PTCP.3SG.F</sub> pass<sub>INF</sub>  
 'A downpour must have been passed.'  
 (NKJP, *Pokój i Diament*, 1948)

The analytic pluperfect sounds archaic in Modern Polish. Nevertheless, examples illustrating its use are very easy to find, also with other MVs:

- (22) Mistrz Li posiał ziarno, które  
 Master Li sow<sub>L-PTCP.SG.M</sub> grain<sub>ACC</sub> which  
 mogło było wydać plon obfitszy.  
 can<sub>L-PTCP.SG.M</sub> be<sub>L-PTCP.SG.M</sub> give<sub>INF</sub> crop<sub>ACC</sub> bountiful<sub>COMPAR</sub>  
 'Master Li sowed a grain that might have brought better results.'  
 (*Polityka* 52/2788, p. 9)

The example given in (22) poses a challenge for the theory according to which the Modern Polish clause is not equipped with TP (cf. Bošković 2012). If Modern Polish does not possess a TP, it remains unclear what the syntactic position of the auxiliary *być* in (21a,b) and (22) is. Wayles Browne (pers. comm.) draws my attention to the fact that MVs occurring with the analytic pluperfect usually inflect for the 3rd person singular, 1st and 2nd persons, are in turn more rare. A plausible explanation might be that 1st and 2nd persons require additional presence of auxiliary clitics merging with MVs:

- (23) Wcale nie musiałem byłem tego postu czytać.  
 at.all NEG must<sub>L-PTCP.M.1SG</sub> be<sub>L-PTCP.M.1SG</sub> this post read<sub>INF</sub>  
 'I didn't have to read this post at all.'  
 (NKJP, an internet forum, 19/3/2001)

As Migdalski (2006: 228) states, “the singular and plural variants of the 3rd person are morphologically null.” As it turns out, Polish non-epistemic and epistemic MVs are com-patible with all past tense forms.

*The Analytic Future Tense.* The analytic future tense imposes no restrictions on the interpretability of the embedded modal either. As (24a) and (4) - repeated here as (24b) - illustrate, the analytic future auxiliary *będzie* does not disambiguate the modal reading of *musieć*:

- (24) a. Gmina będzie musiała pokryć wszystkie koszty.  
 town will<sub>3SG</sub> must<sub>L-PTCP.3SG.F</sub> cover<sub>INF</sub> all costs<sub>ACC</sub>  
 'The town will have to cover all costs.'  
 (NKJP, *Dziennik Zachodni*, 24/1/2008)
- b. PO będzie musiała w końcu zacząć popełniać błędy.  
 PO will<sub>3SG</sub> must<sub>L-PTCP.3SG.F</sub> finally begin<sub>INF</sub> make<sub>INF</sub> mistakes  
 '≈ I suppose that PO (= a political party) will finally start to  
 make mistakes.' (NKJP, *Dziennik Zachodni*, 9/10/2008)

The examples given in (4) and (4') clearly demonstrate that the epistemic MV *musieć* can precede and follow the future auxiliary *będzie* and that the structural position of the modal does affect its interpretation. However, it is not always the case that a MV can follow a tense auxiliary. We cannot reverse the word order of *nie musiałem* and *byłem* in (23):

- (23') \*Wcale byłem nie musiałem tego postu czytać.

Note that the contrast between (23) and (23') does not come from the presence and the position of *nie*:<sup>14</sup>

- (25) Prawdopodobnie mogłeś byłeś to naprawić.  
 probably can<sub>L-PTCP.M.2SG</sub> be<sub>L-PTCP.M.2SG</sub> this fix<sub>INF</sub>  
 'Probably you might have been able to fix this.'
- (25') \*Prawdopodobnie byłeś mogłeś to naprawić.

<sup>14</sup> At this moment, I have no concrete explanation for why this is so. I speculate that the ungrammaticality of (23') and (25') is linked to the *l*-participle morphology of the auxiliary *być*. Presumably, its  $\phi$ -features cannot be checked, once they have been checked by a MV first and then erased. In this case *być* could not establish a probe-goal relation to get values, making the derivation crash.

Summarizing, the semantic interpretation of Polish MVs does not follow from the external temporal syntax, as has been commonly assumed in the literature on Germanic and Romance modals. Polish MVs can occur in all temporal environments and their syntactic position with respect to tense auxiliaries is rather free. Table 2 gives a general overview:

		non-epistemic	epistemic
1.	Present Tense	✓	✓
2.	Synthetic Past Tense	✓	✓
3.	The Past <i>-no</i> Suffix	✓	✓
4.	Analytic Pluperfect	✓	✓
5.	Analytic Future	✓	✓

Table 2: Polish *musieć* 'must' and its compatibility with tense forms

In the next section, I outline a new account of the data presented above.

#### 4 A New Account

So far I have defined the class of MVs in Polish and demonstrated that they can occur in all temporal environments. If their semantics cannot be determined by the presence/absence of a tense auxiliary, it does not seem to be reasonable to posit two distinct structural positions, a higher one for epistemic MVs and a lower one for their non-epistemic counterparts. Mainly, I argue that Polish MVs (i) are base-generated as V-heads,<sup>15</sup> (ii) move to one of the ModPs, and (iii) their particular interpretation (non-epistemic vs. epistemic) is contextually determined:

---

<sup>15</sup> Zagona (2008) adopts a similar approach to English MVs. Her main claim is that English modals occupy a single syntactic position within TP and that their interpretation (non-epistemic vs. epistemic) depends on (un)interpretability of features and on "the properties of the phase in which the modal is merged. (...) root modals are interpreted in the v\*P phase and epistemic modals in the CP phase. It is argued that modals can be merged in either phase, according to the inflectional features that are added to the lexical item as it enters the syntax" (ibid. 274-5).



situation holds for the analytic pluperfect with the auxiliary verb *być* bearing the *l*-participle morphology (cf. 21a,b). The only difference is that the latter pattern does not allow the order [*być*<sub>*l*-participle</sub> + *MV*<sub>*l*-participle</sub>] (see footnote 14 for a possible explanation of this restriction). Now let us consider a more complex example:

- (28) *Prawdopodobnie mogłeś byłeś to naprawić.*  
 probably can<sub>L-PTCP.M.2SG</sub> be<sub>L-PTCP.M.2SG</sub> this fix<sub>INF</sub>  
 'Probably you might have been able to fix this.'

I ignore the semantics of *mogłeś* and concentrate on the derivation mechanism. What appears to be intriguing about (28) is the PF realization of the clitic *-ś* both on the modal and on the pluperfect auxiliary. Małgorzata Szajbel-Keck (pers. comm.) provides more corpus examples from Colloquial Polish:

- (29) *Ale mogłeś byłeś napisać o co chodzi.*  
 but can<sub>L-PTCP.M.2SG</sub> be<sub>L-PTCP.M.2SG</sub> write<sub>INF</sub> about what matter<sub>3SG</sub>  
 'But you could have written what you had in mind.'
- (30) *Ale mogłeś byłeś to jakoś*  
 but can<sub>L-PTCP.M.2SG</sub> be<sub>L-PTCP.M.2SG</sub> this somehow  
*delikatniej ująć w słowa.*  
 more.mildly capture<sub>INF</sub> in words  
 'But you could have put this differently, more mildly.'

I assume that *móc* moves from within VP up to the higher ModP. On its way, the modal merges with the clitic *-ś* within TP. In this connection, the question arises how the clitic adjoins to the pluperfect auxiliary when it has moved higher in the structure with the modal. In order to account for this fact, I argue that *-ś* attached to *był-* in (28) is an overt copy of the clitic which has not been deleted at the PF level after the movement had taken place. We observe a similar situation in some varieties of English in which the auxiliary verb *have* may be duplicated:<sup>16</sup>

<sup>16</sup> Nunes (2004: 43-50) also discusses other cases of overt copies in natural languages, in particular clitic duplication in some dialects of Argentinean Spanish, verb duplication in Vata, a Niger Congo language of the Kru family, and postposition duplication in Panara, a Brazilian indigenous language.

- (31) They might've not have left.  
(Nunes 2004: 170, fn 48)

If there is no need to spell-out two copies overtly, as in the case of the future tense with *będzie*, only one of them is pronounced at PF. Following one of the Chain Reduction Principles proposed in Nunes (2004), we delete all but the copy with the fewest unchecked features. In (28), in turn, the spell-out of the lower copy is optional:

- (28') Prawdopodobnie mogłeś byłeś to naprawić.  
(32) Mogłeś był widzieć ślad moich bosych nóg.  
can<sub>L-PTCP.M.2SG</sub> be<sub>L-PTCP.M.3SG</sub> see<sub>INF</sub> trace my bare legs<sub>GEN</sub>  
'You could have seen a trace of my bare feet.'  
(NKJP, Stefan Żeromski, 1900, *Ludzie Bezdomni*)

What triggers this optionality still remains to be investigated. Finally, we have to ask how to derive a particular modal interpretation of MVs in Polish, if they are not sensitive to temporal environments. Kratzer (1977, 1981, 1991) convincingly illustrates that a conversational background determines the set of worlds MVs quantify over, meaning that the particular interpretation (non-epistemic vs. epistemic) follows from the context. This leads us to the conclusion that external syntax is not powerful enough in Polish to disambiguate elements merging in ModPs.

## 5 Conclusions

In this paper, I have demonstrated that the interpretation of Polish MVs does not follow from the linear word order of tense and modal operators and that Polish MVs can occur in all synthetic as well as analytic tense forms, making them considerably different from their Germanic and Romance counterparts. To the best of my knowledge, the indifference of Polish (epistemic) modals to tense operators has so far gone unnoticed in the literature on MVs in general. I have argued that interpretative differences in the semantics of Polish MVs follow solely from the modal base and the conversational background that a MV takes.

### Primary sources

- BKR – Szymon Hołowina & Marcin Prokop. 2011. *Bóg, kasa i rock'n'roll* ['God, cash and rock'n'roll].
- NKJP – Narodowy Korpus Języka Polskiego ['National Corpus of Polish']. Available at: <http://nkjp.pl>.
- Polityka*. ['Politics']. A weekly magazine.
- UwRz - *Uważam Rze* ['I think t(h)at']. A weekly magazine.

### References

- Abraham, Werner, Peter Meihsner, Jadwiga Piskorz & Kinga Piskorz. 2011. Zur paradigmatischen Ausbildung von Modalverben, von Deontik und Epistemik im jüngeren gesprochenen Polnisch. In *Germanistik in Polen*, ed. Andrzej Kątny & Katarzyna Lukas, 159-190. Frankfurt am Main: Peter Lang.
- Błaszczak, Joanna, Patrycja Jabłońska, Dorota Klimek-Jankowska & Krzysztof Migdalski. 2010. The Riddle of the future tense in Polish: how much “future” is there in “future tense”? Talk and handout delivered at the 43rd SLE Meeting, Vilnius University, Workshop on “Future Tense(s)/Future Time(s)”, 2-5 September.
- Booij, Geert & Jerzy Rubach. 1987. Postcyclic versus postlexical rules in lexical phonology. *Linguistic Inquiry* 18: 1-44.
- Borgonovo, Claudia & Sarah Cummins. 2007. Tensed modals. In *Coreference, Modality, and Focus*, ed. Luis Eguren & Olga Fernández Soriano, 1-18. Amsterdam: John Benjamins.
- Bošković, Željko. 2012. On NPs and clauses. In *Discourse and Grammar. From Sentence Types to Lexical Categories*, ed. Günther Grewendorf & Thomas Ede Zimmermann, 179-242. Berlin: de Gruyter.
- Butler, Jonny. 2004. *Phase Structure, Phrase Structure, and Quantification*. PhD thesis, The University of York.
- Butler, Jonny. 2006. The structure of modality and temporality (or, Towards Deriving like a Cinque Hierarchy). *Linguistic Variation Yearbook* 6: 161-201.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads. A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Cinque, Guglielmo. 2006. *Restructuring and Functional Heads*. Oxford: Oxford University Press.

- Condoravdi, Cleo. 2001. Temporal interpretation of modals: modals for the present and for the past. In *Stanford Papers on Semantics*, ed. David Beaver, Stefan Kaufmann, Brady Clark & Luis Casillas, 1-30. Stanford: CSLI Publications.
- Dziwirek, Katarzyna. 1994. *Polish Subjects*. New York: Garland.
- Embick, David. 1995. Mobile inflections in Polish. In *NELS 25: Proceedings of the 25th Annual Meeting of the North East Linguistics Society*, ed. Jill Beckman, 127-142. Amherst, MA: GLSA.
- Hacquard, Valentine. 2006. *Aspects of Modality*. PhD thesis, MIT.
- Hacquard, Valentine. 2010. On the event relativity of modal auxiliaries. *Natural Language Semantics* 18: 79-114.
- Homer, Vincent. 2010. Epistemic modals: high ma non troppo. (<http://ling.auf.net/lingbuzz/001086>).
- Kibort, Anna. 2004. *Passive and Passive-Like Constructions in English and Polish*. PhD thesis, University of Cambridge.
- Kibort, Anna. 2008. Impersonal in Polish: An LFG Perspective. *Transaction of the Philological Society* 106: 246-289.
- Kissine, Mikhail. 2008. Why *will* is not a modal. *Natural Language Semantics* 16: 129-155.
- Klein, Katarina. 2009. Semi-modal variation. In *Describing and Modeling Variation in Grammar*, ed. Andreas Dufter, Jörg Fleischer & Guido Seiler, 297-324. Berlin: de Gruyter.
- Kratzer, Angelika. 1977. What *must* and *can* must and can mean. *Linguistics and Philosophy* 1: 337-355.
- Kratzer, Angelika. 1981. The notional category of modality. In *Words, Worlds, and Contexts. New Approaches in Word Semantics*, ed. Hans-Jürgen Eikmeyer & Hannes Rieser, 38-74. Berlin: de Gruyter.
- Kratzer, Angelika. 1991. Modality. In *Semantik: Ein internationales Handbuch zeitgenössischer Forschung*, ed. Arnim von Stechow & Dieter Wunderlich, 639-650. Berlin: de Gruyter.
- Krzek, Małgorzata. 2010. Some aspects of subjects of impersonal constructions in Polish. *Newcastle Working Papers in Linguistics* 16: 66-87.
- Kupść, Anna & Jesse Tseng. 2005. A new HPSG approach to Polish auxiliary constructions. In *Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar, Department of Informatics, University of Lisbon*, ed. Stefan Müller, 253-273. Stanford, CA: CSLI Publications.

- Laca, Brenda. 2012. On modal tenses and tensed modals. In *Building a Bridge between the Linguistic Communities of the Old and the New World. Current Research in Tense, Aspect, Mood and Modality*, ed. Chiyo Nishida & Cinzia Russi, 163-198. Amsterdam: Rodopi.
- Mari, Alda & Susan Schweitzer. 2010. Calculating the epistemic interpretation of past modals via K. In *WCCFL 28 Online Proceedings*, ed. by Mary Byram Washburn, Sarah Ouwayda, Chuoying Ouyang, Bin Yin, Canan Ipek, Lisa Marston & Aaron Walker (<https://sites.google.com/site/wccfl28pro/mari-schweitzer>).
- Migdalski, Krzysztof. 2006. *The Syntax of Compound Tenses in Slavic*. PhD thesis, Universiteit van Tilburg.
- Migdalski, Krzysztof. 2013. Diachronic source of two cliticization patterns in Slavic. In *Challenging Clitics*, ed. Christine Meklenborg Salvesen & Hans Petter Helland, 135-158. Amsterdam: John Benjamins.
- Nunes, Jairo. 2004. *Linearization of Chains and Sideward Movement*. Cambridge, Mass.: MIT Press.
- Palmer, Frank. 2001. *Mood and Modality*. Cambridge: Cambridge University Press.
- Picallo, Carme. 1990. Modal verbs in Catalan. *Natural Language and Linguistic Theory* 8: 285-312.
- Reis, Marga. 2007. Modals, so-called semi-modals, and grammaticalization. *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* 12: 1-57.
- Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell.
- Sussex, Roland. 1980. On agreement, affixation and enclisis in Polish. In *Morphosyntax in Slavic*, ed. Catherine V. Chvany & Richard D. Brecht, 187-203. Columbus, OH.
- Taleghani, Azita H. 2008. *Modality, Aspect and Negation in Persian*. Amsterdam: John Benjamins.
- Wurmbrand, Susanne. 2001. *Infinitives. Restructuring and Clause Structure*. Berlin: de Gruyter.
- Zagona, Karen. 2008. Phasing in modals: phases and the epistemic/root distinction. In *Time and Modality*, ed. Jacqueline Guéron & Jacqueline Lecarme, 273-291. Berlin: Springer.

## **On the Degrammaticalization of Pronominal Clitics in Slavic\***

*Hakyung Jung*  
Seoul National University

*Krzysztof Migdalski*  
University of Wrocław

This paper investigates the change of pronominal clitics into weak pronouns in the Slavic languages. Specifically, it intends to show that pronominal clitics, which are heads, were degrammaticalized into weak pronouns as phrases in Old Polish, Old Russian and in some contexts in Macedonian. This paper also establishes the trigger of this process in these languages and shows the way it proceeded. The investigation eventually demonstrates that grammaticalization can be disturbed and reversed by other changes in the same linguistic system.

### **1 Degrammaticalization of Clitics: The Reverse of XP to X<sup>0</sup> Change**

#### *1.1 Grammaticalization and Unidirectionality Hypothesis*

Grammaticalization is the change of a lexical item to a grammatical one, and a grammatical item to a more grammatical one (Meillet 1912, Kuryłowicz 1975), which typically involves phonological and semantic weakening (Heine and Reh 1984). It has been widely assumed that

---

\* We would like to thank the FASL participants, especially Željko Bošković and Wayles Browne, for comments and discussion. We are also grateful to the two anonymous reviewers for their helpful suggestions. The usual disclaimer applies.

grammaticalization is a unidirectional process from lexicon to grammar and until recently it has been argued that changes in the opposite direction are impossible. The chain of changes in (1) is a representative example of the directionality of grammaticalization.

- (1) Cline of grammaticality (Hopper and Traugott 2003:7)  
 content item > grammatical word > clitic > inflectional affix (> Ø)

Grammaticalization is motivated by economy principles. Within a generative perspective, Head Preference Principle by van Gelderen (2004) states that grammaticalization opts for external merge rather than internal merge. In other words, it is a change of a phrase to a head, XP to X<sup>0</sup> (“Be a head”). Upwards Reanalysis (Roberts and Roussou 2003) and Late Merge (van Gelderen 2008) also reflect the same insight. Merge costs less than Move since Move implies Merge. Merging as late/high in the structure as possible is preferred. In this respect, grammaticalization is a change “up the tree” (Roberts and Roussou 2003).

### *1.2 Against Unidirectionality: Degrammaticalization*

Although grammaticalization has been thought to be directional based on economy principles, it has also been argued that this generalization is too strong. Unidirectionality of grammaticalization is only a statistical universal and not an absolute one, hence reversible (Campbell 1991, Ramat 1992, Haspelmath 2004, Traugott 2001, Norde 2009). Such reversal is called degrammaticalization. Just like in grammaticalization, the constructional identity of the degrammaticalized item is preserved within an ambiguous context, which allows for reanalysis (Norde 2009: 8; see Haspelmath 1999, van Gelderen 2011 for grammaticalization). Like other changes, degrammaticalization is also triggered by a change(s) in the same linguistic system.

In this paper, we specifically focus on the “head to phrase” change, i.e., the reversal of the directional “phrase to head” change. We claim that the change of pronominal clitics into weak pronouns in Slavic instantiates degrammaticalization as a X<sup>0</sup> to XP change. Clitics’ dual status as X<sup>0</sup>/XP (Chomsky 1995) provides an adequate condition for the reanalysis.

## 2 Assumptions about Tense and Cliticization

In this section we present our assumptions about cliticization and its relation to the availability of tense morphology. Following Chomsky (1995: 249), we take clitics to be ambiguous categories that share XP and X<sup>0</sup> properties. Verb-adjacent clitics move from argument phrasal positions within VP and adjoin to T<sup>0</sup> as heads. A number of proposals have been put forward in the literature as to the motivation for the clitic T (or I)<sup>0</sup>-adjunction, which is sometimes referred to as Kayne's Law, in relation to Kayne's (1991) suggestion that clitics must attach to the heads of functional categories. For example, Nash and Rouveret (2002) claim that clitics adjoin to T<sup>0</sup>, as they need to be licensed by adjoining to a category endowed with active  $\phi$ -features. Bošković (2012b) argues that verb-adjacent clitics are D+*pro* complexes (as in Jaeggli 1986). *Pro* needs to be licensed; this can be done via verbal morphology, the way the subject *pro* is licensed in null-subject languages. Consequently, the D-clitics in the form of a D+*pro* complex must raise to V+T. What most of these proposals have in common is that they relate cliticization to the availability of T<sup>0</sup>. In this paper, we take the presence of T<sup>0</sup> to be a necessary condition for verb-adjacent cliticization. We also assume (as in Bošković (2012a), Osawa (1999) for Old English, Lin (2010) for Chinese, Todorović (2014) for Serbian) that TP is not a universal projection. Rather, it exists only in the languages with unambiguous morphological tense marking. This means that in Slavic it is only available in Bulgarian and Macedonian (though see section 4 on Macedonian), which are the only Slavic languages that have preserved tense morphology. Thus, as shown in (2), Bulgarian has two designated tenses, the aorist and the imperfect, which are used to unambiguously describe past events. Both of them can be marked for perfective or imperfective aspect.

- (2) a. Včera pročetox knjigata. (aorist, perfective verb)  
 yesterday read<sub>AOR.ISG.PRF</sub> book-the  
 'I read the book yesterday.'
- b. Včera četjax knjigata. (imp, imperfective verb)  
 yesterday read<sub>IMP.ISG.IMPRF</sub> book-the  
 'I was reading the book yesterday.'

The structures presented in (2) are reminiscent of the tense system of Old Church Slavonic. All other Slavic languages have lost the aorist and imperfect and use compound tenses formed with the auxiliary BE and the tenseless, non-finite *l*-participle as the main verb to refer to past events, as exemplified in (3) for Serbo-Croatian.<sup>1</sup> Russian does not use the auxiliary but only uses the *l*-participle to denote the past.

- (3) Ja sam čitao knjigu.  
 I am<sub>AUX</sub> read<sub>PTCP.M.SG</sub> book  
 ‘I (have) read the book.’

Moving back to cliticization properties, a possible follow-up of the assumption that TP is not present in all languages is that verb-adjacent cliticization is only possible in languages with tense. This assumption receives synchronic support within Slavic: verb-adjacent clitics are found only in Bulgarian and Macedonian, the languages with the aorist and imperfect tenses.<sup>2</sup> An example of verb-adjacent cliticization for Bulgarian is given in (4).

- (4) a. Az *im* *ja* preporučvam.  
 I them<sub>DAT</sub> her<sub>ACC</sub> recommend<sub>1SG</sub>  
 ‘I am recommending her to them.’  
 b. [TP Az [T <*im*<sub>DAT</sub>><sub>i</sub> + <*ja*<sub>ACC</sub>><sub>j</sub> + T] ... [VP V t<sub>i</sub> t<sub>j</sub>]]

Migdalski (2013) shows that this idea is also supported by diachronic considerations: in Serbian pronominal clitics shifted from verb-adjacency to second position (2P) and this change was contemporaneous with the loss of morphological tense distinctions, which he analyzes as the loss of TP. Once T<sup>0</sup> is lost, there is no suitable head for pronominal clitics to adjoin to and they raise to separate maximal projections in 2P. This property manifests in many syntactic contrasts between second position

<sup>1</sup> Some dialects of Serbo-Croatian have preserved the aorist to various degrees, but it does not have an unambiguous past tense interpretation any more. See Todorović (this volume).

<sup>2</sup> An anonymous reviewer asks whether this generalization can be extended outside Slavic. We leave this issue for future research, noting that Philippine languages may support our generalization as they have 2P clitics and express tense distinctions via aspectual marking.

and verb-adjacent cliticization observed in the literature, such as the impossibility of partial clitic ellipsis under VP-identity in Bulgarian (see Bošković 2002: 331 and Stjepanović 1998: 530–532 for corresponding Serbian data), the ban on verb-adjacent pronominal clitic splits by a parenthetical (Bošković 2001: 50, 189), or the lack of strong Person Case Constraint effects in languages with 2P clitics (Migdalski 2006 ch. 4). Due to space constraints, we are not able to discuss these contrasts in detail, but they all point to the conclusion that whereas verb-adjacent clitics adjoin to a single head together with the verb, 2P clitics do not form a single syntactic constituent and each of them targets a separate specifier above the VP as an XP-element.

The shift of pronominal clitics to second position is not the only potential repercussion of the loss of  $T^0$  observed in Slavic. This paper shows that it may also lead to the reinterpretation of pronominal clitics as weak pronouns. This part of the paper has the following organization. Section 3 outlines properties of weak pronouns, on the basis of data from Modern Polish. Section 4 investigates the degrammaticalization of clitics into weak pronouns that, as we argue, is currently taking place in some contexts in Macedonian. Sections 5 and 6 turn to diachronic evidence, focusing on Old Polish and Old Russian, respectively.

### 3 Properties of Weak Pronouns

In general, weak pronouns present an intermediate stage between clitics and strong pronouns. Their characteristics are detailed in Cardinaletti and Starke (1999); here we focus on a few of their properties that differentiate them from pronominal clitics, using Polish data. See Rappaport (1988), Witkoś (1998), Franks and King (2000), and Cetnarowska (2003) for a more in-depth analysis.

In comparison to pronominal clitics in South Slavic, weak pronouns are more mobile and typically display characteristics of XP-nominals. For instance, they exhibit more robust scrambling possibilities. As shown in (5a), weak pronouns in Polish may occupy virtually any position in the structure as long as they are not clause-initial. They also avoid the clause-final position, unless a sentence consists of just one other constituent and the final position is the only position available (see 5b).

- (5) a. (\*Go) często (*go*) spotykam (*go*) na ulicy  
 him<sub>ACC</sub> often him<sub>ACC</sub> meet<sub>PRES.1SG</sub> him<sub>ACC</sub> on street  
 ‘I often meet him on the street.’  
 b. Spotykam *go*  
 ‘I meet him.’ (see Spencer 1991: 367–368)

Furthermore, unlike pronominal clitics in South Slavic, where the dative clitic must precede the accusative clitic, weak pronouns permit both DAT-ACC and ACC-DAT orders (see 6), though the former may be more common. Cetnarowska (2003) suggests that the choice of a particular order depends on information structure requirements.

- (6) Jan w końcu *go jej /jej go/* oddał.  
 Jan in end it<sub>ACC</sub> her<sub>DAT</sub> her<sub>DAT</sub> it<sub>ACC</sub> lend<sub>PTCP.M.SG</sub>  
 ‘Jan eventually returned it to her.’

Furthermore, they do not need to be adjacent either to an element of a specific category (such as a verb) or to other pronominal forms, see (7).

- (7) Jan *mu* wczoraj chciał *go* wypożyczyć  
 Jan him<sub>DAT</sub> yesterday wanted it<sub>ACC</sub> lend<sub>INF</sub>  
 a nie sprzedać.  
 and not sell<sub>INF</sub>  
 ‘Jan wanted to lend it to him rather than sell it yesterday.’

#### 4 Degrammaticalization of Clitics into Weak Pronouns in Macedonian

In section 2, example (2), we show that Bulgarian allows the aorist and imperfect tenses to be combined with both perfective and imperfective aspect forms. Although Macedonian has also retained both past tenses, its tense system is more restricted: the aorist is the default past tense for perfective verbs, whereas the imperfect is the default past tense for imperfective verbs. This is a recent impoverishment of the tense system. Friedman (2002: 267) states that until the middle of the 20<sup>th</sup> century, imperfective verbs in the aorist were possible in Macedonian.

We assume the reduction of the tense system in Macedonian to be a significant fact, as it coincides with the modification of its cliticization

strategies. Although Macedonian has verb-adjacent clitics, when they occur in non-verbal predicates (APs, NPs, and passive participles), they are in general found in second position (see 8a). Interestingly, Korubin (1974), Tomić (1997, 2000) and Baerman and Billings (1998) observe that recently some speakers of Macedonian started to permit clitics clause-initially in the contexts of adjectival predicates and passive participles (see 9b and 10a). Furthermore, some speakers also allow clitics to be located below second position (see 9c and 10c) in these contexts. We would like to propose that this means that the clitics in non-verbal predicates are being reinterpreted as weak pronouns. Significantly, this process is accompanied by a recent impoverishment of tense distinctions in Macedonian. In line with the assumptions made in this paper, this fact leads us to suggest that the modification of the Macedonian cliticization pattern is due to a (gradual) loss of  $T^0$ , which precludes head-adjunction of clitics. As a result, they become reinterpreted as weak pronouns.

- (8) a. \* Petko tatko *mi e*.  
 Petko father me<sub>DAT</sub> is  
 b. \* *Mi e* tatko.  
 me<sub>DAT</sub> is father  
 c. Petko *mi e* tatko.  
 Petko me<sub>DAT</sub> is father  
 ‘Petko is my father.’ (Tomić 2000: 295; Bošković 2001: 255)
- (9) a. Mil *si mu*.  
 dear<sub>M.SG</sub> be<sub>2SG</sub> him<sub>DAT</sub>  
 ‘He likes you.’  
 b. %*Si mu* mil.  
 c. Petko sekogaš *mi e* mil.  
 Petko always me<sub>DAT</sub> be<sub>3SG</sub> dear<sub>M.SG</sub>  
 ‘Petko is always dear to me.’ (cf. Franks and King 2000: 86)
- (10) a. %*Mu e* rečeno da bide točen poveke pati.  
 him<sub>DAT</sub> is tell<sub>PASS</sub> to be<sub>SUBJ</sub> punctual more times  
 ‘He was told to be punctual more than once.’

- b. Rečeno *mu e* da bide točen poveќе pati.
- c. Na Petreta (*mu e*) od strana na komisijata  
 to Peter<sub>DAT</sub> him<sub>DAT</sub> is from side of commission-the  
 (*mu e*) poveќе pati (*mu e*) rečeno da bide točen.  
 him is more times him<sub>DAT</sub> is tell<sub>PASS</sub> to be punctual  
 ‘Peter was more than once told by the commission to be  
 punctual.’ (Tomić 2000: 296–299)

The data presented in (8-10) have not received enough attention in the literature and often have been treated as somewhat exceptional. For instance, Bošković (2001: 254–264) states that Macedonian may represent an intermediate stage between languages with verb-adjacent and 2P clitics. However, the idea that the change taking place in Macedonian is related to the syntactic position of the clitics is problematic, given that the Bulgarian variant of (9c), with a pronominal clitic located lower than second position, is acceptable. It is more likely that the modification is related to the reinterpretation of clitics as XP-elements, thus, their strengthening, which gives rise to more robust scrambling possibilities. Crucially, the Bulgarian counterpart of (10c) is ungrammatical: as shown in (11), Bulgarian requires the pronominal forms to be adjacent to the passive participle (Vesela Simeonova, p.c.).

- (11) Na Petūr *mu e* kazvano mnogo pūti ot strana  
 to Peter him<sub>DAT</sub> is tell<sub>PASS.N</sub> many times from side  
 na komisijata da būde točen.  
 of commission-the to be<sub>SUBJ.3SG</sub> punctual  
 ‘P. was told by the commission to be punctual many times.’

## 5 Old Polish

This section investigates the syntactic properties of pronominal elements in Old Polish. It analyzes three texts from different periods. A cursory examination of the data shows that the process of clitic degrammaticalization that is currently occurring in Macedonian was completed in Old Polish and was accompanied by the loss of tense morphology.

We begin the examination with *Holy Cross Sermons (Kazania Świętokrzyskie)*, the oldest Polish prose text from the late 13<sup>th</sup>/early 14<sup>th</sup> century. The style of the sermons is very formal, as they were aimed at

the educated public, whereas its grammar is rather archaic, as it still contains verb forms in the aorist and the imperfect (see bolded in 12), which are virtually not found in later texts. Regardless, the aorist and imperfect forms are less common than the compound tense constructed with the *l*-participle and the auxiliary BE, which eventually replaces the simple forms altogether in the later stages of language history.

- (12) a. **jemuż** **biesze** imię Symeon, święty, prawdziwy.  
 him<sub>DAT</sub>+FOC be<sub>IMP.3SG</sub> name Simon holy true  
 ‘His name was Simon, holy, true.’  
 (Sermon III, *On St. Michael’s Day*)
- b. **pośpieszyczą** **się** do kościoła na modlitwę przed  
 hurry<sub>AOR.3PL</sub> REFL to church to prayer because  
 Boga wszemogącego i **poczęcą** **się** modlić.  
 God Almighty and start<sub>AOR.3PL</sub> REFL pray<sub>INF</sub>  
 ‘They hurried to church for a prayer to God Almighty, and they started to pray.’  
 (Sermon VI, *The Cleansing Of The Blessed Virgin Mary*)

In this text, pronominal clitics tend to appear in 2P or are verb-adjacent (see 13). However, the fact that they may be introduced by a preposition (see bolded in 14) suggests that they are becoming reanalyzed as strong elements. In Modern Slavic, clitics may not occur as complements of a preposition; rather, a strong form is required in such contexts. We take this fact to mean that although their phonological make-up was that of a clitic, they were already morphosyntactically interpreted as non-clitic elements.

- (13) a. a **togodla** **ji** we złe chustki ogarnęła.  
 and therefore him<sub>ACC</sub> in bad cloth wrap<sub>PTCP.F.SG</sub>  
 ‘and therefore she wrapped him in bad cloth.’  
 (Sermon III, *On St. Michael’s Day*)
- b. Należli **ji**, **prawi**, pieluszkami ogarnienego.  
 find<sub>PTCP.M.PL</sub> him<sub>ACC</sub> true nappies<sub>INST</sub> wrapped  
 ‘They found him wrapped in nappies.’  
 (Sermon III, *On St. Michael’s Day*)

- (14) Sam, prawi, **przez** *mię* przysiągł jeśm.  
 He say<sub>AOR.3SG</sub> without me<sub>ACC</sub> swear<sub>PTCP.M.SG</sub> be<sub>1SG</sub>  
 ‘He said that he has sworn without me...’  
 (Sermon III, *On St. Michael’s Day*)

The slightly more recent texts we have consulted, such as *Queen Sophia’s Bible/Sárospatak Bible (Biblia Królowej Zofii)* from 1433–1455 and *Gniezno Sermons (Kazania Gnieźnieńskie)* from the early 15<sup>th</sup> century, exhibit very few simple past tense forms and instead they employ the compound tense constructed with the auxiliary BE and the *l*-participle as the default past tense. Pronominal elements are found more frequently than in *Holy Cross Sermons*, especially in *Gniezno Sermons*. In fact, strong forms of pronouns are increasingly more common (see bolded in 15) and they often co-occur with weak/clitic forms within the same pragmatic or semantic contexts.

- (15) Tegdy wziął Pan Bog człowieka i postawił *ji*  
 then took God man<sub>ACC</sub> and put him<sub>ACC</sub>  
 w raju rozkoszy, aby działał a  
 in paradise bliss<sub>GEN</sub> so-that worked and  
 ostrzegał **jego**. I przykazał **jemu**...  
 protected him<sub>ACC</sub> and commanded him<sub>DAT</sub>  
 ‘The Lord God took the man and put him in the Garden of Eden to work it and take care of it. And the Lord God commanded him...’  
 (*Queen Sophia’s Bible, Genesis, 2,15-16*)

The placement of the pronominal forms in (15) suggests that clitics receive the same morphosyntactic interpretation as strong pronouns, as both of them occur post-verbally. Correspondingly, the clitics in (16) show a remarkable freedom of distribution, given that they do not need to be in 2P or verb-adjacent.

- (16) a. I przywiódł *je* przed Adama, aby  
 and brought them<sub>ACC</sub> before Adam so-that  
*je* opatrzył a *jimiona jim* dał.  
 them<sub>ACC</sub> saw and names them<sub>DAT</sub> give<sub>PTCP.M.SG</sub>  
 ‘He brought them to the man to see what he would name them.’  
 (*Queen Sophia’s Bible, Genesis, 2, 19*)

- b. Nazwał jest Adam jimiona *jich*  
 name<sub>PTCP.M.SG</sub> is<sub>AUX</sub> Adam names them<sub>GEN</sub>  
 wszelikiemu stworzeniu zwierzęcemu.  
 all beings animal  
 ‘Adam gave names to all the livestock.’ (Genesis, 2, 20)

A conspicuous characteristic of some of the *Gniezno Sermons* is a frequent use of ethical datives (*ci* in 17), which regularly target second position, after the first word. In this way they share the distribution of ethical datives in other Slavic languages (see Bošković 2001: 60–61 for Serbo-Croatian). By contrast, argumental pronouns (such as *je* in 17b) do not appear in a specific position in the clause structure.

- (17) a. *tenci się jest był w łonie u swe*  
 this<sub>+DAT</sub> REFL is<sub>AUX</sub> be<sub>PART</sub> in womb at his  
 miły matuchny panny Maryje.  
 kind mother virgin Mary  
 ‘Who had been in the womb of his kind mother Virgin Mary.’
- b. *cożci jest je przezeń był nasz*  
 what<sub>+DAT</sub> is<sub>AUX</sub> them<sub>ACC</sub> because be<sub>PTCP.M.SG</sub> our  
 miły Kryst czynił drzewie.  
 kind Christ made earlier  
 ‘that because of them our kind Christ had made earlier.’

Summarizing, our cursory survey of the three Old Polish texts indicates that the decline of tense morphology in Old Polish is accompanied by the strengthening of pronominal forms (understood here as degrammaticalization), which become syntactically more mobile and start to occur in virtually the same syntactic contexts as strong pronouns.

## 6 Old Russian

### 6.1 The Old Russian Clitic System (the 11<sup>th</sup>–15<sup>th</sup> cc.)

The clitic system of Old Russian during the 11<sup>th</sup>–15<sup>th</sup> centuries is characterized by 2P placement. In the earliest Old Russian manuscripts (the 11<sup>th</sup>–12<sup>th</sup> cc.), pronominal clitics are distributed in the second position of an intonational phrase as a cluster in the order DAT-ACC, as exemplified in (18).

- (18) a. ože *mi*    *sę*        jeste    jali        pomagati...  
 as me<sub>DAT</sub> REFL<sub>ACC</sub> are<sub>AUX.2PL</sub> take<sub>PTCP.PL</sub> help<sub>INF</sub>  
 ‘Since you undertook to help me...’  
 (*Hypatian Chronicle* 1149, 140, Zaliznjak 2008: 35)
- b. poklanęju    *ti*        *sę*.  
 bow            you<sub>DAT</sub>      REFL<sub>ACC</sub>  
 ‘I bow to you.’  
 (*BBL* No. 907, 11<sup>th</sup>–12<sup>th</sup> cc., Zaliznjak 2004: 255)

However, distributions as weak pronouns are also observed from the earliest time. First, an ACC clitic appears as the object of a preposition.

- (19) **za** *tę*        golovy    svoi        sękladyvaëmь.  
 for you<sub>ACC</sub> head<sub>ACC.PL</sub> own<sub>ACC.PL</sub> lay down<sub>1PL</sub>  
 ‘We bow down to you.’  
 (*Hypatian Chronicle* 1177, Zaliznjak 2008: 36)

Zaliznjak (2008: 36) contends that pronominal clitics combined with prepositions are residues from the preceding, prosodically independent stage. However, a hypothesis that *za tę*<sub>ACC</sub> was an intermediate stage of the change from clitic to pronoun is just as possible as the opposite direction, as illustrated in (20). Clitics appearing in the PP can be the first signal of the weakening of pronominal clitics’ cliticness.

- (20) P + weak pronoun ↔ *za tę*<sub>ACC</sub> ↔ P + clitic

Janin and Zaliznjak (1993: 289) observe that proclitics could function as hosts for enclitics (ex. *ne li jesi dalь*). Prepositions are proclitics, and the combination of a preposition and a weak pronoun/clitic constitutes a prosodically independent phonetic unit. It is difficult to determine whether in the string *za tę*<sub>ACC</sub> the morphologically reduced ACC form was prosodically dependent or not, but it is clear that the ACC form occupied an argument position as the object of the preposition. This ambiguity provides an adequate condition for the reanalysis of the reduced form as a prosodically independent element.

Another piece of evidence for the weak pronoun status is that pronominal clitics sometimes appear in non-second positions. In (21) the ACC reflexive clitic *sę* occupies the initial position, given that the

conjunction *a* cannot function as a host.

- (21) *a*    *se*            *ego*    *zapritʹ*.  
 and REFL<sub>ACC</sub> him<sub>ACC</sub> shut<sub>3SG</sub>  
 ‘And ... locks himself up.’<sup>3</sup>  
 (*Gramoty Velikogo Novgoroda i Pskova*, No. 28, 1190-)

Third, the cluster ordering rule was violated. In Old Russian, the relative order of pronominal clitics and AUX was DAT-ACC-AUX<sub>1/2P</sub>, which sharply contrasts Old Russian with modern South Slavic languages, the latter showing the order AUX<sub>1/2/3P</sub>-DAT-ACC-AUX<sub>3S</sub>. In example (22) the accusative reflexive clitic appears right to AUX. This shows that the ordering rule began to be violated in 13<sup>th</sup>-century Old Russian.<sup>4, 5</sup>

- (22) *i*    *jela*            *jesmo*    *se*            *jemu*    *po ruku*  
 and take<sub>PTCP.F.SG</sub> am<sub>AUX.1SG</sub> REFL<sub>ACC</sub> him<sub>DAT</sub> for hand  
 ‘I promised him...’  
 (*BBL* No. 731, early 13<sup>th</sup> c., *Zaliznjak* 2004: 392)

The concrete diachronic process of the loss of pronominal clitics can be formulated based on the data from Birchbark letters (*BBL*) in Old North Russian from the 11<sup>th</sup>-15<sup>th</sup> centuries in (23-25). During this period, the 2P pattern was coexistent with such deviating patterns as in (23-25), the former declining toward the 15<sup>th</sup> century.

<sup>3</sup> The interpretation of this data is not clear due to the co-existence of the ACC reflexive clitic and the ACC pronoun. However, the clitic's position clearly deviates from the 2P pattern.

<sup>4</sup> *Zaliznjak* (2004:393) assumes that this example shows a micro-dialectal tendency to shift the clitic ordering DAT-ACC-AUX to AUX-DAT-ACC, which modern South Slavic languages have.

<sup>5</sup> A reviewer asks whether the ordering rule only applies to ‘clitic’ clusters, as the AUX in (22) is not marked as clitic (not italicized). *Zaliznjak* (2004) treats AUXs in Birchbark letters as clitics while the opposite has also been argued for *Jung* (2013). What matters here is that regardless of the prosodic status of AUX, the relative order between AUX and pronominal clitics was consistently DAT-ACC-AUX<sub>1/2P</sub> in Old Russian. So, the order AUX<sub>1P</sub>-ACC<sub>CL</sub> in (22) clearly deviates from the regular Old Russian pattern.

- (23) **12<sup>th</sup> c.-early 13<sup>th</sup> c.: pronominal clitics after prepositions**  
 postrъčътъ užь na mę i na moe deti  
 provoke<sub>3SG</sub> already against me<sub>ACC</sub> and against my children  
 ‘(He) then provokes against me and against my children.’  
 (BBL No. 831, mid-12<sup>th</sup> c., Zaliznjak 2004: 302)
- (24) **late 13<sup>th</sup> c.-15<sup>th</sup> c.: clitics in non-second position**  
 a. a na koni prišili mi v grivni serebra.  
 and for horses send<sub>IMV</sub> me<sub>DAT</sub> two grivna silver  
 ‘And for horses send me 2 grivnas of silver.’  
 (BBL No. 775, late 13<sup>th</sup> c., Ibid. 502)
- b. tako prišli mi colověkъ  
 thus send<sub>IMV</sub> me<sub>DAT</sub> man<sub>ACC</sub>  
 ‘Thus send me a man.’ (BBL No. 43, late 14<sup>th</sup> c., Ibid. 651)
- (25) **late 14<sup>th</sup> c.: clitics mixed with full forms**  
 a jęzo tobě sę klanęju.  
 and I<sub>NOM</sub> you<sub>DAT</sub> REFL<sub>ACC</sub> bow<sub>1SG</sub>  
 ‘And I bow to you.’ (BBL No. 186, late 14<sup>th</sup> c., Ibid. 618)

The sentence in (25) was formulated based on the template *jęzo ti<sub>DAT</sub> sę<sub>REFL.ACC</sub> klanęju*, from which the clitic *ti* was replaced by the full form *tobě* (Zaliznjak 2004: 618). Here, the ACC clitic *sę* is not in second position. This example indicates that the distribution rules of pronominal clitics are no longer functional and that the use of pronominal clitics became no more than conventional.

From the diachronic process of losing 2P cliticization, as reflected in (23–25), we can conclude that a 2P system coexisted with and gave way to a weak pronoun system. When the reduced ACC form after a preposition, which was ambiguous in terms of its prosodic independence, was reanalyzed as a weak pronoun, the string <# verb + *tę<sub>CL</sub>*> must also have been reanalyzed as <# verb + *tę<sub>WP</sub>*>.

It has been reported that the ACC reflexive clitic *sę* tended to follow verbs more often than other enclitics (Janin and Zaliznjak 1993: 169, Zaliznjak 2008: 292). We take this phenomenon as an indication that the clitic *sę* came to occupy the object DP position as a weak pronoun. The loss of prosodic dependence of reduced forms and the shift from clitic to weak pronoun help us to understand why pronominal clitics, unlike operator clitics (e.g. *li*, *že*), disappeared from Old Russian. Through the change of pronominal clitics to weak pronouns, the dual

system of full and clitic forms became unmotivated, and full forms were generalized in every position.

### 6.2 *The Loss of TP in Old Russian*

We assume that this Old Russian clitic system change was, as in Macedonian and Old Polish, triggered by the loss of TP. In spoken Old Russian, the imperfect and the aorist were out of use at the latest in the 12<sup>th</sup> century (Issatschenko 1983: 355-356, Uspenskij 1987: 144-151). This indicates that TP was lost before the 12<sup>th</sup> century.

In Old Russian manuscripts from the 11<sup>th</sup> century on, the stage preceding the 2P system is not attested, while OCS features verb-adjacent clitics. If we assume that Slavic dialects maintained a high degree of homogeneity in the period of the OCS canon (the 9<sup>th</sup> c.), we can posit a verb-adjacent system for the prehistoric period of East Slavic.

Due to the early loss of TP in Old Russian, pronominal clitics could not raise and adjoin to T<sup>0</sup> as X<sup>0</sup>s but remained as XPs in argument positions, either resulting in 2P clitics or turning themselves into weak pronouns. The 2P system in Old Russian became increasingly unstable and clitics gradually disappeared from written materials until the 15<sup>th</sup> century.

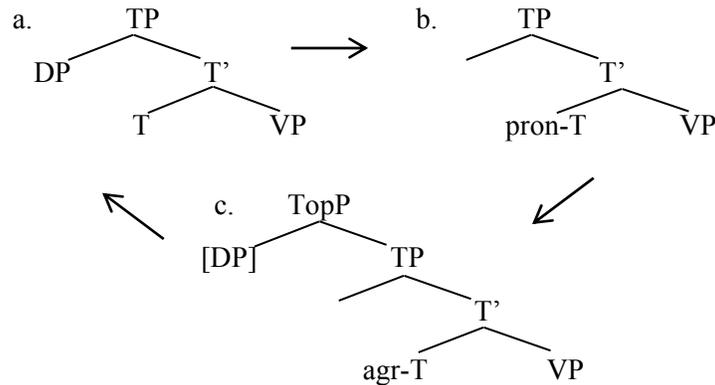
### 6.3 *The I-Perfect Auxiliary as a Subject Pronoun*

In addition to pronominal clitics, the present tense forms of the *I*-perfect auxiliary, such as *jesm'* and *jesi*, also show the status change from head to phrase in Old Russian. The inflected forms of the perfect auxiliary may be analyzed as pronominal because they purely represent person feature, being referential. They originally merged in T<sup>0</sup>, but were reanalyzed as pronoun subjects located in Spec,AGR<sub>S</sub>P, and came to trigger agreement on finite verbs, as shown in (26). In (26a), *jesm'*, the 1<sup>st</sup> SG form of AUX, appears in the presence of the finite verb *xosču*, triggering *phi*-agreement on it, which is nominative subject's typical behavior. The same is observed in (26b), where *jesi*, the 2<sup>nd</sup> SG form of AUX, controls agreement on the finite verb *budeš*. In (26c), *jesmi* triggers agreement on the finite verb *znaju*. In these examples, the 1<sup>st</sup> and 2<sup>nd</sup> person AUX forms behave like subject pronouns, such as *ja* and *ty*, respectively. In (26b) and (26c), the auxiliary forms are higher than the negation, which is ungrammatical in Modern Russian. This also indirectly indicates that the auxiliary verb is at the IP level.

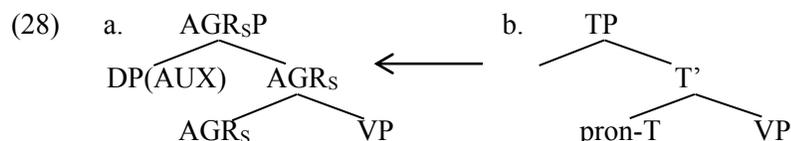
- (26) a. povestouju, čto jesmь nyně na dorogu jexati xoščü.  
 tell<sub>1SG</sub> comp be<sub>1SG</sub> now on road go<sub>INF</sub> want<sub>1SG</sub>  
 ‘I tell you that I will leave now.’  
 (*Pskovian chronicle III*, 1473, Zaliznjak 2004: 179)
- b. a ženy ne vidělъ jesi budešъ vъ sně.  
 and woman<sub>GEN</sub> NEG see<sub>PTCP.M.SG</sub> be<sub>2SG</sub> be<sub>FUT.2SG</sub> in dream  
 ‘And you will not have seen a woman in a dream.’  
 (*Kirik’s Queries*, mid-12<sup>th</sup> c., Ibid.)
- c. a togo žь jesmi ne znaju, u kogo kupilъ.  
 and that FOC be<sub>1SG</sub> NEG know<sub>1SG</sub> from whom buy<sub>PTCP.M.SG</sub>  
 ‘And I don’t know from whom I bought.’ (*PTL*, 1473, Ibid.)

This change of the present tense AUX instantiates the reverse of van Gelderen’s Subject Agreement Cycle (2011) in (27) and this reversed cycle constitutes X<sup>0</sup> to XP degrammaticalization.

- (27) From van Gelderen (2011: 42, Figure 2.1)



In Old Russian, the stage in (27b), in which the pronominal auxiliary (with D-feature and person feature) merges in T<sup>0</sup> as a head, reverted to a (27a)-like structure, as illustrated in (28). Structure (28a) is almost the same as (27a) but contains AGR<sub>S</sub>P instead of TP since Old Russian lost TP. In (28a) the pronominal auxiliary occupies Spec,AGR<sub>S</sub>P as a subject DP, replacing the null subject (Jung 2013).



## 7 Conclusion

Thus far, we have shown that the Macedonian, Old Polish, and Old Russian data provide empirical evidence against the idea of the irreversibility of grammaticalization, the historical directionality of the pronoun weakening cycle (pronoun>clitic>verbal affix), and universal directionality of language change.

The degrammaticalization of pronominal clitics into weak pronouns was triggered by the loss of TP, which is indicated by the loss of inflected tense categories such as the aorist and the imperfect. Verb-adjacent clitics first shifted to 2P clitics and then were subject to the weakening of their cliticness. This process can be traced on the basis of changes in their distributional patterns.

## References

- Baerman, Matthew, and Loren Billings. 1998. Macedonian clitics and the trisyllabic stress window. *University of Trondheim Working Papers in Linguistics* 31: 13–32.
- Bošković, Željko. 2001. *On the Nature of the Syntax–Phonology Interface. Cliticization and Related Phenomena*. Amsterdam: Elsevier.
- Bošković, Željko. 2002. Clitics as non-branching elements and the linear Correspondence Axiom. *Linguistic Inquiry* 33: 329–340.
- Bošković, Željko. 2012a. On NPs and clauses. In *Discourse and grammar: From sentence types to lexical categories*, G. Grewendorf and T. E. Zimmermann (eds), 179–242. Boston: De Gruyter Mouton.
- Bošković, Željko. 2012b. On clitic doubling and second position cliticization: how much does D matter? Paper presented at *Clitics and Beyond*, 3–5 May 2012, University of Göttingen.
- Cardinaletti, Anna and Michal Starke. 1999. The typology of structural deficiency: a case study of the three classes of pronouns. In *Clitics in*

- the Languages of Europe*, H. van Riemsdijk (ed.), 185–234. Berlin: Mouton de Gruyter.
- Campbell, Lyle. 1991. Some grammaticalization changes in Estonian and their implications. In *Approaches to Grammaticalization I*, E. Traugott and B. Heine (eds), 285–299. Amsterdam/Philadelphia: John Benjamins.
- Cetnarowska, Bożena. 2003. On pronominal clusters in Polish. In *Investigations into Formal Slavic Linguistics: Contributions of the Fourth European Conference on Formal Description of Slavic Languages - FDSL IV held at Potsdam University, November 28-30, 2001*, P. Kosta, J. Błaszczak, J. Frasek, L. Geist, and M. Zygis (eds), 13–30. Frankfurt am Main: Peter Lang.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: The MIT Press.
- Franks, Steven, and Tracy Holloway King. 2000. *A Handbook of Slavic Clitics*. New York: Oxford University Press.
- Friedman, Victor, A. 2002. Macedonian. In *The Slavonic Languages*, B. Comrie and G. G. Corbett (eds), 249–305. London: Routledge.
- Gelderen, Elly van. 2004. *Grammaticalization as Economy*. Amsterdam/Philadelphia: John Benjamins.
- Gelderen, Elly van. 2008. Where did late merge go? Grammaticalization as feature economy. *Studia Linguistica* 62(3): 287–300.
- Gelderen, Elly van. 2011. *The Linguistic Cycle: Language Change and the Language Faculty*. Oxford: Oxford University Press.
- Haspelmath, Martin. 1999. Why is grammaticalization irreversible? *Linguistics* 37(6): 1043–1068.
- Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization. In *Up and Down the Cline: The Nature of Grammaticalization*, O. Fischer, M. Norde, and H. Perridon (eds), 17–44. Amsterdam/Philadelphia: John Benjamins.
- Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.
- Hopper, Paul and Elizabeth Traugott 1993/2003. *Grammaticalization*. 2<sup>nd</sup> edition. Cambridge: Cambridge University Press.
- Issatschenko, A.V. 1983. *Geschichte der russischen Sprache*. Vol. 2. H. Birnbaum, L. Āurovič, and E. Salnikow-Ritter (eds.). Heidelberg: Winter.

- Janin, V. L. and A. A. Zaliznjak. 1993. *Novgorodskie gramoty na bereste: iz raskopok 1984-1989 gg.* Moscow: Nauka.
- Jaeggli, Osvaldo. 1986. Three issues in the theory of clitics: case, doubled NPs, and extraction. In *Syntax and Semantics 19, The Syntax of Pronominal Clitics*, H. Borer (ed.), 15–42. Orlando, FL: Academic Press.
- Jung, Hakyung. 2013. On the syntax of the so-called auxiliary clitic in Old North Russian. Paper presented at Formal Description of Slavic Languages 10, University of Leipzig, December 6, 2013.
- Kayne, Richard S. 1991. Romance clitics, verb movement, and PRO. *Linguistic Inquiry* 22: 647–686.
- Korubin, Blagoja. 1974. Mesto glagola sum kak vspomogatel'nogo i kak svjazki v makedonskom literaturnom jazyke.” In *Grammatičesko opisane slavjanskix jazykov: koncepcii i metody*, N. Švedova et al. (eds), 244–50. Moskva: Nauka.
- Kuryłowicz, Jerzy. 1965/1975. The evolution of grammatical categories. In *Esquisses Linguistiques II*, 38-54. Munich: Wilhelm Fink Verlag.
- Kuteva, Tania. 2001. *Auxiliation: An Enquiry into the Nature of Grammaticalization*. Oxford: Oxford University Press.
- Lehmann, Christian. 1982/1995. *Thoughts on Grammaticalization*. Munich/Newcastle: Lincom Europa.
- Lin, Jo-wang. 2010. A tenseless analysis of Mandarin Chinese revisited: a response to Sybesma 2007. *Linguistic Inquiry* 41(2): 305–329.
- Meillet, Antoine. 1912/1926. L'évolution des formes grammaticales. In *Linguistique Historique et Linguistique Générale*, I, 130–148. Paris: Edouard Champion.
- Migdalski, Krzysztof. 2006. *The Syntax of Compound Tenses in Slavic*. Ph.D. dissertation, Tilburg University. Downloadable at: <http://www.lotpublications.nl>
- Migdalski, Krzysztof. 2013. Diachronic source of two cliticization patterns in Slavic. In *Challenging Clitics*, C. M. Salvesen and H. P. Helland (eds), 135–158. Amsterdam/Philadelphia: John Benjamins.
- Nash, Léa and Alain Rouveret. 2002. Cliticization as unselective Attract. *Catalan Journal of Linguistics* 1: 157–199.
- Norde, Muriel. 2009. *Degrammaticalization*. Oxford: Oxford University Press.

- Osawa, Fuyo. 1999. The relation between tense and aspect: the emergence of the T system. *UCL Working Papers in Linguistics* 11: 521–544.
- Ramat, Paolo. 1992. Thoughts on degrammaticalization. *Linguistics* 30: 549–560.
- Rappaport, Gilbert. 1988. On the relationship between prosodic and syntactic properties of pronouns in the Slavic languages. In *American Contributions to the Tenth International Congress of Slavists*, A. M. Schenker (ed.), 301–327. Bloomington: Slavica.
- Roberts, Ian and Anna Roussou. 2003. *Syntactic Change*. Cambridge: Cambridge University Press.
- Spencer, Andrew. 1991. *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Oxford: Blackwell.
- Stjepanović, Sandra. 1998. On the placement of Serbo-Croatian clitics: evidence from VP ellipsis. *Linguistic Inquiry* 29: 527–537.
- Todorović, Neda. 2014. Aspectual tenses in Serbian and parallelism with the nominal domain. Paper presented at FASL 23, University of California Berkeley, May 2-4, 2014.
- Tomić, Olga. 1997. Non-first as a default clitic position. *Journal of Slavic Linguistics* 5: 301–323.
- Tomić, Olga. 2000. On clitic sites. In *Clitic Phenomena in European Languages*, F. Beukema and M. den Dikken (eds), 293–316. Amsterdam: Benjamins.
- Traugott, Elizabeth. 2001. Legitimate counterexamples to unidirectionality. Paper presented at Freiburg University, October 17, 2001.
- Uspenskij, B. A. 1987. *Istorija ruskogo literaturnogo jazyka (XI–XII vv.)*. München: O. Sagner.
- Witkoś, Jacek. 1998. *The Syntax of Clitics. Steps Towards a Minimalist Account*. Poznań: Motivex.
- Zaliznjak, A. A. 2004. *Drevnenovgorodskij dialect*. Moscow: Jazyki slavjanskoj kul'tury.
- Zaliznjak, A. A. 2008. *Drevnerusskie ènklitiki*. Moscow: Jazyki slavjanskoj kul'tury.

hakyungj@snu.ac.kr  
migdalski@yahoo.com

## **The Effect of Prosody on Availability of Inverse Scope in Russian**

*Tatiana Luchkina*

University of Illinois at Urbana-Champaign

*Tania Ionin*

University of Illinois at Urbana-Champaign

This paper examines the interpretation of scopally ambiguous double-quantifier sentences in Russian, a free word order language. We describe the acoustic-prosodic features that successfully discriminate between the productions disambiguating double-quantifier sentences in Russian in favor of surface vs. inverse scope. We argue that the prosodic contour under which inverse scope is available in Russian is associated with a highly salient information structural configuration in which the preverbal QP is interpreted as focused.

### **1 Introduction**

This paper examines the interpretation of double-quantifier sentences in Russian, a free word order language. We experimentally assess the role of information structure (IS), prosody, and word order on the availability of inverse scope in Russian. In languages like English, double-quantifier sentences like (1) give rise to more than one reading and are characterized as scopally ambiguous. Reversing the order of the quantifiers reverses what constitutes surface vs. inverse scope, as shown in (2).

- (1) One dog frightened every man.
- a. surface scope: One specific dog frightened every man (ONE>EVERY)
  - b. inverse scope: Every man was frightened by a potentially different dog (EVERY>ONE)
- (2) Every dog frightened one man.
- a. surface scope: Every dog frightened a potentially different man (EVERY>ONE)
  - b. inverse scope: One specific man was frightened by every dog (ONE>EVERY)

The two possible readings of (1) illustrated in (1a) and (1b) arise due to one of the two quantifiers, indefinite or universal, taking wide scope. The surface-scope reading in (1a) arises because the subject quantifier phrase (QP) ‘one dog’ takes scope over the object QP ‘every man’. The inverse-scope reading in (1b) arises when the universal object QP scopes over the indefinite subject QP. The inverse scope reading in (1b) has been shown to be more difficult to process than the surface scope reading (Tunstall 1998, Anderson 2004). One way to account for the processing difficulty (e.g., Anderson 2004) is to tie it to the processing costs incurred by covert quantifier raising (QR), which takes place at LF and raises the object QP to a position higher than the subject QP (May 1985, Fox 2000).

English is one language which exhibits covert QR as the means for deriving inverse scope. However, not all languages do QR covertly. Languages such as German, Greek, and Russian, among others, allow for variable word orders. Relative word order freedom makes overt movement of the sentence constituents, including QPs, possible, as shown in the Russian translations of (1) and (2) given in (3) and (4) below.

- (3) a. *SVO, canonical word order:*  
 Odna sobaka napugala každogo čeloveka.  
 one<sub>NOM</sub> dog<sub>NOM</sub> frightened every<sub>ACC</sub> man<sub>ACC</sub>
- b. *OVS, scrambled word order:*  
 Každogo čeloveka napugala odna sobaka.  
 every<sub>ACC</sub> man<sub>ACC</sub> frightened one<sub>NOM</sub> dog<sub>NOM</sub>

- (4) a. *SVO, canonical word order:*  
 Každaja sobaka napugala odnogo čeloveka.  
 every<sub>NOM</sub> dog<sub>NOM</sub> frightened one<sub>ACC</sub> man<sub>ACC</sub>
- b. *OVS, scrambled word order:*  
 Odnogo čeloveka napugala každaja sobaka.  
 One<sub>ACC</sub> man<sub>ACC</sub> frightened every<sub>NOM</sub> dog<sub>NOM</sub>

Word order flexibility allows Russian to alter the order of subject and object QPs overtly, reversing what constitutes surface vs. inverse scope. Corroborating this expectation is the fact that a change in word order reverses quantifier scope in other ‘free word order’ languages (e.g., German: Beck 1996, Sæbø’ 1997, Bobaljik & Wurmbrand 2012; Greek: Baltazani 2002; Japanese: Miyagawa 1997), where overt constituent movement is motivated by grammatical features, such as object scrambling for case agreement (Japanese) or preferred information structural configuration (Greek).

In Russian, canonically an SVO language (Bailyn 1995), word order variability is discourse-constrained (Slioussar 2011) and is used to encode the categories of given vs. novel information, as well as to signal especially prominent information in discourse, i.e., words which are contrastively focused or emphasized (Neeleman & Titov 2009). To illustrate, in (3b), the direct object ‘every man’ appears pre-verbally, whereas the subject ‘one dog’ is sentence-final. It is traditionally believed that the sentence-final position in Russian is reserved for discourse-novel information (new information focus), and that it is felicitous for old or given information (Topic) to precede the novel information (Calhoun 2010). Thus, one may view scrambling in Russian as a means for a constituent to appear outside the ‘nuclear focus domain’ (Rosengren 1993) aligned with the right sentence periphery and reserved for discourse-novel information, and to form a Topic domain or a separate (contrastive) focus domain. In accordance with the preferred IS configuration, in (3b) it is plausible to associate the object phrase ‘every man’ with given, previously established information or sentence Topic (the preceding discourse must be about men). The subject phrase ‘one dog’ is then perceived as discourse-novel and presents the new information focus.

Ionin (2003) analyzed scope in Russian double-quantifier sentences which exhibit scrambling driven by topicalization under neutral (non-

contrastive) prosody. Ionin proposed that in Russian, the close tie between word order and discourse function is what constrains the availability of inverse scope in sentences such as (3) and (4), for the following reasons:

1. In non-emotive matrix clauses in Russian, the leftmost position is reserved for the Topic element associated with the previously established information in discourse.
2. Since Topics are established entities, they have previous mention in discourse and must be interpreted first. This precludes scope reversal, which requires that the sentence-final QP be interpreted first.

Ionin (2003) proposed that the only IS configuration under which inverse scope is available in Russian is when the pre-verbal QP is non-topical and may therefore be interpreted in the scope of the post-verbal QP. This configuration holds under (contrastive) focus in Russian. Contra Ionin's (2003) proposal, Antonyuk (2006) argued that inverse scope for Russian sentences such as (3) and (4) is freely available, and is derived by covert QR, as in English (May 1985, Heim & Kratzer 1998).

## 2 Previous Experimental Work on Inverse Scope Availability

Experimental investigations of quantifier scope in Russian (Stoops & Ionin 2013; Ionin, Luchkina & Stoops 2014) have tested the availability of inverse scope in simple transitive SVO and OVS sentences presented in written form and out of context. This work documented a preference for surface-scope readings in double-quantifier sentences such as (3) and (4), more so with SVO than OVS word order. Ionin et al. (2014) also compared surface and inverse scope accessibility for written transitive sentences in Russian and English. In line with Anderson's (2004) processing-based account of scope availability, Ionin et al. (2014) found very similar patterns of results in the two languages, namely, that inverse scope readings are available but are dispreferred relative to surface scope readings. However, this prior work, by focusing on the written modality, did not consider the role of prosody in scope disambiguation.

That a link between the information structure of an utterance in relation to discourse, and its prosodic realization, may help listeners disambiguate scope has been previously established for English, a rigid

word order language (Jackendoff 1972), as well as German, Japanese, and Greek, relatively free word order languages (Baltazani 2002, Hwang 2006, Hirotani 2004, Hirose & Kitagawa 2007, Büring 1997, Krifka 1998, Sæbø' 1997). Production-perception studies of scopally ambiguous sentences in these languages have found that speakers reliably signal the desired scope interpretation by means of acoustic-prosodic cues, such as segment (vowel) duration or pitch, the perceptual correlate of fundamental frequency (e.g., Hirose & Kitagawa 2007). During comprehension, listeners have been shown to reliably utilize the prosodic encoding to determine the scopal relation intended by the speaker. More specifically, specific scope disambiguation effects have been reported. Baltazani (2002) tested the effects of word order and prosodic prominence on scope disambiguation in Greek and found that, independently of word order, placement of prosodic prominence affects the scope reading in double-quantifier sentences. In her work, Baltazani particularly emphasizes the importance of the preceding context on the interpretability of the prosodic effects in Greek: she argues that the accent pattern alone has no direct effect on scope interpretation. Baltazani also reports that select intonation contours in Greek are salient enough to enable listeners to reconstruct the context in which the utterance could be used based solely on its prosodic contour, and hence successfully disambiguate in favor of the intended scope reading.

More recently, Antonyuk-Yudina (2011) and Ionin & Luchkina (in press) have advanced the work on inverse scope availability in Russian double-quantifier sentences by examining the role of prosodic prominence. Recall that with regard to Russian, Ionin (2003) proposes that inverse scope, dispreferred in emotively-neutral sentences, is available under a contrastive focus configuration. Antonyuk-Yudina (2011) tested the scope disambiguating effect of prosodic grouping and contrastive pitch accent. Antonyuk-Yudina used simple transitive SVO and OVS sentences featuring an indefinite subject and a universal object QP. Sentences were presented in contexts biasing readers towards a surface or an inverse scope reading, and prosodic features of eight native speakers' reading performance were impressionistically analyzed. Antonyuk-Yudina did not report a word order effect on inverse scope availability (cf. Ionin & Luchkina in press), however she found that an intermediate phrase break and a perceptual strong, contrastively sounding pitch accent are associated with the surface scope interpretation. A

series of downstepped pitch accents and prosodically neutral realization of the object universal QP, combined with a contrastive pitch accent on the indefinite subject QP, bring about the inverse scope reading. In the perception component of Antonyuk-Yudina's study, native speakers of Russian had to disambiguate the scope double-quantifier sentences based on the prosodic information provided in the reading performance of the model speaker. Relatively poor performance was recorded for inverse scope disambiguation (17% success, compared to 77% success for surface scope disambiguation). A very common result in Antonyuk-Yudina's study is incorrectly disambiguated inverse scope prosody: i.e., respondents had a tendency to match the prosodic realization characteristic of the inverse scope reading with surface scope biasing contexts. Antonyuk-Yudina concluded that the prosodic realization associated with the inverse scope reading in her study is also compatible with the surface scope reading. Ionin & Luchkina (in press) conducted a judgment study (summarized below) and found that prosodic prominence realized on the indefinite quantifier and OVS surface order are both required to derive inverse scope in Russian.

### 3 Research Goals

In the present work, we build on the findings of the judgment study of Russian scope reported in Ionin & Luchkina (in press), focusing primarily on the prosodic characteristics of inverse-scope readings. We pursue the following research goals:

- To determine the acoustic-prosodic features of utterances in which inverse scope is accessible, relative to those which only yield the surface scope interpretation; *and*
- To examine the acoustic-prosodic properties of the prosodically prominent indefinite QP when it is produced in a context that supports a contrastive interpretation, vs. out of context, in order to investigate whether context affects the acoustic-prosodic realization of indefinite QP.

In what follows, we summarize the experimental task used in Ionin & Luchkina (in press), and highlight the key finding of that study, that

prosodic prominence is required to obtain the inverse scope reading in double-quantifier sentences in Russian. We then focus on the prosodic properties of the stimuli used in Ionin & Luchkina (in press) and look in detail at the acoustic-prosodic characteristics that lie at the heart of inverse scope availability in Russian.

#### 4 Experimental Study

Ionin & Luchkina (in press) conducted an auditory sentence-picture verification task (SPVT) in which double-quantifier sentences were presented auditorily along with a picture that illustrated each sentence (see Figures 1 and 2 below). Native speakers of Russian had to listen to each sentence and decide whether it matched the picture by selecting either YES or NO. Target sentences were equally divided between control items, where the sentence is unambiguously true in the context of the picture (e.g., (5) in the context of Figure 1), and test items, where the sentence is true on the surface OR the inverse scope reading, but not both, in the context of the picture (e.g., (5) in the context of Figure 2). For more details about the types of control and test pictures, and their distribution across the four sentence types in (3) and (4), see Ionin and Luchkina (in press).

- (5) Odná sobáka napugála každygo čeloveka.  
 one<sub>NOM</sub> dog<sub>NOM</sub> frightened every<sub>ACC</sub> man<sub>ACC</sub>



Fig. 1: Sample control picture: makes (5) unambiguously true

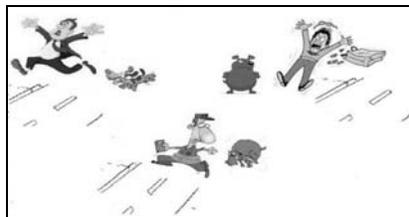


Fig. 2: Sample test picture: makes (5) false on the surface-scope reading and true on the inverse-scope reading

Performance on the test items was analyzed to determine which scope reading, surface or inverse, the respondents accessed for a given sentence type. Two word orders, SVO and OVS, were crossed with two quantifier configurations, indefinite subject with universal object (3) and universal subject with indefinite object (4). Four different SPVT versions were prepared. In the Baseline SPVT, the stimulus sentences were presented out of context and with neutral prosody, i.e., the entire sentence was contained within one prosodic phrase and the nuclear pitch accent was aligned with the sentence-final word. In the Emphasis SPVT, the stimulus sentences were presented out of context and with a perceptually salient pitch accent always on the indefinite quantifier (in follow-up work, Ionin & Luchkina, under review, we have manipulated the prosodic prominence of the universal quantifier as well;). Depending on the word order, the pitch accent could appear sentence-initially, e.g., with the subject quantifier *one* in example (5), or be aligned with the penultimate word in a sentence where the indefinite quantifier appeared in the penultimate position (see (3b) and (4a)). The Topic SPVT differed from the Baseline SPVT only in that the test sentences were preceded with a short two-sentence context which set up the Topic reading of the indefinite quantifier, as shown in (6). The Focus SPVT differed from the Emphasis SPVT in that the test sentences were preceded with a short two-sentence context which set up the contrastive focus reading of the indefinite quantifier. Additionally, the focus particle *vsego* ‘only’ preceded the prosodically prominent indefinite QP, as shown in (7) below.

(6) Topic SPVT:

- Anna: V parke guljali tri mušćiny,  
 In park walked three men  
 potom tuda pribežali tri sobaki.  
 then there ran in three dogs  
 ‘Three men were walking in the park. Then there ran in  
 three dogs.’
- Vera: I čto, sobaki isportili progulku?  
 And what dogs ruined walk  
 ‘And did the dogs ruin the walk?’
- Anna: Nu, odna sobaka napugala každygo človeka.  
 Well one dog frightened every man  
 ‘Well, one dog frightened every man.’

## (7) Focus SPVT:

- Anna: V parke guljali tri mužčiny,  
 in park walked three men  
 potom tuda pribežali tri sobaki.  
 then there ran in three dogs  
 ‘Three men were walking in the park. Then there ran in  
 three dogs.’
- Vera: A dal’še? Navernoe, každogo čeloveka  
 and next probably every man,  
 napugali vse sobaki?  
 frightened all dogs  
 ‘And what happened next? Probably every man was  
 frightened by all the dogs?’
- Anna: Net, čto ty, vsego odna sobaka  
 No what you only one dog  
 napugala každogo čeloveka.  
 frightened every man  
 ‘Oh no, only one dog frightened every man.’

All Baseline and Emphasis SPVT stimuli were recorded by the first author, while the Topic and Focus stimuli were recorded as dialogues by both authors (both are native Russian speakers). Recordings were made in a sound-proof recording booth using a Marantz PDM 750 solid state recorder and a head-mounted microphone. The model speakers were trained to produce prosodic prominence stimuli such that the indefinite quantifier would be audibly prominent.

The study participants were 117 adult native Russian speakers born in Russia or another country where Russian is spoken widely. They ranged in age from 18 to 54 (mean = 23). Results of the SPVT (see Ionin & Luchkina in press for more details) indicate an overall strong preference for the surface scope reading, regardless of prosody or word order. The availability of the inverse scope reading is most evident in the responses to the Emphasis and Focus SPVTs. Statistical analysis showed that prosodic prominence interacts significantly with (1) quantifier configuration (indefinite or universal in the subject position) and (2) word order. Ionin and Luchkina (in press) report that the SVO stimuli do not demonstrate considerable availability of the inverse scope reading, regardless of prosody. The availability of the inverse scope reading

increases significantly whenever the indefinite QP is in the preverbal object position in the OVS order and is prosodically prominent. Despite the fact that the context preceding the test sentences in the Focus (or Topic) SPVT clearly sets up the IS category of the indefinite QP as focus (or topic), it plays no significant role in the availability of surface or inverse scope readings. For the remainder of this paper, we focus on the prosodic characteristics related to surface vs. inverse scope availability.

## 5 Prosodic Analysis

To address our first research goal, namely, what acoustic-prosodic features characterize the oral productions in which inverse scope is accessible, relative to those which only yield the surface scope interpretation, prosodic features of the test sentences from each SPVT version were automatically extracted and analyzed. In the following analyses, we exclude the data extracted from the Topic SPVT. Because in both Baseline and Topic SPVTs, the target sentences are produced with normal tempo and emotively-neutral prosody, we do not anticipate systematic differences in the prosodic properties of the target sentences extracted from these SPVTs. Instead, we focus our attention on the systematic differences in the acoustic-prosodic features of the Baseline SPVT, which offered no prosodic cues, and the Focus and Emphasis SPVTs, which featured a prosodically prominent indefinite QP.

Figures 3 & 4<sup>1</sup> show the averaged time-normalized pitch contours of the SVO and OVS test sentences produced with non-emotive vs. contrastive prosody.

---

<sup>1</sup> Figures 3, 4, and 5 illustrate experimental stimuli in which the indefinite quantifier is in the object position. Ionin and Luchkina (in press) found that prosody affected scope interpretation only when the indefinite was in object position, not when it was in subject position. However, the prosodic contours look very similar regardless of the syntactic position of the indefinite. To save space, we only report the contours for the sentences where the indefinite is in object position.

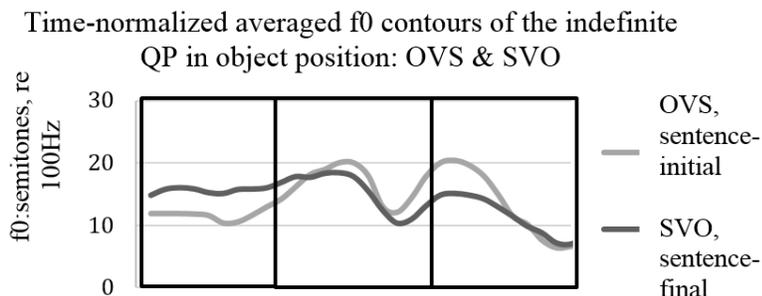


Fig. 3: OVS test sentences: indefinite QP in sentence-initial object position. Each rectangle sector encloses one word as in *odnu sobaku*, ‘one dog<sub>ACC</sub>’.

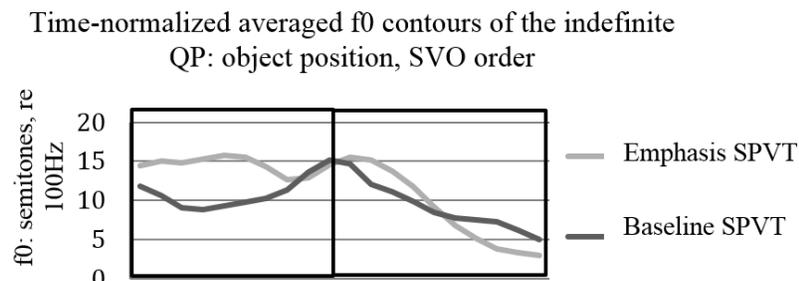


Fig. 4: SVO test sentences: indefinite QP in sentence-final object position. Each rectangle sector encloses one word, as in *odnu sobaku*, ‘only one dog<sub>ACC</sub>’.

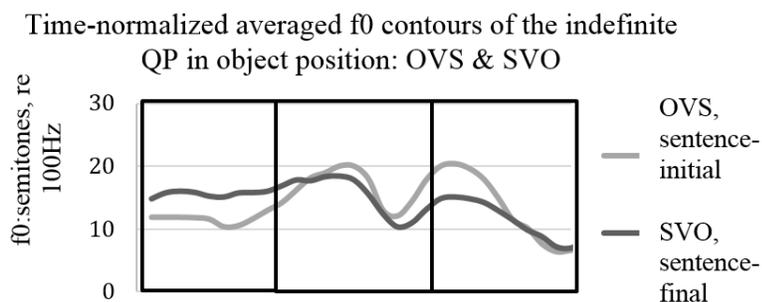


Fig. 5: Focus SPVT: time-normalized averaged f0 contours of SVO & OVS test sentences with indefinite QP in object position. Each rectangle sector encloses one word, as in *vsego odnu sobaku*, ‘only one dog<sub>ACC</sub>’.

Note that independently of the word order or sentence position of the indefinite quantifier, the resulting prosodic contour looks very consistent in the Emphasis SPVT. The pitch accent peak and valley in the left sector in Figures 3 and 4 visually illustrate the perceptually salient (contrastive) reading of the indefinite quantifier *odin* followed with a gradual downstep in f0 in the right sector, corresponding to the noun that is adjacent to the quantifier. In the Focus SPVT (see Fig. 5), the indefinite QP is contrastively focused and prosodically prominent. Additionally, the focus marker *vsego*, shown in the leftmost sector in Fig. 5, precedes the indefinite QP and expands its focus domain. Time-normalized averaged pitch contours of the indefinite QPs in the Focus SPVT are comparable to the contours obtained for the Emphasis SPVT: the contrastive pitch accent aligns with the quantifier shown in the middle sector in Fig. 5 and is followed with a downstep in f0 in the following noun, shown in the rightmost sector in Fig. 5. Next, we compared select prosodic qualities of experimental sentences across prosodic conditions.

We chose to work with the cross-linguistically attested acoustic-prosodic correlates of prosodic prominence (Ladd 2008): f0 maxima and minima, intensity and duration of the stressed vowel. We also measured the distance from the midpoint of the vowel to its tonal center of gravity (henceforth, TCoG distance)<sup>2</sup>. All acoustic-prosodic measurements were taken from the stressed syllable of each content word in the target sentences of the Baseline, Emphasis, and Focus SPVT versions. Measurements were extracted automatically in Praat (Boersma & Weenink, 2013). The values of max f0 and max intensity were taken from the center region of the stressed vowel in order minimize the influence of the adjacent segments at the voice onsets and intersegmental transitions. All f0 outputs were transformed to semitone values relative to a fixed value of 100 Hz. Intensity and duration values were normalized using the natural logarithm scale. For the analyses reported below, we only use prosodic data extracted from the indefinite QP. Differences in the means and distributions of the acoustic-prosodic measurements for the indefinite quantifier *odin* across the SPVT versions and the focus marker *vsego* used in the Focus SPVT are shown in Figures 6-8.

---

<sup>2</sup> Luchkina & Cole (2014) report that the distance between the vowel midpoint to its tonal center of gravity is an effective correlate of prosodic prominence in Russian.

The distributions plotted in Figures 6-8 illustrate that compared to the Baseline SPVT, the Emphasis and the Focus SPVTs feature rather distinctive prosodic realizations of the indefinite quantifier, with greater mean values of f0 range, duration, and intensity. The acoustic-prosodic qualities of the focus marker *vsego* confirm that it is prosodically prominent and belongs to the contrastive focus domain in the Focus SPVT.

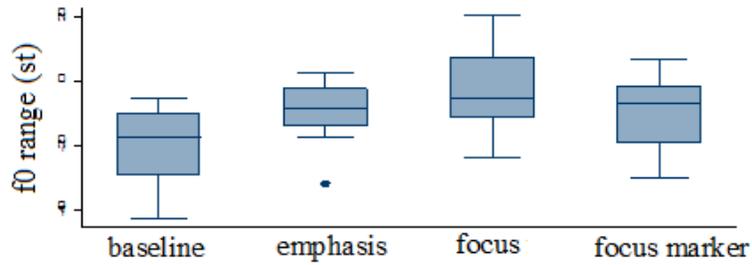


Fig. 6: Means and distributions of the f0 range (f0 max-f0 min) across the SPVT versions. Measurements shown from the indefinite quantifier and the focus marker (focus marker present in the Focus SPVT only).

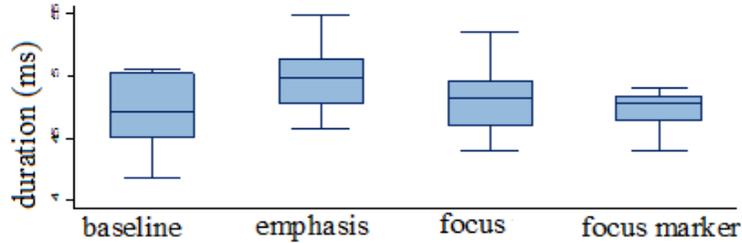


Fig. 7: Means and distributions of the stressed vowel duration across the SPVT versions. Measurements shown from the indefinite quantifier and the focus marker (focus marker present in the Focus SPVT only).

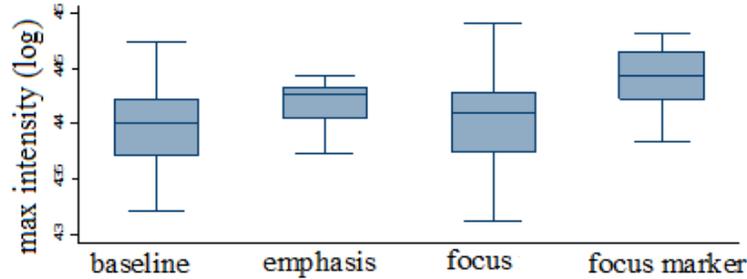


Fig. 8. Means and distributions of the stressed vowel max intensity across the SPVT versions. Measurements shown from the indefinite quantifier and the focus marker (focus marker present in the Focus SPVT only).

So far we have argued for systematic acoustic-prosodic differences between the emotively-neutral vs. contrastively-prominent productions of the indefinite QP and presented some preliminary evidence that robust prosodic cues play a critical role in making the inverse scope readings available in the Emphasis and Focus SPVTs. However, when used in natural speech, such cues are interpretable relative to discourse in which they occur. With regards to scope relations, Baltazani (2002) argues that in Greek, prosody by itself does not influence scope relations, rather it is the IS configuration, contextually set up and supported by distinctive prosodic cues, that renders the inverse scope reading available. Similarly, Ionin (2003) proposes that in Russian, scope is interpreted based on the IS category of the indefinite QP: inverse scope is unavailable when the indefinite QP occupies the pre-verbal sentence Topic position, however a contrastively focused indefinite QP makes the inverse scope reading available. Ionin & Luchkina (in press) tested this prediction using Topic and Focus SPVTs. They found no differences in the preference for surface scope in Topic SPVT relative to Baseline SPVT, as well as in Focus SPVT relative to the Emphasis SPVT. Recall that in addition to a context which set up the contrastive reading of the indefinite QP, Focus SPVT also featured the focus marker *vsego* ‘only’, which reinforced the focus reading of the indefinite QP and expanded its focus domain. Ionin & Luchkina’s results indicate that neither of these cues has a significant effect on inverse scope availability. Given highly similar acceptability rates for the inverse scope readings obtained in these SPVT versions, we

predict highly similar prosodic expression of prominence in the Emphasis SPVT and the Focus SPVT. To test this prediction, we subjected our prosodic data to a rigorous statistical test of each acoustic-prosodic parameter's ability to predict a major prosodic condition of the SPVT study. By assessing the differences in the magnitude of the predictive power for each of our measurements, we are able to gauge the overall similarity between the oral productions of the indefinite QP extracted from Baseline, Focus, and Emphasis SPVTs. To this end, the model speakers' production data were submitted to a multinomial logistic regression analysis which determined how well each prosodic condition could be predicted based off the acoustic-prosodic parameters of f0 (maxima and minima), vowel intensity and duration, and TCoG distance. The dependent variable in the analysis was prosodic condition (3 levels: Baseline, Emphasis, Focus). The Baseline and Focus prosodic condition were each used as the reference level of the dependent variable. The model was fit in STATA. To save space, we only report the output with Focus prosody as the reference level of the dependent variable. Results of the multinomial logistic analysis are summarized in Table 1.

Condition	Parameter	Regression coefficient
Baseline (relative to Focus)	intensity	-21.7***
	duration	3.37***
	f0 max	-1.56**
	f0 min	1.06***
	TCoG distance	608.9*
Emphasis (relative to Focus)	intensity	-9.38*
	duration	ns
	f0 max	-.59*
	f0 min	.49**
	TCoG distance	ns

Table 1. Multinomial regression ( $\chi^2(18)=103.99$ ,  $p<0.001$ ) estimates and significance levels for the acoustic parameters of f0, intensity, duration, and TCoG distance. Measurements from Focus SPVT are used as the reference level of the dependent variable.

The results indicate that all acoustic-prosodic parameters chosen for the analysis robustly differentiate between Focus and Baseline SPVTs, which represent two extremes of the experimental design: whereas the Baseline version offers no prosodic or context cues, the Focus version offers both. The magnitude of the regression coefficients, suggestive of each factor's effect size, shows that the distance between the central point of the stressed vowel and its tonal center of gravity, i.e., the region with the maximum area in the f0 curve over that vowel, is the strongest discriminant between Baseline and Focus prosody. Additionally, vowel intensity, duration, and f0 minima and maxima all robustly discriminate between Baseline and Focus prosodic conditions. Table 1 also indicates that the prosodic differences between the Emphasis and the Focus SPVTs are more subtle: neither duration nor TCoG distance discriminate between these two SPVT versions, and intensity and f0 maxima and minima, while all significant, have considerably smaller effect sizes (see the regression coefficients associated with these variables in Table1), confirming our expectation that the prosodic realizations of the indefinite QP presented in the Emphasis and Focus SPVTs are highly similar.

To summarize, we have visually inspected the f0 contours of the target sentences presented in emotively-neutral Baseline SPVT, as well as in the Emphasis and Focus SPVTs, in which the indefinite QP is prosodically prominent. We then examined the discriminability of the oral productions of the indefinite QP extracted from Baseline, Emphasis, and Focus SPVTs. Results of our analyses indicate that (1) more eventful f0 contours and (2) qualitatively distinct values of intensity, duration, and f0 characterize the indefinite QP produced in Emphasis and Focus SPVTs, but not in Baseline SPVT. Finally, we have highlighted the similarities between the indefinite QP qualities in the Focus and Emphasis SPVTs. These similarities are particularly important since regardless of other design differences, such as context availability, these SPVT versions yielded comparable rates of inverse scope availability in Ionin & Luchkina's study.

## 6 Discussion

In this work, we continue to analyze the interpretability of double-quantifier sentences in Russian, a free word order language. In languages like English, transitive sentences with two quantifiers, one existential and

the other universal, are scopally ambiguous. However, inverse scope in English has been found to be dispreferred to surface scope, possibly because it incurs a greater processing cost (Anderson, 2004). Inverse scope availability in Russian has been characterized as even more limited: Ionin & Luchkina (in press) recently reported that while the surface scope reading is consistently preferred over the inverse scope reading, inverse scope readings are more accessible in OVS sentences in which the indefinite quantifier is preverbal and is prosodically prominent. Building on the findings of Ionin & Luchkina (in press), in this study, we analyze the acoustic-prosodic features of the indefinite QPs used in different prosodic conditions in Ionin & Luchkina (in press). We find that a number of systematic acoustic-prosodic correlates of pitch accent, such as  $f_0$  range, vowel segment intensity and duration, and distance between the vowel midpoint and its tonal center of gravity reliably differentiate the prosodic realization of the indefinite QP in the Focus and Emphasis SPVTs from those in the Baseline SPVT. Our results are in line with Ionin's (2003) proposal that a contrastive reading of the indefinite QP is needed for the inverse scope to become available.

One unexpected finding of Ionin & Luchkina (in press) is that the inverse scope accessibility in the Emphasis SPVT, which features only prosodic prominence, matches that obtained in the Focus SPVT, where context helps set up a contrastive focus reading of the indefinite QP. Recall Baltazani's (2002) argument that in Greek, prosodic prominence should be seen as secondary to context, which is necessary to motivate the presence of the contrastive pitch accent and set up the IS configuration matching the focus reading of the indefinite QP. According to Baltazani, listeners 'proceed from prosodic structure to information structure to scope calculation' (2002:73). In the present study, we have experimentally shown that regardless of the context manipulation and availability of a focus marker, the prosodic features of the indefinite QP produced in the Emphasis and Focus SPVTs are highly similar. Following Ionin & Luchkina (in press), we propose that Russian speakers access inverse scope readings from contrastive prosody alone, and reconstruct the IS from prosody. This proposal agrees with the view expressed by Baltazani (2002) that it is possible for listeners to interpret (particularly salient) prosodic contours out of context, and thus successfully access both readings of the scopally ambiguous sentence. We conclude that two IS configurations are available in Russian double-

quantifier sentences, one in which the pre-verbal QP is interpreted as the sentence Topic (under neutral prosody) and one in which the pre-verbal QP is interpreted as contrastively focused. Our results suggest that the preference for surface scope in Russian may be overridden by IS considerations: surface scope is strongly preferred when the preverbal NP is the topic, but inverse scope becomes available when the preverbal NP is in contrastive focus. This is consistent with the proposal of Neeleman and Titov (2009) that contrastively focused NPs reconstruct to their base position in the scope of the subject, making the inverse scope readings available.

## 7 Conclusion

Building on the findings reported in Ionin & Luchkina (in press) that inverse scope availability is significantly higher in scrambled OVS sentences which feature a contrastive pitch accent aligned with the indefinite QP, we have described the acoustic-prosodic features which successfully discriminate between the productions disambiguating surface and inverse scope in Russian. We also determined that the role of discourse-level information in inverse scope availability in Russian is secondary to the role played by the prosodic cues. We concluded our investigation with a tentative argument that the prosodic contour under which the inverse scope is available is associated with a highly salient information structural configuration in which the pre-verbal QP is interpreted as focused. Such an IS configuration is unavailable under neutral prosody, where the pre-verbal QP is in Topic position.

A still-unresolved question, at present, is how our findings on the relationship between scope and prosody compare to those on other languages. Specifically, Bobaljik and Wurmbrand (2012) argue that languages such as Japanese and German, inverse scope becomes available whenever the IS and the LF are in conflict (i.e., whenever the IS topic-focus configuration does not match the surface-scope configuration). Our data do not align with those discussed by Bobaljik and Wurmbrand for German and Japanese, and our prosodic configuration (prosodic prominence on the indefinite quantifier) is different from that discussed by Bobaljik and Wurmbrand (a rise-fall contour, cf. Krifka 1998). In follow-up work (Ionin & Luchkina, under review), we examine the effects of a rise-fall prosodic contour on

Russian scope, in order to allow for a cross-linguistic comparison of the effects of prosody on scope.

### References

- Anderson, Catherine. 2004. The Structure and Real-Time Comprehension of Quantifier Scope Ambiguity. PhD dissertation, Northwestern University.
- Antonyuk, Svitlana. 2006. The Scope of Quantifier Phrases in Russian: A QR Analysis. Linguistics in the Big Apple.
- Antonyuk-Yudina, Svitlana. 2011. Why prosody matters. An extended 4-page abstract of the presentation, Linguistic Society of America, Annual Meeting, Pittsburg.
- Bailyn, John. 1995. A Configurational Approach to Russian 'Free' Word Order. PhD dissertation, Cornell University.
- Baltazani, Maria. 2002. The prosodic structure of quantificational sentences in Greek. In Andronis, M., E. Debenport, A. Pycha, and K. Yoshimura (eds.) *Proceedings from the 38th Meeting of the Chicago Linguistic Society*, 63-78. Chicago, IL: Chicago Linguistic Society.
- Beck, Sigrid. 1996. Quantified structures as barriers for LF movement. *Natural Language Semantics* 4: 1-56.
- Bobaljik, Jonathan & Susi Wurmbrand. 2012. Word order and scope: transparent interfaces and the  $\frac{3}{4}$  signature. *Linguistic Inquiry* 43: 371-421.
- Boersma, Paul & David Weenink. 2013. *Praat: Doing Phonetics by Computer* [Computer program]. Version 5.3.60.
- Büring, Daniel. 1997. The great scope inversion conspiracy. *Linguistics and Philosophy* 20: 175-194.
- Calhoun, Sasha. 2010. The centrality of metrical structure in signaling information structure: a probabilistic perspective, *Language*, 86(1): 1-42.
- Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press and MITWPL.
- Heim, Irene & Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Malden, MA: Blackwell.

- Hirotsu, Masako. 2004. Prosody and LF: Processing of Japanese Wh-questions, Ph. D. dissertation, University of Massachusetts at Amherst.
- Hirose, Yuki & Kitagawa, Yoshihisa. 2007. Prosody-scope correlation in wh-interrogatives: production and perception, Paper presented at International Conference on Processing Head-final Structures, Rochester Institute of Technology, Rochester, NY.
- Hwang, Hyun Kyun. 2006. Intonation patterns of wh-interrogatives in South Kyungsang Korean and Fukuoka Japanese. *Eoneohak, Journal of the Linguistic Society of Korea*, 45: 39-59.
- Ionin, Tania. 2001. The one girl who was kissed by every boy: scope, scrambling and discourse function in Russian. In Marjo van Koppen, Joanna Sio, and Mark de Vos (eds.), *Proceedings of ConSOLE X*, Leiden University: Leiden, The Netherlands, 65-80.
- Ionin, Tania, Tatiana Luchkina & Anastasia Stoops. 2014. Quantifier scope and scrambling in the second language acquisition of Russian. In Chin, C.-Y., Coughlin, C. E., Prego, C. L., Minai, U., Tremblay, A. (eds.), *Proceedings of GALANA-4*, Cascadilla Proceedings Project, Somerville, MA, 169-179.
- Ionin, Tania. & Tatiana Luchkina. in press. One reading for every word order: Revisiting Russian Scope. *Proceedings of WCCFL 2014*.
- Ionin, Tania & Tatiana Luchkina. under review. Focus on Russian scope: an experimental investigation of the relationship between quantifier scope, prosody, and information structure.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.
- Krifka, Manfred. 1998. Scope inversion under the Rise-Fall contour in German. *Linguistic Inquiry* 28: 662-688.
- Ladd, D. Robert. 2008. *Intonational Phonology*. Cambridge, Cambridge University Press.
- Luchkina, Tatiana & Jennifer S. Cole. 2014. Structural and prosodic correlates of prominence in free word order language discourse. In Campbell, N., Gibbon, D., and Hirst, D. (eds.), *Proceedings of the 2014 meeting of the Speech Prosody Conference (SP 7)*, Dublin, Ireland, 1119-1123.
- May, Robert. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.

- Miyagawa, Shigeru. 1997. Against optional scrambling. *Linguistic Inquiry* 28: 1-26.
- Neeleman, Ad & Titov, Elena. 2009. Focus, contrast and stress in Russian. *Linguistic Inquiry* 40, 514-524.
- Rosengren, Inger. 1993. Wahlfreiheit mit Konsequenzen – Scrambling, Topikalisierung und FHG im Dienste der Informationsstrukturierung, in Reis, M. (ed.) *Wortstellung und Informationsstruktur*, Tübingen: Niemeyer, 251–312.
- Stoops, Anastasia & Ionin, Tania. 2013. Quantifier scope and scrambling in Russian: an experimental study. In Franks, S., Dickinson, M., Fowler, G., Whitcombe, M., and Zanon, K. (eds.), *Formal Approaches to Slavic Linguistics #21: The Third Indiana Meeting 2012*. Ann Arbor, MI: Michigan Slavic Publications, p. 344-358.
- Slioussar, Natalia. 2011. Processing of a free word order language: the role of syntax and context. *Journal of Psycholinguistic Research*, 40: 291-306.
- Sæbø, Kjell Johan. 1997. Topic, focus and quantifier raising. In P. Dekker, M. Stokhof, & Y. Venema (ed.), *Proceedings of the 11th Amsterdam Colloquium*, 67-72. Amsterdam: Institute for Logic, Language and Computation.
- Tunstall, Susanne Lynn. 1998. The Interpretation of Quantifiers: Semantics and Processing. Ph.D. thesis, UMass.

luchkin1@illinois.edu  
tionin@illinois.edu

## **4000 Measure NPs: Another Pass Through the *шлюз*\***

*Ora Matushansky*

SFL (CNRS/Université Paris-8)/UiL OTS/Utrecht University

*E.G. Ruys*

UiL OTS/Utrecht University

In this paper we will argue that the default neuter singular agreement with cardinal-containing NPs in Russian is a diagnostic for the degree denotation of these NPs. We will then advance a hypothesis as to how the correlation between the denotation and agreement can be realized formally.

### **1 Evidence for the Degree Denotation of Numeral NPs**

The semantic type of degrees has been introduced in the context of scalar adjectives, such as *tall*, but also extends to other measurement contexts, such as the direct object position of verbs like *weigh*, *last*, etc. Degrees are abstract representations of measurement, which are generally taken to be points or intervals on a totally ordered scale (Seuren 1973, Cresswell 1976, Hellan 1981, von Stechow 1984, Bierwisch 1989, Heim 1985, etc.), or equivalence classes (Krifka 1990), and naturally correspond to measure phrases in examples like (1). It is easy to see that any semantic environment requiring degree denotation is compatible not only with NPs headed by inherent measure nouns, but also with cardinal-containing

---

\* For discussion and useful comments we are grateful to the audiences at RALFe 2012 (Paris VIII), FDSL 10 (Leipzig) and FASL 23 (Berkeley), where different parts of this work were presented, as well as to Masha Polinsky and to Katja Jasinskaja.

NPs headed by other nouns (henceforth, numeral NPs), as in (2). In other words, a degree can be constructed on the basis of any numeral NP.

- (1) a. The track is **four hundred meters** long.  
 b. The presentation lasted **two hours**.
- (2) a. V Komi **tremja počëtnymi graždanami** stalo bol'se.  
 in Komi three<sub>INS</sub> honorary<sub>INS</sub> citizens<sub>INS</sub> became more  
 'Komi has [acquired] three honorary citizens more.'  
 b. Èta serija **na pjat' knig** dlinnee.  
 this series on five book<sub>GEN</sub> longer  
 'This series is five books longer.'

The hypothesis that NPs in argument positions can denote degrees as well as entities has also been proposed to account for the modified cardinal construction in English (3) by Gawron 2002, as well as by Billings 1995, in order to deal with the Russian approximative inversion.

- (3) Barry Bonds hit a career-best 73 home runs. Gawron 2002

Russian can furthermore be argued to distinguish degree-denoting NPs from their entity-denoting counterparts when the lexical head noun is animate. Animacy in Russian is generally determined by the lexical semantics of a noun and can be diagnosed by the surface realization of accusative case for plural and second-declension nouns: accusative is syncretic with nominative for inanimates and with genitive for animates, as shown for plural NPs in (4) and for numeral NPs in (5).<sup>1</sup> However, as Mel'čuk 1980a, b shows, numeral NPs behave as inanimate in showing up with surface nominative after accusative-assigning prepositions in "quantity" readings (6), even if their lexical head is an inherently animate noun. Degree denotation therefore has a clear effect on the syntax of a numeral NP in an area other than agreement.

---

<sup>1</sup> The phenomenon is only observed for cardinals lower than five; higher cardinals either belong to the declension classes that do not reflect animacy or, like *million*, are treated as inanimate nouns.

- (4) Ty videla stol/ mal'čika.  
 you saw table<sub>ACC=NOM</sub> boy<sub>ACC=GEN</sub>  
 'You saw a/the table/boy.'
- (5) a. Ty videla \*dvux stolov/ dva stola.  
 you saw two<sub>ACC=GEN</sub> table<sub>GEN.PL</sub> two<sub>ACC=NOM</sub> table<sub>PAUC</sub>  
 'You saw two tables.'  
 b. Ty videla dvux mal'čikov/ \*dva mal'čika.  
 you saw two<sub>ACC=GEN</sub> boy<sub>GEN.PL</sub> two<sub>ACC=NOM</sub> boy<sub>PAUC</sub>  
 'You saw two boys.'
- (6) a. bol'se na dva mal'čika Mel'čuk 1980b  
 more on two<sub>ACC=NOM</sub> boy<sub>PAUC</sub>  
 'two boys more'  
 b. siloj rovno v tri medvedja  
 strength<sub>INS</sub> exactly in three<sub>ACC=NOM</sub> bear<sub>PAUC</sub>  
 'as strong as exactly three bears'

Matushansky and Ruys [to appear] confirm the hypothesis that default agreement and lack of animacy can be used as diagnostics for the degree denotation of numeral NPs. Evidence for this comes from a number of other cases where the degree denotation can be argued for on semantic grounds and where numeral NPs trigger default agreement and behave as inanimates even when their lexical head is animate.

The first piece of evidence comes from numeral NPs headed by inherent measure nouns, which, as noted by Rothstein and Khrizman 2013, must trigger default agreement (unless they are specific or definite), as shown in (7).<sup>2</sup> Second, the same is true for the so-called event-oriented readings of numeral NPs (8), which Krifka 1990 argues to involve degrees: event-oriented readings are only compatible with default agreement:

<sup>2</sup> It is unclear whether the animacy diagnostic works for inherent measure nouns, which are naturally inanimate. The only possible candidate for verification is the classifier-like item *čelovek* 'person' (see Sussex 1976, Yadroff 1999 for a discussion of such items in Russian), which does behave as inanimate (i) and which Matushansky [to appear] assigns to the category of measure nouns for independent reasons:

(i) nanjali četyre čeloveka / \*četyrëx čelovek učenyx  
 hired four<sub>ACC=NOM</sub> person<sub>PAUC</sub> four<sub>ACC=GEN</sub> persons<sub>GEN</sub> scientists<sub>GEN</sub>  
 '[they] hired three scientists'

- (7) Prošlo/\*prošli pjat' let.  
 went<sub>NSG/PL</sub> five years<sub>GEN</sub>  
 'Five years passed.'
- (8) a. 4000 korablej prošli čerez šljuz. individual ships only  
 4000 ships passed<sub>PL</sub> through lock  
 '4000 ships passed through the lock.'  
 b. 4000 korablej prošlo čerez šljuz. no commitment  
 4000 ships passed<sub>NSG</sub> through lock  
 '4000 ships passed through the lock.'

Accumulative verbs formed with the prefix *na-* (Pereltsvaig 2006) give us another semantic diagnostic for degree denotation, since their direct object provides the measure of the event. As discussed in Matushansky and Ruys [to appear], such direct objects exhibit the inanimate pattern for the accusative case and, when passivized, require default agreement if the surface subject is a numeral NP:

- (9) a. Vsego u nix bylo nabrano tysjača slov.  
 all.in.all at them was<sub>NSG</sub> ACM taken<sub>NSG</sub> thousand words<sub>GEN</sub>  
 'Overall, a thousand words was collected by them.'  
 b. \*Vsego u nix byli nabrany tysjača slov.  
 all.in.all at them was<sub>PL</sub> ACM taken<sub>PL</sub> thousand words<sub>GEN</sub>

Finally, approximative inversion, which Rothstein and Khrizman 2013 convincingly argue to yield the measure interpretation of NP, also gives rise to the default agreement pattern, as well as to the inanimate pattern of accusative case syncretism.

Having thus established the fact that, in Russian, degree-denoting numeral NPs can be distinguished by their syntax, two questions arise. (i) Do the inanimate pattern of accusative case syncretism and/or default agreement truly distinguish degree denotation from entity denotation? (ii) How can the two be formally implemented?

Before addressing both, we will first discuss the semantics of degree-denoting numeral NPs in the subject position.

## 2 Krifka's Compositional Semantics for Degrees in Argument Positions

Krifka's analysis of examples like (8) presupposes a definition of degrees as quantized predicates, i.e., the more or less regular definition of a numeral NP compounded with the assumption that any noun can be used as a measure noun (10). The latter assumption is however not crucial, as the same result is achieved by using the semantics of cardinals in Ionin and Matushansky 2006, which allows a non-individuated reading of regular lexical nouns by making the non-intersection condition part of the semantics of the cardinal itself (see Ruys 2014 for discussion).

- (10) a.  $\llbracket 60 \text{ tons of waste} \rrbracket = \lambda u . [\text{waste}'(u) \ \& \ \text{ton}'(u) = 60]$   
 b.  $\llbracket 4000 \text{ ships} \rrbracket = \lambda u . [\text{ship}'(u) = 4000]$

In order to saturate an argument position of a predicate, such quantized predicates need to be first combined with a determiner. The traditional entity denotation is straightforwardly achieved by the usual existential quantifier, which is phonologically null and asserts the existence of the (plural) individual that the numeral NP measures (11):

- (11)  $\emptyset = \lambda Q \lambda R \lambda e \exists u [R(e, u) \ \& \ Q(u)]$                       null existential quantifier

The event-related reading, on the other hand, measures events in the terms of the numeral NP by using the special determiner OEMR (and assuming therefore that degree-denoting NPs are DPs).<sup>3</sup>

---

<sup>3</sup> The lexical entry in (11) does not account for the fact that event-related readings can also be obtained with the definite article (i). To avoid this issue the definite article should be defined as an identity function rather than as the iota operator -- a move that is equally warranted by the availability of the definite article on predicate definites.

(i) The 4000 ships that passed through the lock last year stretched its capacity to the limit.

(12) Let  $\Sigma$  be a quantized predicate of the lattice sort and  $\alpha$  an event relation then OEMR ( $\alpha$ ) is defined as the smallest relation  $\sigma$  between an event and a quantized predicate of the lattice sort  $\Sigma$  such that (for any event  $e$  and quantized predicates  $\beta, \beta'$ )

(i) (Standardization)

$$\neg \text{ITER}(e, \alpha) \rightarrow [\sigma(e, \beta) \leftrightarrow \exists u[\beta(u) \ \& \ \alpha(e, u)]]$$

(ii) (Generalization)

$$\neg e \circ_{\Sigma} e' \ \& \ \sigma(e, \beta) \ \& \ \sigma(e', \beta'') \rightarrow \sigma(e \cup_{\Sigma} e', \beta +_{\Sigma} \beta')$$

In other words, if the event is not iterative (i.e., e.g., no ship passes the lock more than once), then OEMR yields true of an event and a measure if there exists an entity having that measure (i.e., a plural individual consisting of 4000 ships) that participated in the event. If, on the other hand, the event is iterative, it is separated into a number of non-iterative non-overlapping pass-through-the-lock sub-events and the measure is separated into the corresponding number of sub-measures and together they add up to the main event and 4000 ships, respectively.

From our perspective, the entity denotation of a numeral NP corresponds to the existential quantifier, while the degree denotation is handled by OEMR. The question now arises of why the two denotations yield two different agreement patterns and how this is achieved formally.

### 3 Degree Denotation and Individuation Hierarchies

It is well established that Russian non-agreeing numeral NP subjects lack some interpretations that a numeral NP is expected to have. Pereltsvaig 2006 shows that they cannot be definite or specific (13), outscope other quantifiers (14) or (with certain caveats, on which see Matushansky and Ruys [to appear]) control PRO and bind freestanding reflexives.

- (13) a. Pjat' knig na stole byli/ \*bylo moi.  
 five bookson table were<sub>PL</sub>/ was<sub>N<sub>NSG</sub></sub> mine<sub>PL</sub>.  
 'The five books on the table were mine.'  
 b. Kakie-to tri knigi prodajutsja/ \*prodajetsja deševle.  
 some three books sell<sub>3<sub>PL</sub>.REFL</sub>/ \*sell<sub>3<sub>SG</sub>.REFL</sub> cheaper  
 'Some three books are being sold cheaper.'

- (14) a. Každyj raz pjat' xirurgov operirovali Bonda.  $\checkmark \forall > 5$ ,  $\checkmark 5 > \forall$   
 every time five surgeons operated<sub>PL</sub> Bond  
 'Every time five surgeons operated on Bond.'  
 b. Každyj raz pjat' xirurgov operirovalo Bonda.  $\checkmark \forall > 5$ ,  $*5 > \forall$   
 every time five surgeons operated<sub>NSG</sub> Bond  
 'Every time five surgeons operated on Bond.'

The hypothesis that non-agreeing numeral NPs denote degrees rather than entities explains these facts. As a matter of fact, inherent measure phrases (i.e., unambiguous degrees) in argument possessions of measure verbs are limited in a very similar way: they cannot be quantified (Adger 1996), they cannot be extracted across weak islands (Rizzi 1990)<sup>4</sup> and they cannot be specific:

- (15) a. \*It weighs a certain five pounds.  
 b. \*The talk lasted every hour.

It can be objected that inherent measure phrases can be definite, as in (16). While in (16a) the measure phrase can be argued to denote an entity rather than a degree, corresponding to a specific period of time, there is no identifiable weight entity in (16b) and in (16c) the measure phrase can hardly be argued to denote an existing entity.

- (16) a. The team lasted the twenty-three years of his pontificate.  
 b. She weighs the same fifty-eight pounds (as before/\*Sue).  
 c. I will never gain the five kilos they want me to gain.

Furthermore, asserting that such definite measure phrases do not denote degrees in environments that semantically select for these is not only theoretically suspicious, but also leaves us with the question of how to diagnose degree denotation. We therefore conclude that degrees, just like entities, can be both definite and specific. The fact that default agreement is impossible with definite and specific measure NPs (17) does not rule out the hypothesis that it is in fact sensitive to degree denotation.

---

<sup>4</sup> This diagnostic is not applicable to Russian non-agreeing numeral NPs subjects: on the one hand, subjects are not easy to move to begin with, and on the other, Russian does not easily allow extraction out of finite clauses.

- (17) (Èti) dva djujma, na kotorye ja vyše, ne igrajut/\*igraet roli.  
 these two inches on which I taller NEG play<sub>PL/\*NSG</sub> role  
 ‘These two inches by which I am taller play no role.’

There is furthermore independent cross-linguistic evidence that links agreement to degree denotation. Thus in Dutch (Klooster 1972) measure NP subjects systematically trigger singular agreement on the verb; it is also possible in Basque (Etxeberria and Etxepare 2008, 2012):

- (18) a. Er **staat/\*staan** drie liter water op tafel. Dutch  
 there stand<sub>SG/PL</sub> three liter<sub>SG</sub> water on table  
 ‘There are three liters of water on the table.’  
 b. Hiru litro ardo edan du/ditu. Basque  
 three liter wine drunk AUX<sub>SG/AUX<sub>PL</sub></sub>  
 ‘He/she drank three liters of wine.’

We propose to explain these facts by proposing that degree-denoting NPs are located at the bottom of various individuation hierarchies (Silverstein 1976, Comrie 1981, Aissen 1999, 2003, Bickel 2008), which may differentiate between animate vs. inanimate, specific vs. non-specific, pronominal vs. nominal, etc., NPs with respect to agreement, number marking and case-marking. To provide just one example, in the Chadic language Miya (Schuh 1989, 1998) agreement and concord are sensitive to animacy: as examples (19-20) from Schuh 1998:197 show, while with plural animate nouns the demonstrative takes the plural form and therefore bears no indication of gender (i.e., nouns of both genders appear with the same plural forms), with plural inanimate nouns the demonstrative does not agree in number (though it does agree in gender). Attributive APs, relative pronouns, linkers, etc., behave the same, showing that [-animate] nouns do not trigger number agreement.

- (19) a. níykin dzáfə animate: number agreement  
 this<sub>PL</sub> man<sub>PL</sub>  
 b. níykin təmakwìy  
 this<sub>PL</sub> sheep<sub>PL</sub>

- (20) a. *nákən víyayúw-awàw*      inanimate: gender agreement only  
           this<sub>M.SG</sub> fireplace<sub>M.PL</sub>  
       b. *tákən tlərkáy-ayàw*  
           this<sub>F.SG</sub> calabash<sub>F.PL</sub>

The Miya case is extremely revealing in that it demonstrates that default agreement or lack of agreement for number can coincide with agreement for gender, in other words, that it is not the agreement process itself that fails to take place with inanimate nouns in Miya. This in turn suggests that default agreement in Russian may also result not from some putative invisibility of degree-denoting NPs to the syntactic process of agreement, but perhaps rather from the failure of the probing phi-features on T° to find a proper target. This is in fact the analysis that we intend to pursue.

#### 4 The Syntax of Non-Agreeing Numeral NPs: Preliminaries

As discussed above, a purely semantic approach to non-agreeing numeral NPs is untenable: definite or specific numeral NPs cannot trigger default agreement even if they clearly denote degrees (17), nor can numeral NPs headed by *one* (21a), unless their lexical NP is headed by an inherent measure noun (21b). On the other hand, a purely syntactic approach is also impossible, since a numeral NP can be indefinite and non-specific, without yet triggering default agreement (22):

- (21) a. *Prišël/\*prišlo odin student.*  
           came<sub>MSG/NSG</sub> one<sub>M</sub> student<sub>M</sub>  
           ‘There came one student.’  
       b. *Našlas/našlos' odna tonna zerna.*  
           found<sub>REFL.FSG/NSG</sub> one<sub>F</sub> ton<sub>F</sub> grain  
           ‘There was found one ton of grain.’
- (22) *Na každoe tvoe slovo u nego najdutsja pjat', vse necenzurnye.*  
       on each your word at him find<sub>REFL.PL</sub> five all obscene  
       ‘He will find five words to each one of yours, all of them obscene.’

We therefore need a mixed account, which will take into consideration the denotation of the numeral NP (degree vs. entity), its referentiality (i.e., whether it is specifically definite) and also the crucial role played by the presence of a cardinal and its choice.

#### 4.1 The Structure of a Numeral NP

Non-agreeing NP subjects in Russian always contain a cardinal other than *one* or a similar item. Following the tradition, we concentrate mostly on non-agreeing NPs headed by a cardinal, but the class of non-agreeing subjects is in fact larger. Vague numerals, such as *mnogo* 'many', and pseudo-partitives headed by one of certain semantically bleached group nouns, such as *rjad* 'series', *para* 'couple', *kuča* 'heap', etc. (Crockett 1976; for a corpus-based study see Graudina et al. 1976), can also trigger default agreement:

- (23) Bylo namečeno rjad konkretnyx voprosov.  
 was<sub>NSG</sub> sketch<sub>PPT.NSG</sub> series concrete<sub>GEN</sub> questions<sub>GEN</sub>  
 'There was sketched a series of concrete questions.'

Following Ionin and Matushansky 2006, we assume the following structure for a numeral NP (with the exception of numeral NPs containing the cardinal *odin* 'one', which we assume to be adjectival not only in its morphology, but also in its syntax):

- (24)
- 
- ```

graph TD
    NP1[NP] --- N1[N°]
    NP1 --- NP2[NP]
    N1 --- sem[sem']
    NP2 --- N2[N°]
    NP2 --- NP3[NP]
    N2 --- tysjač[tysjač]
    NP3 --- PL[seven thousandPL.GEN]
    NP3 --- GEN[examplesGEN]
  
```

Given these facts, we hypothesize that the crucial syntactic property of non-agreeing subjects is that their head noun is phi-deficient in a way to be made precise below. Cardinals, therefore, are also assumed to be deficient nouns (cf. Ionin and Matushansky 2006). As a result, if a numeral NP headed by a cardinal other than *one*, a vague numeral or a vague group noun denotes in the *d* domain (degrees), it doesn't trigger syntactic phi-agreement, but also, not being a semantic plurality, cannot trigger semantic agreement either. On the other hand, when the same numeral NP denotes in the *e* domain (entities), it denotes an aggregate (i.e., plural) entity and thus can trigger semantic agreement. A formal account of this intuition is unfortunately complicated by the fact that NP-

internal agreement differs from predicate agreement in that the former is unconditional, as we will presently show.

#### 4.2 NP-Internal Agreement: The Case of Modifying APs

A demonstrative, quantifier or adjective higher than the cardinal must show plural agreement with NPs headed by cardinals higher than *one*, irrespective of the denotation of a numeral NP. NPs headed by inherent measure nouns (25) are no exception, and the effect is not limited to adjectives agreeing in case with the entire numeral NP, as the so-called pre-quantifiers (Babby 1987) also appear with plural marking where relevant, regardless of agreement on the verb (26):

- (25) a. Èti/kakie-nibud'/poslednie pjat' let prošli nezametno.  
 these<sub>PL</sub>/any<sub>PL</sub>/last<sub>PL</sub> five years passed<sub>PL</sub> unnoticed<sub>ADV</sub>  
 'These/any/last five years passed unnoticed.'  
 b. \*èto/kakoe-nibud'/poslednee pjat' let  
 this<sub>NSG</sub>/any<sub>NSG</sub>/last<sub>NSG</sub> five years
- (26) a. Prošlo kakix-to/žalkix tri goda.  
 passed<sub>NSG</sub> some<sub>GEN.PL</sub>/meager<sub>GEN.PL</sub> three years  
 'A bare/meager three years passed.'  
 b. \*kakogo-to/žalkogo tri goda  
 some<sub>GEN.SG</sub>/meager<sub>GEN.SG</sub> three years

The two case-marking options on the modifying adjective correlate, we claim, with the denotation of the numeral NP. Case-agreeing adjectives include possessive and demonstrative adjectives, quantifiers (e.g., *some*), ordinals, and sequentials (e.g., *last*, *next*) (Mel'čuk 1985:97). NPs case-agreeing modifiers appear in are clearly entity-denoting. Genitive-marked adjectives, on the other hand, form a closed class,<sup>5</sup> semantically seem to be similar to the modified cardinal construction (*a meager three people*, see Ionin and Matushansky 2004), and the numeral NPs that they

<sup>5</sup> To the best of our knowledge, the list consists of all the indefinite quantifiers (*kakix-to* 'some', *kakix-nibud'* 'any', etc.) and the adjectives *celyx* 'whole', *dobryx* 'good', *(ne)polnyx* '(not)full', *žalkix* 'meager', *nesčastnyx* 'poor', *bityx* 'broken', *lišnix* 'spare' and *dolgix* 'long'. It is the last one, which is only compatible with numeral NPs denoting measures of time and distance, that shows that genitive-marked APs do not specify quantity and therefore do not modify the cardinal itself (contra Babby 1985), but rather combine with the entire numeral NP.

modify systematically provide a measure of the event. Confirming our primary hypothesis, case-agreeing (nominative) APs require plural marking on the verb with cardinals higher than *one*,<sup>6</sup> while genitive APs also allow default agreement. Plural agreement with genitive pre-quantifiers (27b) furthermore appears to give rise to the intuition that the entity corresponding to the provided measure is presupposed to exist, which further links the pre-quantifier construction to degree-denoting NPs, which, as discussed above, appear with plural agreement only when referential:

- (27) a. *Eslikakie-nibud' pjat' let i prošli/\*prošlo nezametno...*  
 if any<sub>PL</sub>/last<sub>PL</sub> five years EMPH passed<sub>PL/\*SG</sub> unnoticed<sub>ADV</sub>  
 'If any five years passed unnoticed [then it was 1915-1920].'  
 b. *Na poxorony prišli/prišlo kakix-to tri čeloveka.*  
 on funeral came<sub>PL/\*SG</sub> some<sub>GEN.PL</sub> three person<sub>PAUC</sub>  
 'A meager three people came to the funeral.'

The question still remains how the two patterns of case-marking can be achieved. Answering it, Babby 1987 proposes that genitive-marked pre-quantifiers combine with the cardinal rather than with the NP as a whole. Several issues arise with the solution.

Our first and primary problem lies in our inability to adopt the hypothesis that a complex cardinal forms a constituent to the exclusion of the lexical NP (see Ionin and Matushansky 2006 for a discussion). But even if this problem didn't arise, Babby's constituency fails to explain a number of facts. The first one is paucal case-assignment: if the cardinal combines with a pre-quantifier, then cardinals have to be introduced as specifiers and therefore cannot assign case – how then can the fact that the lower cardinals surface with the paucal case (generally identical to genitive singular) while the higher cardinals combine with genitive plural, be accounted for? Furthermore, the simple fact of combining with a cardinal does not in any way explain why the pre-quantifier is marked genitive plural – while case-marking can be attributed to the cardinal, the number

---

<sup>6</sup> Crockett 1976:335 also claims that nominative adjectives require plural agreement on the verb and attributes this observation to traditional Russian grammars, but all examples she provides are either definite or universally quantified.

cannot be. Moreover, the same class of lexical items can appear with exactly the same meaning with measure NPs not containing a cardinal:

- (28) *kakoj-to/ žalkij litr vodki*  
*some<sub>SG.NOM</sub>/ meager<sub>SG.NOM</sub> liter<sub>M</sub> vodka<sub>GEN</sub>*  
 ‘a meager liter of vodka’

Finally, as with the modified cardinal construction in English (Ionin and Matushansky 2004), the hypothesis that the pre-quantifier combines with a cardinal does not account for its ability to combine with a conjoined numeral NP, as in (29):

- (29) *žalkix [ dva litra vodki i tri ogurca]*  
*meager<sub>PL.GEN</sub> two liter<sub>PAUC</sub> vodka<sub>GEN</sub> and three cucumber<sub>PAUC</sub>*  
 ‘a meager two liters of vodka and three cucumbers’

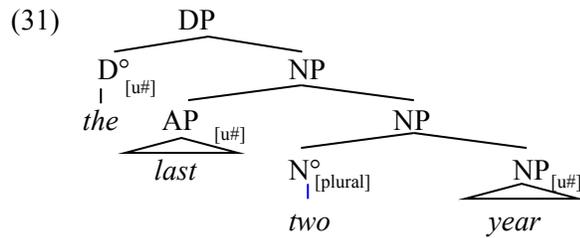
We conclude therefore that the pre-quantifier combines with the entire numeral NP, but is its surface position also its base position? Could it have been merged below the cardinal and then moved? Unfortunately, this hypothesis is also untenable, and not only because it would not give us the correct truth-conditions: as the following example shows, with the paucal cardinal *pol-* an adjective contained in the lexical NP surfaces in the genitive singular, while the pre-quantifier is genitive plural:

- (30) *kakix-to pol-svetovogo goda*  
*some<sub>PL.GEN</sub> half light<sub>ADJ.GEN</sub> year<sub>GEN</sub>*  
 ‘a meager half a light-year’

To link the appearance of a genitive plural pre-quantifier with the degree denotation of the numeral NP we propose that its case is assigned by the null OEMR determiner ensuring the compatibility of a predicative NP with an argument position (see section 2). Assuming that this kind of genitive case is quantificational in nature (cf. Bailyn 2004) explains why it does not surface on the cardinal, which is itself specified [Q], and on its lexical NP sister, and why, like the genitive assigned by cardinals, in oblique-case environments it is replaced by the corresponding oblique.

#### 4.2 NP-Internal Agreement and Syntactic Agreement

Unconditional agreement is obviously unsurprising and it would seem that no special proviso should be made for it. However, as we mentioned before, we assume that cardinals take semantically singular NPs as their complements and project (Ionin and Matushansky 2006).



On the assumption that cardinals combine with semantically singular lexical NPs (Ionin and Matushansky 2006), the first target for number agreement is the lexical NP itself, which we assume to be specified for an uninterpretable and unvalued number feature ([u#]). Since the lexical NP does not c-command the only available source of number in (31), i.e., the cardinal, the standard minimalist probe-goal approach to agreement does not explain how the number feature on the lexical NP can be valued. To circumvent the issue, we hypothesize, extending Béjar 2003 and Rezac 2003, that when an uninterpretable feature does not find a suitable goal in its domain, it can continue to probe upwards. As a result, the unvalued number feature on the lexical NP eventually finds the number feature on the cardinal and is set to [plural] for cardinals higher than *one*.

The next agreement target is the AP. Here also, the locus of unvalued phi-features does not c-command the agreement controller and the same algorithm can be used: the unvalued phi-features on A° do not find a suitable goal locally and can be assumed to continue probing in the sister of the AP, i.e., in the numeral NP itself. The number feature of the AP is then also set to [plural]. The same is true for adjectival demonstratives and quantifiers, irrespective of whether we locate them in D° or not.

NP-internal agreement in number is therefore unconditional and nothing special has been assumed besides a mechanism to deal with agreement in a configuration where the probe (a modifier AP or the complement of the cardinal) does not c-command the goal (i.e., the cardinal). NP-external agreement in number, however, is a different matter.

#### 4.3 NP-External Number Marking and Semantic Agreement

As discussed above, predicate agreement in Russian is unexpectedly sensitive to the denotation of the subject. Were it the number feature on  $T^\circ$  that probed the cardinal, the standard unconditional agreement would be expected. Furthermore, the lack of number agreement obligatorily correlates with the neuter gender, which shows that the gender feature on  $T^\circ$  is also not valued -- in other words, unlike in Miya, numeral NPs in Russian appear to fail to agree altogether -- by hypothesis, when they denote degrees.

To resolve other cases where agreement markers reflect semantic rather than formal properties of the controller, it has been suggested (Corbett 2006) that agreement can be syntactic (in the framework assumed here, resulting from the standard Agree relation between two formal features) or semantic (resulting from an unspecified mechanism that takes into consideration the semantics of the NP controlling agreement). It can then be claimed that NP-external agreement in Russian is semantic and fails when the subject does not have semantic number or gender -- as is the case with degree-denoting NPs.

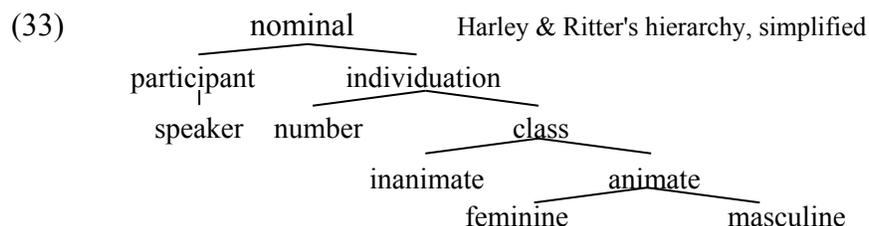
Two problems arise with this hypothesis. On the one hand, semantically singular pluralia tantum nouns trigger plural agreement in Russian:

- (32) Odnī sani/ nožnicy okazalis' lišnimi.  
 one<sub>PL</sub> sleigh<sub>PL</sub> scissors<sub>PL</sub> turn.out<sub>PAST:PL</sub> superfluous<sub>PL</sub>  
 'One sleigh/one pair of scissors turned out to be superfluous.'

On the other, if degree-denoting NPs had no gender feature, it would be unclear why inherent measure NPs can agree for gender in the singular (21b) -- to say nothing about the fact that for inanimate NPs the gender feature is generally considered to be non-semantic. Furthermore, since we have no hypothesis as to how semantic agreement works, attributing default agreement to the failure of semantic agreement does not much advance our understanding. In the next section, therefore, we will sketch our view on how conditioned agreement is achieved.

## 5 An Implementation of Conditioned Agreement

Following the suggestion in Matushansky and Ruys [to appear], we make use of the independently motivated notion of phi-feature hierarchy (33), from Harley and Ritter 2002, and suggest that phi-features are introduced in syntax as structured bundles.

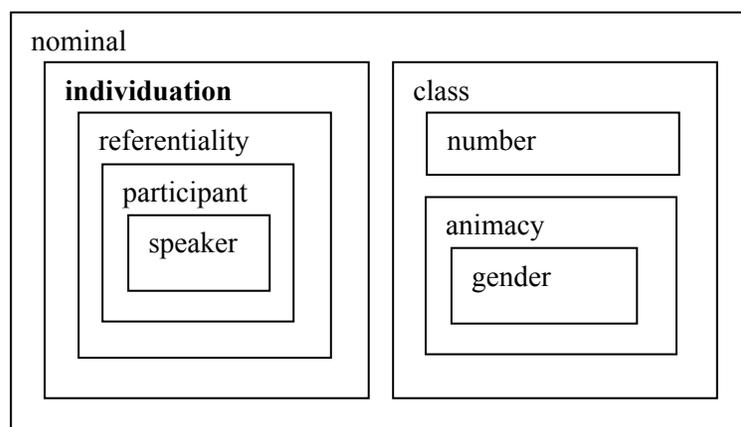


While the hierarchy in (33) has been formulated for phi-features of pronouns, with the presence of a more embedded feature entailing the presence of features that dominate it as a matter of principle,<sup>7</sup> here we propose to encode this by assuming that features may form part of other features. We also extend the Harley and Ritter hierarchy slightly in order to account for the so-called prominence, referentiality, individuation, etc., hierarchies. In particular, while Harley and Ritter label the node containing number and animacy “individuation” and assume that it sorts entities according to their discourse-independent properties, we use for the same purpose the term “class”, and reserve [individuation] for the topmost feature in the set containing specificity in various senses of the term (D-linking, existence presupposition, referentiality, etc.), rigidity (to distinguish demonstrative NPs, pronouns and proper names from all other NPs), definiteness,<sup>8</sup> etc.; individuation distinguishes entities from degrees and is entailed by referentiality, definiteness, etc.:

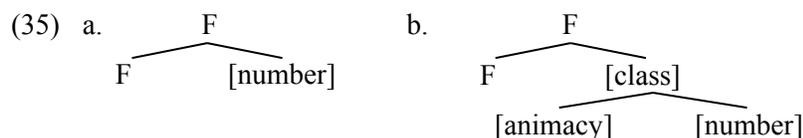
<sup>7</sup> The postulated entailment relation between animacy and gender is far from clear: thus in Russian, gender distinction is not limited to animate NPs in the singular and there are three lexical items (the cardinals *dva/dve* ‘two’ and *poltora/poltory* ‘one and a half’, the quantifier *oba/obe* ‘both’) that distinguish genders in the plural. As this issue is peripheral to our current interests, we leave it for future research.

<sup>8</sup> Definiteness is known to condition case-assignment and agreement; that it is itself a phi-feature is shown by the fact that it can trigger morphologically manifest agreement, both NP-internally (e.g., in Modern Hebrew) and NP-externally (e.g., in Hungarian).

(34)



Crucial for our approach is the hypothesis that one feature can form a structured part of another feature. For instance, the [number] feature may be introduced individually (35a) or as part of the [class] feature bundle (35b):



As a result, syntactic encoding of the feature hierarchy naturally entails a constraint on probing: a phi-feature internal to another feature cannot probe separately (Chomsky 2001). To continue with the same example, the content of the feature [class] reduces to the features [animacy] and [number], which are naturally unvalued in this configuration. The natural consequence of this is a principled way of implementing conditioned agreement: e.g., if  $F^\circ$  is specified for the feature [class] as in (35b), then only NPs that are specified for [animacy] and [number] will trigger  $\phi$ -agreement. Given that [animacy] is a privative feature, only animate NPs will be able to trigger number agreement on  $F^\circ$ .

The conventional unconditional agreement can now be implemented by assuming that the probing head is specified for the number feature only, as we have hypothesized for NP-internal agreement in Russian in the previous section.

To handle conditioned agreement in Russian we now hypothesize that  $T^\circ$  bears not [number] (or [class]) but rather [uN],<sup>9</sup> with the uninterpretable unvalued nominal feature crucially containing the [individuation] feature. An entity-denoting N is [individuated] by virtue of its semantics, while a degree-denoting N is not. As a result, an NP headed by a regular lexical noun, being specified for both [individuation] and [class], triggers full agreement on  $T^\circ$ , while a measure NP, which is specified for [class] only, fails to trigger agreement (since not all of its sub-features find their goals). Adopting the hypothesis advanced in Preminger 2011, we assume that agreement failure does not crash the derivation but rather that the phi-features involved receive their default values.

In addition, the presence of the [individuation] feature is entailed by any of the features embedded under it in the hierarchy in (34), i.e., a definite or specific NP is specified for [individuation] and the valuation of [uN] on  $T^\circ$  is successful.

Finally, implementing the hypothesis that numeral NPs are phi-deficient, we propose that (most) cardinals are deficient nouns: they bear an interpretable number feature obligatorily, but the individuation feature only optionally. We further assume that the presence of this latter correlates with the interpretation in the same way as for regular nouns: an individuated numeral NP denotes in the *e* domain, while a non-individuated one is a degree. The [uN] feature on  $T^\circ$  therefore cannot be valued with the former (yielding default agreement) and must be, with the latter.

The cardinal *one*, known to be adjectival in Russian, fits naturally into this description. As it does not bear interpretable phi-features, it does not represent a possible (defective) goal for the [uN] feature on  $T^\circ$  and thus yields regular agreement, unless the lexical NP that it combines with is itself non-individuated (i.e., headed by an inherent measure noun).

To sum up, postulating the individuation feature distinguishing entities and degrees makes it possible for us to account for the default agreement with degree-denoting subjects in Russian by assuming the "all-or-none" approach to agreement (Chomsky 2001). The particular implementation of this approach that we suggest allows us to use the

---

<sup>9</sup> Given that categorial features are generally supposed to be privative, there would seem to be no such thing as an unvalued nominal feature ([uN]). The hypothesis that a feature may have other features as its content provides a natural way out of this dilemma.

entailment relation between features in order to deal with both phi-feature hierarchies and conditioned agreement.

## 6 Conclusion

In this paper we have argued that agreement with Russian numeral NP subjects is sensitive both to the denotation of these numeral NPs and to their discourse properties: plural marking is possible only with definite or specific NPs or with NPs denoting in the *e* domain.

We have formalized this intuition by a particular implementation of conditioned agreement based on the phi-feature hierarchy proposed by Harley and Ritter. On the assumption that a feature bundle can only be valued as a whole and that a cardinal is phi-deficient in that it does not have to bear the individuation feature we can account for agreement failure with numeral NPs in a principled way that can also be extended to other cases of differential marking and conditioned agreement.

## Bibliography

- Adger, David. 1996. Aspect, agreement and measure phrases in Scottish Gaelic. In *The Syntax of the Celtic Languages*, ed. by Robert D. Borsley and Ian Roberts, 200-222. Cambridge: Cambridge University Press.
- Aissen, Judith. 1999. Markedness and subject choice in optimality theory. *Natural Language & Linguistic Theory* 17:673-711.
- Aissen, Judith. 2003. Differential object marking: iconicity vs. economy. *Natural Language & Linguistic Theory* 21:435-448.
- Babby, Leonard H. 1985. Prepositional quantifiers and the direct case condition. In *Issues in Russian Morpho-syntax*, ed. by Michael S. Flier and Richard D. Brecht, 91-117. Columbus: Slavica.
- Babby, Leonard H. 1987. Case, pre-quantifiers, and discontinuous agreement in Russian. *Natural Language & Linguistic Theory* 5:91-138.
- Bailyn, John. 2004. The case of Q. In *Formal Approaches to Slavic Linguistics #12*, ed. by Olga Arnaudova, Wayles Browne, Maria-Luisa Rivero and Danijela Stojanovic, 1-36. Ann Arbor, Michigan: Michigan Slavic Publications.

- Béjar, Susana. 2003. *Phi-Syntax: A Theory of Agreement*, Doctoral dissertation, University of Toronto.
- Bickel, Baltasar. 2008. On the scope of the referential hierarchy in the typology of grammatical relations. In *Case and Grammatical relations. Papers in honor of Bernard Comrie*, ed. by Greville Corbett and Michael Noonan, 191-210. Amsterdam: John Benjamins.
- Bierwisch, Manfred. 1989. The semantics of gradation. In *Dimensional adjectives*, ed. by Manfred Bierwisch and Ewald Lang, 71-261. Berlin: Springer-Verlag.
- Billings, Loren Allen. 1995. *Approximation in Russian and the Single-Word Constraint*, Doctoral dissertation, Princeton University. Princeton.
- Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: a Life in Language*, ed. by Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.
- Comrie, Bernard. 1981. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- Corbett, Greville G. 2006. *Agreement*. Cambridge: Cambridge University Press.
- Cresswell, M. J. 1976. The semantics of degree. In *Montague Grammar*, ed. by Barbara H. Partee, 261-292. New York: Academic Press.
- Crockett, Dina B. 1976. *Agreement in Contemporary Standard Russian*. Bloomington, Indiana: Slavica Publishers, Inc.
- Etxeberria, Urtzi, and Ricardo Etxepare. 2008. Number agreement with weak quantifiers in Basque. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. by Natasha Abner and Jason Bishop, 159-167. Somerville, Massachusetts: Cascadilla Proceedings Project.
- Etxeberria, Urtzi, and Ricardo Etxepare. 2012. Number agreement in Basque: counting and measuring. In *Noun Phrases and Nominalization in Basque: Syntax and Semantics*, ed. by Urtzi Etxeberria, Myriam Uribe-Etxebarria and Ricardo Etxepare, 149-178. Amsterdam: John Benjamins.
- Gawron, Jean Mark. 2002. Two kinds of quantizers in DP. Paper presented at *LSA Annual Meeting*
- Graudina, Ljudmila Karlovna, Viktor Aleksandrovič Ickovič, and Lia Pavlovna Katlinskaja. 1976. *Grammatičeskaja pravil'nost' russkoj reči. Stilističeskij slovar' variantov*. Moscow: Nauka.

- Harley, Heidi, and Elisabeth Ritter. 2002. Structuring the bundle: a universal morphosyntactic feature geometry. In *Pronouns: Grammar and Representation*, ed. by Heike Wiese and Simon J. Horst, 23-39. Amsterdam: John Benjamins.
- Heim, Irene. 1985. Notes on comparatives and related matters. Ms., University of Texas, Austin.
- Hellan, Lars. 1981. *Towards an Integrated Theory of Comparatives*. Tübingen: Gunter Narr Verlag.
- Ionin, Tania, and Ora Matushansky. 2004. A singular plural. In *WCCFL 23: Proceedings of the 23rd West Coast Conference on Formal Linguistics*, ed. by Benjamin Schmeiser, Vineeta Chand, Ann Kelleher and Angelo Rodriguez, 399-412. Stanford: CSLI.
- Ionin, Tania, and Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23:315-360.
- Klooster, Wim. 1972. *The Structure Underlying Measure Phrase Sentences*. Dordrecht: Reidel.
- Krifka, Manfred. 1990. Four thousand ships passed through the lock: object-induced measure functions on events. *Linguistics and Philosophy* 13:487-520.
- Matushansky, Ora. [to appear]. On Russian approximative inversion. In *Proceedings of FDSL 10, Leipzig, December 5-7, 2013*.
- Matushansky, Ora, and E.G. Ruys. [to appear]. Measure for measure. In *Proceedings of FDSL 10, Leipzig, November 5-7, 2013*.
- Mel'čuk, Igor. 1980a. Animacy in Russian cardinal numerals and adjectives as an inflectional category. *Language* 56:797-811.
- Mel'čuk, Igor. 1980b. O padeže čislovogo vyraženiya v russkix slovosočeta'nix tipa (*bol'she*) na dva mal'čika ili po troe bol'nyx. *Russian Linguistics* 5:55-74.
- Mel'čuk, Igor. 1985. *Poverxnostnyj sintaksis russkix čislitel'nyx vyraženiij*. Wiener slawistischer Almanach. Sonderband 16. Vienna: Institut für Slawistik der Universität Wien.
- Pereltsvaig, Asya. 2006. Small nominals. *Natural Language & Linguistic Theory* 24:433-500.
- Preminger, Omer. 2011. Agreement as a Fallible Operation, Doctoral dissertation, MIT.
- Rezac, Milan. 2003. The fine structure of cyclic Agree. *Syntax* 6:156-182.

- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- Rothstein, Susan, and Keren Khrizman. 2013. Approximative inversion in Russian as a measure construction. Paper presented at *FDSL 10*, Leipzig, December 5-7, 2013.
- Ruys, E.G. 2014. Two Dutch *many*'s and the structure of pseudo-partitives. Ms., UiL OTS/Utrecht University.
- Schuh, Russell G. 1989. Number and gender in Miya. In *Current Progress in Chadic Linguistics: Proceedings of the International Symposium on Chadic Linguistics: Boulder, Colorado, 1-2 May, 1987*, ed. by Zygmunt Frajzyngier. *Current Issues in Linguistic Theory 62*, 171-181. Amsterdam: John Benjamins.
- Schuh, Russell G. 1998. *A Grammar of Miya*. University of California Publications in Linguistics 130. Berkeley: University of California Press.
- Seuren, Pieter A. M. 1973. The comparative. In *Generative Grammar in Europe*, ed. by Ferenc Kiefer and Nicolas Ruwet, 528-564. Dordrecht: D. Reidel.
- Silverstein, Michael. 1976. Hierarchies of features and ergativity. In *Grammatical Categories in Australian Languages*, ed. by R. M. W. Dixon, 112-171. New Jersey: Humanities Press.
- von Stechow, Arnim. 1984. Comparing theories of comparison. *Journal of Semantics 3*:1-77.
- Sussex, Roland. 1976. The numeral classifiers of Russian. *Russian Linguistics 3*:145-155.
- Yadoff, Michael. 1999. Formal Properties of Functional Categories: The Minimalist Syntax of Russian Nominal and Prepositional Expressions, Doctoral dissertation, Indiana University.

o.m.matushansky@uu.nl  
e.g.ruys@uu.nl

## **An Alternative Semantics for Negative Conjunction in Russian\***

*Denis Paperno*  
University of Trento

In this paper, I analyze the semantics of Russian conjunction *ni... ni* ‘neither... nor’. I discuss the problems faced by the analysis in terms of the classical theory of cross-categorical coordination in formal semantics. I then propose a novel compositional account based on alternative semantics. The analysis proposed here is similar to Agafonova’s (2010, 2011) account of Russian *a* ‘and’ motivated by the semantics of gapping constructions.

### **1 Introduction**

*Ni... ni*, which can often be translated into English as *neither...nor*, is a coordinating construction in Russian and other Slavic languages that combines two or more constituents of the same category, whereby each of them is preceded by the clitic *ni*. Coordinate structures with *ni...ni* are negative concord items (NCI), and exhibit all the distributional properties of the class of indefinites formed with *ni* (such as *nikto* ‘nobody’ or *nikogda* ‘never’). *Ni...ni* generally requires an overt *clausemate* negation, which however can be absent in certain specific cases such as fragment answers and existential statements:

---

\* I thank Ed Keenan, Jessica Rett, and anonymous reviewers who provided valuable feedback on this work.

- (1) a. Ni Petja, ni Maša \*(ne) prišli. (*ni...ni*)  
 NI Peter NI Mary not came.  
 ‘Neither Peter nor Mary came.’
- b. Ni-kto \*(ne) prišel. (*ni-word*)  
 NI-who not came  
 ‘Nobody came.’
- (2) a. Kogo ty vybereš’?  
 who you choose  
 ‘Who will you choose?’
- b. Ni togo, ni drugogo. (*ni...ni*)  
 NI that NI other.  
 ‘Neither one nor the other.’
- b’. Ni-kogo. (*ni-word*)  
 NI-who  
 ‘Nobody.’
- (3) a. V “Vojne” – ni gumanizma, ni pacifizma. (*ni...ni*)  
 in War NI humanism NI pacifism  
 ‘[Film] *War* contains no humanism or pacifism.’ [RNC<sup>1</sup>]
- b. V “Vojne” – ni-kakogo podteksta. (*ni-word*)  
 in War NI-which hidden.meaning  
 ‘[Film] *War* contains no hidden meaning.’

Exceptions from the clause mate negation can all arguably be analyzed as ellipsis involving the overt negation marker.

An interesting property of *ni...ni* coordination is that it can include *i* ‘and’ before the last conjunct:

- (4) Nel’zja ni videt’ i ni slyšat’.  
 impossible NI see and NI hear  
 ‘It’s impossible to see or to hear.’ [RNC]

The added *i*, however, doesn’t have any effect on the truth conditions of the sentence, so I ignore those kinds of examples from consideration below when discussing the semantics of the construction.

Besides the local negation requirement, *ni...ni* has the typical syntactic behavior of a conjunction. *Ni...ni* is polymorphic (combines constituents

<sup>1</sup> Russian National Corpus, <http://ruscorpora.ru/search-main.html>.

of various syntactic categories, including but not limited to full sentences, as well as noun, adjective, and verb phrases), it allows Across The Board extraction and occurs in Non-Constituent Coordination and gapping environments.

Semantic analysis of polymorphic coordination is an old problem, and there is a classical solution (Rooth and Partee 1983; Keenan and Faltz 1985; Gazdar 1980), based on the assumption that logical operators denoted by conjunctions distribute over function application (the property of *distributivity*). The basic meanings of connectives *or*, *and*, *neither...nor* are taken to be the corresponding truth functional operators of propositional calculus ( $\wedge$ ,  $\vee$ ,  $\downarrow$ , respectively). Function application relates predicates to (individual) arguments and generalized quantifiers to predicates, so distributivity correctly predicts truth conditional equivalences like the following:

- (5) a. Peter [dances and sings]  $\equiv$  Peter dances and Peter sings  
       b. [Many students and few teachers] sing  $\equiv$  Many students sing and few teachers sing
- (6) a. Peter [dances or sings]  $\equiv$  Peter dances or Peter sings  
       b. [Either many students or few teachers] sing  $\equiv$  Either many students sing or few teachers sing
- (7) Peter is [neither a doctor nor a lawyer]  $\equiv$  It's not the case that Peter is a doctor and it's not the case that Peter is a lawyer.

There are known cases – such as the so-called ‘non-Boolean’ coordination (Krifka 1990) or some instances of disjunction (Alonso-Ovalle 2006) – where the classical theory of polymorphic coordination does not suffice and modifications are required. The case of *ni...ni*, as discussed below, is another instance of a coordination pattern problematic for the classical Boolean theory that can be interpreted in a much more straightforward and enlightening way with the semantic apparatus of alternative semantics (Hamblin 1973).

## 2 Boolean Interpretations for *ni... ni* and the Empirical Challenges They Face

Unified Boolean semantics of coordination could be extended to *ni...ni* in more than one way that is consistent with the interpretation of the simplest

examples like (1). The three major analytic options for *ni...ni* are: a) conjunction AND taking scope above negation, b) disjunction in the scope of negation (OR), or c) negative NOR (as in Doetjes 2005). All three approaches have their problems when applied to *ni...ni*. The complications arise from the scope interaction of conjunction, negation, and quantifiers, because, as we will see below, coordinate phrases with *ni...ni* sometimes seem interpreted above and sometimes below negation.

In simple cases like conjunction of individual-denoting noun phrases the three semantic analyses produce equivalent truth conditions, but use different ingredients. The difference between the three options is illustrated below, where I provide the denotations of two main parts of example (1), *Ni Petja, ni Maša ne prišli*: the coordinate noun phrase and the rest of the sentence. Under the OR analysis I have to provide a somewhat artificial notation with a variable (**Q** in 8 below) for the denotation of the subject NP, since in this analysis negation does not constitute a syntactic and semantic unit with the verb to the exclusion of the NP.

- (8)  $[[Ni\ Petja, ni\ Maša]]$      $[[ne\ prišli]]$
- AND: conjunction outscoping negation  
 $\lambda P.I_p(P) \& I_m(P)$              $\lambda x. \neg \text{come}(x)$
  - OR: disjunction in the scope of negation  
 $\lambda P.I_p(P) \vee I_m(P)$              $\neg \mathbf{Q}(\lambda x. \text{come}(x))$
  - NOR: negative conjunction  
 $\lambda P.(I_p(P))\text{NOR}(I_m(P))$      $\lambda x. \text{come}(x)$

Let us now explore the empirical challenges that each of the approaches faces.

### 2.1 Challenges: The AND Approach

As already mentioned, *ni...ni* may coordinate different syntactic categories and semantic types, which are challenging (and therefore informative) in different ways. Examples where *ni...ni* combines two predicates are problematic for the AND interpretation. Naturally occurring examples of this kind are abundant, compare:

- (9) Pinočet uže ne byl ni prezidentom, ni  
 Pinochet already not was NI president NI  
 glavnokomandujuščim, no eščě ostavalsja senatorom.  
 Commander-in-Chief but still remained senator.  
 ‘Pinochet was no longer the president or the commander in chief but  
 remained a senator.’ [RNC]

If *ni...ni* in (9) were interpreted as a logical AND, the Boolean semantics of cross-categorial coordination predicts *ni prezidentom, ni glavnokomandujuščim* to denote the predicate ‘both president and Commander-in-Chief’. However, the meaning of (9) cannot be paraphrased using that predicate. In particular, (9) does not mean that P. was no longer both president and Commander-in-Chief – actually, he was neither, and a weaker interpretation is just not available for this example. Interestingly, the weaker reading is possible with *i* ‘and’, compare:

- (10) a. Pinočet uže ne byl prezidentom i  
 Pinochet already not was president and  
 glavnokomandujuščim, on byl tol’ko prezidentom.  
 Commander-in-Chief he was only president.  
 ‘Pinochet was no longer the president and the commander in  
 chief, he was just the president.’  
 b. #Pinočet uže ne byl ni prezidentom, ni  
 Pinochet already not was NI president NI  
 glavnokomandujuščim, on byl tol’ko prezidentom.  
 Commander-in-Chief he was only president.

Another difficult fact for the AND analysis is that *ni...ni* coordination can include NPI. The AND hypothesis states that the coordinated constituent is interpreted outside the scope of negation. On the other hand, negative polarity items, such as Russian elements composed of a *wh*-word and the enclitic *libo* (*kto-libo* ‘anybody’, *gde-libo* ‘anywhere’), have to occur in the scope of negation or another downward entailing operator. If a coordinate structure with *ni...ni* contains an NPI, the AND analysis predicts a scope paradox whereby the NPI has to be in the scope of negation and the coordinate structure that contains it has to be outside of the scope of negation! Such examples are indeed possible, and are attested in naturally produced texts:

- (11) Ni Petja, ni kto-libo eščë ne prišël.  
 NI Peter NI anyone else not came  
 ‘Neither Peter nor anyone else came.’
- (12) Odnako ne bylo ni kakogo-libo perevorota, ni  
 however not was NI any coup NI  
 tainstvennogo zagovora.  
 mysterious conspiracy  
 ‘But there was not any coup or a mysterious conspiracy.’ [RNC]

## 2.2 Challenges: the OR Approach

One problematic case for the OR approach is a configuration where each conjunct in the *ni...ni* construction itself contains a negation. The conjuncts can be full negated sentences or negated predicates, compare:

- (13) Ni Petja ne prišël, ni Maša ne pojavilas’.  
 NI Peter not came NI Mary not appeared  
 ‘Peter didn’t come nor did Mary appear.’
- (14) Ni po radio ne prozvučit, ni v viktorinax ne promel’knët  
 NI on radio not sound NI in quizzes not dash  
 korotkaja, nežnaja eë familija.  
 short tender her last.name  
 ‘Her short tender name won’t sound on the radio or dash in quiz shows.’ [Tatjana Tolstaja, *Okkervil*. Quoted from RNC]

The OR approach assumes that *ni...ni* must be under the scope of negation, but in such examples negation (*ne*) is inside each conjunct so it cannot scope above the coordinate structure.<sup>2</sup> In short, coordinated negated constituents create a scope paradox for the OR approach just like coordinated predicates without negation in them do for the AND analysis. For another empirical challenge of the OR approach created by quantified NPs, see below.

---

<sup>2</sup> An anonymous reviewer suggested that *ni...ni* in (14) might “be used for stylistic reasons ... independent from coordination.” However, *ni...ni* in (13) and (14) patterns syntactically with its other usages of doubled conjunctions: it has to be repeated between each conjunct (omitting any single instance of *ni* is impossible), the conjuncts have an equal syntactic status (e.g. inserting a subordinating conjunction like *potomu što* ‘because’ makes the sentence ungrammatical). This is not expected if *ni* were a discourse particle added merely for stylistic reasons.

### 2.3 Challenges: The NOR Approach

Now let us turn to the last of the simple Boolean approaches, the NOR analysis, which postulates that *ni...ni* includes a negative operator in its meaning, making the negative particle *ne* with which it co-occurs semantically vacuous, at least in the context of negative coordination. Therefore, if a sentence happens to contain two instances of the *ni...ni* construction, they feed two negative operators into the semantic interpretation. Therefore, a double negation reading should be possible in sentences with multiple *ni... ni*. Sentences with multiple instances of *ni...ni* are possible:

- (15) a. Ni Petja, ni Maša ne ljubjat ni syr, ni vino.  
 NI Peter NI Mary not like NI cheese NI wine  
 ‘Neither Peter nor Mary like cheese or wine.’
- b. Ni on, ni otec ego ne kopili nenavist’ k  
 NI he NI father his not accumulated hatred to  
 kapitalizmu ni v šaxtax, ni v dymnyx i pyl’nyx  
 capitalism NI in mines NI in smoky and dusty  
 fabričnyx cexax.  
 factory shops  
 ‘Neither he nor his father had been accumulating hate for  
 capitalism in mines or in smoky and dusty factory shops.’  
 [RNC]

For (15a), the NOR approach predicts a double negation reading that can be paraphrased as follows:

- (16) Neither Peter nor Mary is such that they like neither cheese nor wine  
 ≡ Peter likes either cheese or wine, and so does Mary.

In reality, double negation readings are never available for sentences with multiple *ni...ni*.

Finally, coordinated quantified phrases with *ni...ni* constitute yet another kind of a scope problem, compare:

- (17) Ni sekretariat v celom, ni každyj iz ego členov v  
 NI sekretariat in whole NI each of its members in  
 otdel'nosti dlja menja ne avtoritet.  
 separation for me not authority  
 'Neither the sekretariat as a whole nor any [literally: 'each'] of its  
 members individually is an authority for me.'

This is a version of a naturally occurring example, which I simplified for expository purposes:

- (18) Ni ves' sekretariat v celom, ni každyj iz ego členov  
 NI all sekretariat in whole NI each of its members  
 v otdel'nosti ne moguť byť dlja menja avtoritetami ni v  
 in separation not can be for me authorities NI in  
 tvorčeskom, ni tem bolee v nraťstvennom oťnošenií.  
 creative NI that more in moral regard.  
 'Neither the sekretariat as a whole nor any [literally: 'each'] of its  
 members individually can be my authorities in the creative, let alone  
 in moral respect.' [Vladimir Vojnovič. Open letter to the sekretariat  
 of Writers' Union]

Both the OR and the NOR analyses place the conjuncts in the scope of a negative operator (negation of the negative conjunction itself). Therefore the example (18) above is predicted to be paraphrasable as (19):

- (19) 'it is not the case that sekretariat is an authority or that each of sekretariat members is an authority'

Putting aside the unproblematic first conjunct, the second one means

- (20) 'it is not the case that each of sekretariat members is an authority'  $\equiv$   
 $\equiv$  'some sekretariat member is not an authority' ( $\neg > \forall$ ).

In reality, however, (17) has a much stronger implication: no sekretariat member is an authority ( $\forall > \neg$ ). This makes the interpretation of

quantified conjuncts as in (17) problematic for both the OR and the NOR analysis.

Similar examples are found with other quantifiers:

- (21) Ona ne vozmuščaet ni podavljajuščee bol'sinstvo dumsnix  
 it not outrages NI vast majority Duma  
 partij, liš' pered vyborami vspominajuščix o zaščite  
 parties only before elections remembering about advocating  
 interesov izbiratelej, ni drugie obščestvenno-političeskie  
 interests voters NI other social-political  
 sily.  
 forces.  
 'This outrages neither a vast majority of parliamentary parties,  
 which remember about advocating the interests of their voters only  
 during election campaigns, nor other social forces.' [RNC]

vast majority >¬

- (22) Ni on, ni bol'sinstvo členov kollegii takogo ne imejut.  
 NI he NI most members collegium such not have  
 'Neither he nor most collegium members don't have one' [RNC]

most >¬

- (23) Ni emu, ni dvum drugim (Lysomu i Xudomu)  
 NI he NI two others Bald and Thin  
 prisutstvovat' ne razrešili.  
 be.present not allowed  
 'Neither he nor two others (Bald and Thin) were allowed to be  
 present' [RNC]

two >¬

### 3 Proposal

As argued above, the standard cross-categorial logical operator approach with distributivity does not allow us to provide a simple unified account

for all data. Different kinds of examples prove to be problematic for different versions of this analysis, as summarized in Table 1.

|                                                           | OR          | NOR         | AND         |
|-----------------------------------------------------------|-------------|-------------|-------------|
| conjoined predicates                                      | OK          | OK          | problematic |
| NPIs in conjuncts                                         | OK          | OK          | problematic |
| <i>ni...ni</i> conjoining sentences or negated predicates | problematic | OK          | OK          |
| double <i>ni. . . ni</i>                                  | OK          | problematic | OK          |
| quantified NPs                                            | problematic | problematic | OK          |

Table 1: Summary of the data introduced so far.

My proposal relies on the notion of alternative sets. Alternative semantics was originally proposed for the semantic analysis of questions; according to Hamblin's (1973) proposal, the interpretation of a question is a set of alternatives, each of which is a proposition that can be an answer to the question:

(24) [[Is John here?]] = {John is here, John is not here}

For a more modern proposal on the alternative semantics of questions with a wider empirical coverage and a more detailed discussion of the syntax-semantics interface, see Hagstrom (1998).

Alternatives have been useful in the analysis of various semantic phenomena, most notably focus (Rooth, 1985). Indefinites are another prominent class of items for which alternative semantics provides an enlightening account (Kratzer, 2005).

Alternative-based approaches have also been defended for conjunctions: Simons (2005a,b) and Alonso-Ovalle (2006) argued that an analysis in terms of alternatives is more adequate than the classical one for disjunction (*or*); Aloni (2002) relied on alternative-based account of *and* and *or*; Agafonova (2010, 2011) argued for an alternative semantics of Russian *a* 'and' and English *and* based on evidence from gapping constructions.

Following this line of analysis of the semantics of coordination, and motivated by the empirical evidence introduced so far, I propose that

*ni...ni* is not a truth functional operator but rather an alternative-forming operator. In our analysis, *ni...ni* takes two or more conjuncts as arguments and returns a set of Hamblin alternatives that contains their denotations:

$$(25) \quad \lambda x_1 \dots \lambda x_k. \{x_1, \dots, x_k\}$$

The alternative set is then universally closed by a closure operator  $\emptyset_{\forall}$  that always takes scope above the negation.

$$(26) \quad [[\emptyset_{\forall}]] = \lambda P. \forall x \in P. \forall x$$

In the case of sentence coordination, as in example (13), the universal closure operator  $\emptyset_{\forall}$  amounts to conjoining all propositions in the alternative set, which consists of the propositions expressed by the two sentences:

$$(27) \quad [[\text{Ni Petja ne prišel, ni Maša ne pojavilas'}]] \\ = \emptyset_{\forall} \{ \neg \text{came}'(\mathbf{p}'); \neg \text{appear}'(\mathbf{m}') \} = \neg \text{came}'(\mathbf{p}') \& \neg \text{appear}'(\mathbf{m}')$$

In most cases, however, *ni...ni* combines proper subconstituents of a negated sentence rather than full sentences. In those cases, too, the coordinate structure denotes a set of alternatives. For example, *ni Petja, ni Maša* ‘neither Peter nor Mary’ can be interpreted as a set of two alternative objects:

$$(28) \quad [[\text{Ni Petja, ni Maša}]] = \{ \mathbf{p}', \mathbf{m}' \}$$

As is standard in alternative-based approaches since the foundational paper by Hamblin, semantic composition proceeds pointwise with each of the Hamblin-style alternatives:

$$(29) \quad [[\text{Ni Petja, ni Maša ne prišli}]] = \{ \neg \text{came}'(\mathbf{p}'), \neg \text{came}'(\mathbf{m}') \}$$

Finally, the alternative propositions in the set thus obtained are bound by a universal closure operator  $\emptyset_{\forall}$  higher up, which is attached to a phrase or clause with negation:<sup>3</sup>

---

<sup>3</sup> I am leaving open the question of the exact syntactic position of  $\emptyset_{\forall}$ .

$$(29') \quad [[\emptyset_{\forall} \text{ Ni Petja, ni Maša ne prišli}]] = \wedge \{ \neg \text{came}'(\mathbf{p}'), \neg \text{came}'(\mathbf{m}') \} = \\ = \forall p \in \{ \neg \text{came}'(\mathbf{p}'), \neg \text{came}'(\mathbf{m}') \}. \forall p \equiv \neg \text{came}'(\mathbf{p}') \ \& \ \neg \text{came}'(\mathbf{m}')$$

As we see, the alternative-based account predicts the correct truth conditions for the simple case of individual coordination, which was unproblematic to start with. The same compositional mechanism predicts correct interpretations for all the more difficult kinds of examples discussed above. First, let us take the case of coordinated predicates (9):

$$(30) \quad \text{a. } [[\text{ni prezidentom, ni glavnokomandujuščim}]] = \\ = \{ \lambda x. \text{president}'(x), \lambda x. \text{commander-in-chief}'(x) \} \\ \text{b. } [[\emptyset_{\forall} \text{ Pinočet ne byl ni prezidentom, ni glavnokomandujuščim}]] \\ = \emptyset_{\forall} \{ \neg \text{president}'(\mathbf{p}'), \neg \text{commander-in-chief}'(\mathbf{p}') \} = \\ = \neg \text{president}'(\mathbf{p}') \ \& \ \neg \text{commander-in-chief}'(\mathbf{p}')$$

Coordinated quantified phrases are compositionally treated in exactly the same way, giving example (17) the following interpretation:

$$(31) \quad \text{a. } [[\text{ni ves}' sekretariat, ni každyj iz ego členov v otdel'nosti}]] = \\ = \{ \lambda P.P(\text{secretariat}'), \lambda P. \forall x \in \text{secretariat}'. P(x) \} \\ \text{b. } [[\emptyset_{\forall} (\text{ni ves}' sekretariat, ni každyj iz ego členov v otdel'nosti} \\ \text{dlja menja ne avtoritet}]] = \\ = \emptyset_{\forall} \{ \neg \text{authority}'(\text{secretariat}'), \forall x \in \text{secretariat}'. \neg \text{authority}'(x) \} = \\ = \neg \text{authority}'(\text{secretariat}') \ \& \ \forall x \in \text{secretariat}'. \neg \text{authority}'(x)$$

As we saw above, coordinated predicates seem to be interpreted under the scope of negation while coordinated quantifiers seem to be outside the scope of negation, creating a scope paradox for a unified semantic account of *ni...ni* as a boolean operator. But in the framework of alternative semantics this is not a problem because the semantic composition proceeds within each alternative, so predicates are ultimately interpreted under the scope of negation and quantifiers are above, while alternatives are both introduced and bound in exactly the same way in both cases.

Finally, our account (plus the standard assumptions about 'flexible' function application when functions and arguments are alternative sets, Hagstrom 1998) predicts the correct truth conditions without a double negation reading in the case of multiple *ni...ni*.

- (32) a. Ni Petja, ni Maša ne ljubjat ni syr, ni vino.  
 b.  $[[\text{Ni Petja, ni Maša}]] = \{\mathbf{p'}, \mathbf{m'}\}$   
 c.  $[[\text{ni syr, ni vino}]] = \{\mathbf{cheese'}, \mathbf{wine'}\}$   
 d.  $[[\text{ne ljubjat ni syr, ni vino}]] = \{\neg\mathbf{like'}(\mathbf{cheese'}), \neg\mathbf{like'}(\mathbf{wine'})\}$   
 e.  $[[\emptyset_{\vee} \text{Ni Petja, ni Maša ne ljubjat ni syr, ni vino}]] =$   
 $= \emptyset_{\vee} \{\neg\mathbf{like'}(\mathbf{cheese'})(\mathbf{p'}), \neg\mathbf{like'}(\mathbf{wine'})(\mathbf{p'}),$   
 $\neg\mathbf{like'}(\mathbf{cheese'})(\mathbf{m'}), \neg\mathbf{like'}(\mathbf{wine'})(\mathbf{m'})\} = \neg\mathbf{like'}(\mathbf{cheese'})(\mathbf{p'})$   
 $\& \neg\mathbf{like'}(\mathbf{wine'})(\mathbf{p'}) \& \& \neg\mathbf{like'}(\mathbf{cheese'})(\mathbf{m'}) \&$   
 $\neg\mathbf{like'}(\mathbf{wine'})(\mathbf{m'})$

## 5 Conclusions

As argued above, alternative semantics for coordination allows us to achieve a unified semantic analysis of the negative conjunction *ni... ni* in Russian. The analysis is fully compositional, albeit it is a bit more complex than the classic analysis because the meaning of coordination is essentially distributed between the coordinator itself, which introduces alternatives, and the empty operator that binds them. The alternative-based analysis is adequate for coordination of NPs, predicates, full sentences, etc., and unlike the simple proposals discussed in the earlier part of the paper it is capable of unifying the various phenomena that prove to be problematic for different versions of the classical cross-categorial semantic account.

In addition to being elegant and empirically adequate, the alternative-based account of *ni...ni* paves way for further exciting developments. I will only briefly sketch those now, leaving their detailed elaboration for future research.

First, although we only discussed the *ni...ni* pattern on its own right, it will be interesting to explore the relation between alternatives introduced by *ni...ni* and focus alternatives: are those types of alternatives the same or different? Note that in traditional grammar *ni...ni* has been characterized as ‘emphatic’, just like focus particles are. Indeed, as the Russian Academic grammar puts it, *ni...ni* ‘combines the coordination function with that of an emphatic [= focus, D.P] particle’ (AG80, p. 3114).

The perceived ‘emphasis’ of *ni...ni* and other ‘doubled’ conjunctions (e.g. Russian *i...i* ‘both...and’, *ili...ili* ‘either...or’), as formalized in an alternative-based analysis, could potentially explain known grammatical peculiarities of ‘doubled’ conjunctions, see for instance the discussion of *i...i* in Progovac (1999).

Last but not least, alternative semantics can be a natural way of unifying negative conjunction *ni...ni* and negative concord items (*ni*-words). An alternative-based analysis of negative concord elements, treated as a type of indefinites, is implicitly suggested already by Kratzer (2005). Kratzer, who argued for analyzing indefinites within the framework of alternative semantics, cites (although without further analysis) the Latvian paradigm of *ne*-words, which are an analog of Slavic *ni*-words. Indeed, if *nikto* ‘nobody’ denotes the set of all animate objects  $\{x | \mathbf{animate}'(x)\}$ , then the universal closure operator  $\emptyset_{\forall}$  that we have been relying on for interpreting sentences with *ni...ni* also predicts correct interpretation for sentences with *nikto*:

(33)  $[[\emptyset_{\forall} \text{Nikto ne prišēl}]] = \emptyset_{\forall} \{\neg \mathbf{came}'(x) | \mathbf{animate}'(x)\}$ ,

which amounts to the truth of all propositions of the form  $\neg \mathbf{came}'(x)$  where  $x$  is animate.

Full consequences of a unified alternative-based semantic account of *ni...ni* coordination and *ni*-words are yet to be explored. Let me only point out an interesting type of data that such an account should provide a full analysis of, namely coordinate structures where some conjuncts are marked with the coordinating *ni* while others are *ni*-words, usually accompanied by conjunction *i*:

(34) Ne prišēl ni vrač i nikto iz administracii.  
 not came NI doctor and nobody of administration.  
 ‘Neither the doctor nor anyone of administrative staff came.’  
 [A.Solženicyn, *Arxipelag GULag*, cited from RNC]

- (35) Odnako ni Dinara, ni Evgenij i nikto iz drugix učastnikov  
 but NI Dinara NI Evgenij and nobody of other participants  
 amerikanskogo pervenstva, vyšedšix vo vtoroj krug, v svoix  
 American competition exit in second round in their  
 startovyx matčax ne stolknulsja s takim dikim  
 starting matches not faced with such wild  
 soprotivleniem, kakoe prišlos' preodolevat' Maratu.  
 resistance that had to overcome Marat  
 'But neither Dinara, nor Evgenij, nor any other participant of the  
 American competition who entered the second round, had to face in  
 their starting matches the wild resistance that Marat had to  
 overcome.' [RNC]

To summarize, this paper proposes a compositional semantic analysis of *ni...ni* coordination in Russian. General semantic composition principles for alternatives, as proposed by other authors to account for different phenomena, elegantly and without any additional stipulations accommodate various kinds of examples that are problematic for other analyses. Furthermore, semantics of alternatives allows to relate *ni...ni* coordination to other phenomena whose affinity to *ni...ni* is intuitively obvious but has not yet been adequately formalized, including focus and quantificational negative concord items.

## References

- Agafonova, Irina. 2010. On Syntax-Semantics of Gapping Constructions. Paper presented at Chicago Linguistic Society 46.
- Agafonova, Irina. 2011. *On Syntax, Semantics, and Computation in Coordination*. PhD dissertation, Michigan State University.
- AG80: Natalia Shvedova, ed. 1980. *Russkaja grammatika*. Moscow.
- Aloni, Maria. 2002. Free Choice in Modal Contexts. In *Proceedings of Sinn und Bedeutung 7*, ed. M. Weisgerber. University of Konstanz. 25-37.
- Alonso-Ovalle, Luis. 2006. *Disjunction in Alternative Semantics*. PhD dissertation, University of Massachusetts Amherst.
- Gazdar, Gerald. 1980. A Cross-Categorical Semantics for Coordination. *Linguistics and Philosophy* 3: 407-409.

- <http://dx.doi.org/10.1007/BF00401693>.
- Hagstrom, Alan. 1998. *Decomposing Questions*. PhD dissertation, Massachusetts Institute of Technology.
- Hamblin, Charles. L. 1973. Questions in Montague English. *Foundations of Language* 10: 41-53.
- Keenan, Edward Louis and Leonard M. Faltz. 1985. *Boolean Semantics for Natural Language*. Synthese language library, D. Reidel Publishing Company.
- Kratzer, Angelika. 2005. Indefinites and the Operators They Depend On: From Japanese to Salish. In *Reference and quantification: the Partee effect*, eds. Gregory N. Carlson and Francis Jeffry Pelletier. CSLI lecture notes, CSLI Publications. 113-142.
- Krifka, Manfred. 1990. Boolean and Non-Boolean And. In *Papers from the Second Symposium on Logic and Language*, eds. László Kálmán and László Polos. Akadémiai Kiadó, Budapest 1990, 161-188.
- Progovac, Liljana. 1999. Conjunction Doubling and 'Avoid Conjunction' Principle. In *Topics in South Slavic Syntax and Semantics*, eds. Mila Dimitrova-Vulchanova and Lars Hellan, Vol. 172 of Current Issues in Linguistic Theory, 25-40. John Benjamins Publishing Company.
- Rooth, Mats and Barbara Hall Partee. 1983. Generalized Conjunction and Type Ambiguity. In *Meaning, Use, and Interpretation of Language*, eds. Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow, 361-383. Walter de Gruyter.
- Rooth, Mats Edward. 1985. Association with Focus (Montague Grammar, Semantics, Only, Even). PhD dissertation, University of Massachusetts, Amherst. Electronic Doctoral Dissertations for UMass Amherst. Paper AAI8509599.  
<http://scholarworks.umass.edu/dissertations/AAI8509599>.
- Simons, Mandy. 2005a. Dividing Things Up: The Semantics of Or and the Modal/Interaction. *Natural Language Semantics* 13: 271-316.
- Simons, Mandy. 2005b. Semantics and Pragmatics in the Interpretation of Or. In *Proceedings of SALT XV*, ed. E. Georgala & J. Howell, 201-222. Ithaca, NY. Cornell University. Cornell University.

## **Coordinated *Wh*-words in Polish: Monoclausal or Multiclausal?**

*Agnieszka Patejuk*

Institute of Computer Science of the Polish Academy of Sciences

### **1 Introduction**

This paper discusses the issue of how conjuncts in constructions known as lexico-semantic or hybrid coordination should be represented in Polish.<sup>1</sup> It concentrates on examples featuring *wh*-words such as (1), while examples where conjuncts are non-adjacent (as in *Kto zaufał i komu?* where the last conjunct is placed after the verb) remain outside of its scope.

- (1) Kto i komu zaufał?  
wh<sub>NOM</sub> and wh<sub>DAT</sub> trusted  
'Who trusted whom?'

Monoclausal analyses argue that conjuncts belong to the same clause, which means that (1) is treated similarly to multiple questions (as in (2)). By contrast, multiclausal analyses (often referred to in the literature as biclausal, though more than two clauses may be involved) treat (1) as consisting of two different clauses, as in (3) which consists of two distinct questions (see § 2 for discussion of different analyses).

- (2) Kto komu zaufał?  
(3) Kto zaufał i komu zaufał?

---

<sup>1</sup> This paper is a revised and extended version of Patejuk and Przepiórkowski 2014.

A range of tests proposed for languages which use lexico-semantic coordination (these include, apart from Polish, also Hungarian, Romanian and Russian) is presented and the possibility of applying these tests to Polish is discussed. It is perhaps worth noting that this paper is largely based on authentic data – the National Corpus of Polish (NKJP; <http://nkjp.pl/>; Przepiórkowski et al. 2011, 2012), the largest annotated corpus of Polish containing over 1.8 billion segments, served as the main source of data for discussing representation tests.

## 2 Multiclausal Analyses

This paper takes into consideration two recent multiclausal analyses of lexico-semantic coordination, namely that of Tomaszewicz 2011a and Citko and Gračanin-Yüksek 2013. These analyses were chosen because they were originally applied to Polish data and they were also accompanied by an explanation of the resulting representation.

Although the analyses of Tomaszewicz 2011a and Citko and Gračanin-Yüksek 2013 differ with respect to formal devices they use – Tomaszewicz 2011a operates with the notion of deletion (under identity) while Citko and Gračanin-Yüksek 2013 use multidominance (sharing certain branches by clauses) – the intuition behind these analyses is strikingly similar: both offer two-fold analyses which cover nearly equivalent cases.

The first case in Tomaszewicz 2011a ( $T_1$ ) is the one where two conjuncts belong to two distinct clauses and missing arguments (if any) are filled using implicit pronouns – the representation provided in (5)<sup>2</sup> corresponds to (4), more precisely to its single pair reading, while the interpretation accompanying this analysis is provided in (6).

- (4) Kto i co kupić?  
 who and what bought  
(Tomaszewicz 2011a: ex. (1))
- (5) [ who [ $TP$  ~~who~~ bought something ] ] & [ what [ $TP$  *pro* bought ~~what~~ ] ]  
(Tomaszewicz 2011a: ex. (4a))
- (6) Who bought something? And what did they buy?  
(Tomaszewicz 2011a: ex. (4b))

<sup>2</sup> Mismatched brackets (closing bracket missing) in (5) following Tomaszewicz 2011a.

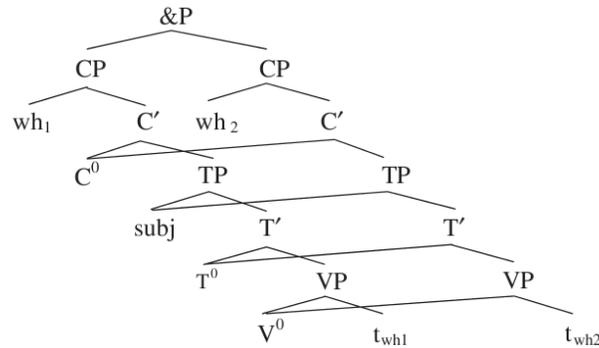


Figure 1: non-bulk sharing structure (Citko and Gračanin-Yüksek 2013: ex. (6c))

While pair list readings are typical of multiple questions (such as ‘Who left when?’) which ask about more than one variable (subject, time) and expect answers consisting of a list of pairs (‘Mary left yesterday, John left two days ago.’), single pair readings are typical of single questions which ask about only one variable (‘Who left yesterday?’) and require single answers (‘Mary left yesterday.’). Since both questions in (5) are single questions (this is reflected in the interpretation in (6)), Tomaszewicz 2011a uses this strategy for obtaining the single pair reading.

The counterpart of this strategy is what Citko and Gračanin-Yüksek 2013 call the non-bulk sharing strategy ( $CGY_1$ ) where “*wh*-words are NEVER shared between the two  $CP$ s (while everything else in the structure is)” – their representation is shown in Figure 1.

The second case used by Tomaszewicz 2011a ( $T_2$ ) involves a coordination of two questions: a single question in the first clause (containing the *wh*-word corresponding to the first conjunct of lexico-semantic coordination) and a multiple question in the second clause (containing both *wh*-words). As explained in Tomaszewicz 2011a “the two identical *wh*-phrases in the two conjuncts undergo ATB movement”, while the second *wh*-phrase stays in the second clause. (7) provides the representation (including implicit pronouns) of the pair list reading of (4), its interpretation is provided in (8). Please note that both include two questions:

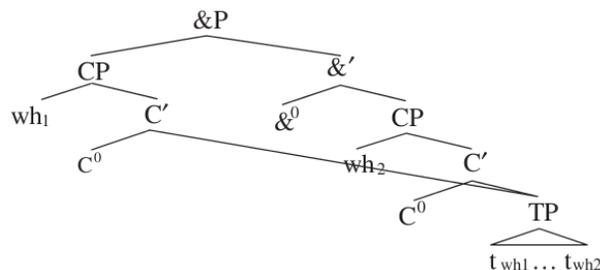


Figure 2: bulk sharing structure (Citko and Gračanin-Yüksek 2013: ex. (6b))

- (7) [ who [ ~~who~~ [<sub>TP</sub> ~~who~~ bought something ] ] & [ ~~who~~ what [<sub>TP</sub> ~~who~~ bought ~~what~~ ] ] ] (Tomaszewicz 2011a: ex. (6a))
- (8) Who bought something? And who bought what? (Tomaszewicz 2011a: ex. (6c))

Citko and Gračanin-Yüksek 2013 use the bulk-sharing strategy (CGY<sub>2</sub>) which is a near counterpart of the pair list multiclausal analysis of Tomaszewicz 2011a. Under this analysis “there IS a point in the derivation when the two *wh*-phrases belong to both CPs, even though in the final representation each *wh*-phrase occupies a specifier of a different CP”, see Figure 2 for an illustration. The difference with respect to the analysis of Tomaszewicz 2011a is that only one of the clauses, the second one, contains a multiple question, while in the analysis of Citko and Gračanin-Yüksek 2013 this is the case with both clauses. As a consequence, the multidominance bulk sharing analysis does not use implicit arguments as there are no missing arguments.

### 3 Representation Tests

#### 3.1 Sentence-level Adverbs

Tomaszewicz 2011a claims that lexico-semantic coordination is multiclausal due to the fact that it is possible to split the conjuncts using a sentence-level adverbial. After providing examples from Bulgarian, she notes: “In Polish the facts are exactly the same as in Bulgarian and the speaker-oriented adverbs include: *najważniejsze* ‘most importantly’,

*zwłaszcza* ‘importantly’, *niestety* ‘unfortunately’, *na szczęście* ‘fortunately’, *o dziwo* ‘surprisingly’.”

For each of the “sentence-level adverbials” listed above, a counter-example was found in NKJP – examples provided below feature such adverbs inside plain NP coordination:

- (9) Zdemontowane zostały [piece, maszyny i co najważniejsze removed were furnaces machines and what most important pompy].  
pumps  
‘Furnaces, machines and, what is most important, pumps were removed.’ (NKJP)
- (10) W domu po prostu zamęczał [matkę i zwłaszcza ojca] [...] at home simply pestered mother and especially father  
‘At home he would simply pester his mother and especially his father.’ (NKJP)
- (11) Z tego tytułu zachowanie [posłów SLD i niestety posłów PSL] jest wyjątkowo złe, naganne [...] for this reason behaviour MPs SLD and unfortunately MPs PSL] is particularly bad reprehensible  
‘For this reason the behaviour of SLD MPs and, unfortunately, PSL MPs is particularly bad, reprehensible.’ (NKJP)
- (12) Dali mi [trochę forsy i na szczęście samochód] gave me some money and fortunately car  
‘They gave me some money and, fortunately, a car.’ (NKJP)
- (13) [Włochy, Grecja, Francja, Niemcy i o dziwo Węgry wraz ze Słowacją] są nastawione prorosyjsko. Italy Greece France Germany and surprisingly Hungary with Slovakia are disposed pro-Russian  
‘Italy, Greece, France, Germany and, surprisingly enough, Hungary with Slovakia have pro-Russian attitude.’ (NKJP)

Examples provided above contain precisely the “sentence-level adverbials” listed by Tomaszewicz 2011a – they clearly can occur in plain coordination where particular conjuncts correspond to the same grammatical function.

It is dubious whether these examples should be analysed as multiclausal simply because a “sentence-level adverbial” is present – examples where such adverbials are placed between conjuncts of the coordinate subject seem to provide strong counterevidence to such claims as the verb displays plural agreement, which would be unexpected under the multiclausal analysis. Examples where the subject is split using such adverbials include (9) and (13), though an example with singular conjuncts could make a stronger argument, see constructed (14), a modified version of (13):

- (14) [Francja i o dziwo Słowacja] są nastawione prorosyjsko.  
 Frances<sub>SG</sub> and surprisingly Slovakia<sub>SG</sub> are disposed<sub>PL</sub> pro-Russian

It seems more likely that “sentence level adverbials” may split conjuncts, making coordination discontinuous. The adverb placed between conjuncts may be analysed as a modifier of the relevant verb.

### 3.2 Clausal Coordinators

Tomaszewicz 2011a claims that *a* is a strictly clausal coordinator in Polish (“*a* never conjoins constituents smaller than a full clause”) and provides the following examples in support of this claim:

- (15) Kto a najważniejsze co mówił o tobie?  
 who and most importantly what said about you  
 ‘Who said something about you and what did they say?’  
 (Tomaszewicz 2011a: ex. (11))
- (16) Jan i/\*a Maria  
 Jan and Maria  
 (Tomaszewicz 2011a: ex. (12a))
- (17) wąski i/\*a długi mostek  
 narrow and long bridge  
 (Tomaszewicz 2011a: ex. (12b))

First, please note that (15) is ungrammatical without *najważniejsze* – Tomaszewicz 2011a mentions in footnote 3 that “the adverb is needed here since “a” is contrastive, and the adverb provides the needed contrast”.

- (18) Kto a \*(najważniejsze) co mówił o tobie?  
 who and most importantly what said about you

If so, the judgement in (16) is controversial – the isolated fragment does not prove that NPs cannot be coordinated using *a*. Maybe it could be improved using some adverbials, as in constructed (19) and authentic (20):

- (19) Jan i/\*a zwłaszcza Maria głośno chrapią.  
 Jan and especially Maria loudly snore  
 ‘Jan and especially Maria snore loudly.’
- (20) Życie, a zwłaszcza śmierć Angeliki de Sancé  
 life and especially death Angelika de Sancé  
 ‘Life and especially death of Angelika de Sancé’ (Google)

Secondly, the judgement provided in (17) seems to be wrong when confronted with corpus data – there are numerous examples in NKJP where it is used to coordinate adjectives in a contrastive manner, so again it does not follow that *a* is an exclusively clausal coordinator:

- (21) Latem umarł ksiądz Józef Tischner – piękny człowiek  
 summer died reverend Józef Tischner beautiful human  
 i piękny mężczyzna: [[wrażliwy i delikatny] a mocny],  
 and beautiful man sensitive and delicate but strong  
 [uśmiechnięty a poważny], [[bardzo mądry] a prosty].  
 smiling but serious very wise but simple  
 ‘This summer reverend Józef Tischner passed away – a beautiful human and a beautiful man: sensitive and delicate yet strong, smiling yet serious, very wise yet simple.’ (NKJP)
- (22) [...] padł ofiarą sprytnego podstępu i przy pomocy [zręcznych  
 fell victim cunning trick and with help clever  
 a fałszywych] argumentów został nakłoniony do udziału [...]  
 but false arguments was persuaded to participation  
 ‘He fell victim to a cunning trick and using clever yet false arguments he was persuaded to take part...’ (NKJP)

The examples provided above show that the conjunction *a* may be used in Polish in plain AP coordination – there is no reason to claim that such examples are multiclausal. As a result, such examples provide evidence against the claim that *a* is a strictly clausal coordinator in Polish and that structures which contain it, such as the lexico-semantic coordination example in (15), must be multiclausal.

### 3.3 Distribution of Question Particles

Tomaszewicz 2011a mentions a test based on the distribution of question particles and discusses it using data from Romanian, cited after Rațiu 2009:

- (23) Oare cine ce va spune?  
PART who what will say

(Tomaszewicz 2011a: ex. (13a))

- (24) Oare cine \*(și) oare ce va spune?  
PART who and PART what will say

‘Who will say something and what will he say?’

(Tomaszewicz 2011a: ex. (13b))

As the examples show, in Romanian the question particle *oare* can be used only once with multiple *wh*-questions, but it can be used with each conjunct under lexico-semantic coordination. This suggests that while multiple *wh*-questions are monoclausal, the structure of lexico-semantic coordination of *wh*-words is multiclausal in Romanian.

While the test itself seems convincing, it cannot be applied directly to Polish as it does not use any question particle which could be a counterpart of the Romanian one. However, it is possible to consider the distribution of other elements which may occur only once per clause in Polish. Potential candidates include mood markers such as *BY* (conditional) and *NIECH* (imperative), the reflexive marker *SIE*, the negative particle *NIE* and agglutinate forms of the verb *BYĆ* ‘be’ (such as *-ś* in *Coś zrobił?* ‘What have you done?’). To verify whether items listed above may be used with each *wh*-word conjunct under lexico-semantic coordination, the following base query was used for searching NKJP:

- (25) [base="kto|co|gdzie|jak|kiedy"  
& (case=\$1 | case!=".\*")] VAR i  
[base="kto|co|gdzie|jak|kiedy"  
& (case!=\$1 | case!=".\*")] VAR

VAR is a variable which is to be substituted (twice) for a relevant query element from the list provided above (mood marker, reflexive marker, negative particle, agglutinate verb form). The query fragment [base="kto|co|gdzie|jak|kiedy" & (case=\$1 | case!=".\*")] matches a segment whose base form is KTO

‘who’, CO ‘what’, GDZIE ‘where’, JAK ‘how’ or KIEDY ‘when’ (base="kto|co|gdzie|jak|kiedy"), its case value is assigned to the \$1 variable (case=\$1) or it has no case at all (case!=".\*"). While [base="kto|co|gdzie|jak|kiedy" & (case!=\$1 | case!=".\*")] matches the same base forms, it requires that the case of this segment is not the same as the one assigned to the variable \$1 (case!=\$1) or that the segment has no case whatsoever (case!=".\*").

A sample query resulting from substituting by for VAR in (25) is provided below:

```
(26) [base="kto|co|gdzie|jak|kiedy"
      & (case=$1 | case!=".*")] by i
      [base="kto|co|gdzie|jak|kiedy"
      & (case!=$1 | case!=".*")] by
```

Some constructed examples that would match the query in (26):

```
(27) *Kto by i kogo by uderzył?
      whNOM COND and whACC COND hit
      ‘Who would hit whom?’ (intended)
```

```
(28) *Kto by i kiedy by uderzył?
```

```
(29) *Gdzie by i kiedy by uderzył?
```

The following table provides a short summary of the results of relevant NKJP queries (the entire corpus, NKJP1800M, was searched for results):

| (30) <b>variable used in (25)</b> | by | niech | się | nie | [pos=aglt] |
|-----------------------------------|----|-------|-----|-----|------------|
| <b>NKJP results</b>               | 0  | 0     | 0   | 0   | 0          |

The results summarised in (30) show that there is no evidence that it is possible to use elements which normally occur only once per clause (markers, particles, clitics) in Polish with each element of questions featuring lexico-semantic coordination. This suggests that there is no evidence supporting the multiclausal analysis.

While it might be the case that counterexamples exist in larger text collections, constructed examples, such as in (27)–(29), are unacceptable.

### 3.4 Auxiliary between *Wh*-phrases

Bílbíie and Gazdik 2012 advance an argument based on the following example from Hungarian:

- (31) Mit akarunk és hol vacsorázni?  
 what want<sub>3</sub> and where eat for dinner  
 ‘What do we want to eat for dinner and where?’  
 (Bílbíe and Gazdik 2012: ex. (48))

They claim that “it can be argued that *akar* ‘want’ is an auxiliary in Hungarian” because “it can interrupt the infinitive following it and appear between the verbal particle (if there is one) and the verbal stem” – if it is assumed, the argument goes, that the auxiliary and the main verb must belong to the same clause, it follows that the structure of such examples must be monoclausal.

Let us see how this test can be applied to Polish:

- (32) Kto będzie komu pomagać?  
 wh<sub>NOM</sub> AUX wh<sub>DAT</sub> help
- (33) Komu będzie kto pomagać?  
 wh<sub>DAT</sub> AUX wh<sub>NOM</sub> help
- (34) \*Kto będzie i komu pomagać?  
 wh<sub>NOM</sub> AUX and wh<sub>DAT</sub> help
- (35) \*Komu będzie i kto pomagać?  
 wh<sub>DAT</sub> AUX and wh<sub>NOM</sub> help

Judgements for sentences (32)–(35) seem to find support in NKJP: the query [base=kto & case=\$1] [pos=bedzie] i [base=kto & case!=\$1] matches the coordination of two *wh*-words whose lemma is KTO ‘who’, requiring that the case of the first conjunct (assigned to the variable \$1: case=\$1) must be different than the case of the second conjunct (case!=\$1). This query yielded 0 results in the entire corpus. By contrast, its modified version with the conjunction removed did return one good result (more were found using Google), supporting judgements which accept splitting multiple *wh*-questions with an auxiliary:

- (36) [...] chodzi jednak o to, kto będzie komu służył  
 matters still about this wh<sub>NOM</sub> AUX wh<sub>DAT</sub> serve  
 ‘It’s more about who will serve whom.’ (NKJP)
- (37) Pytanie kto będzie kogo sponosał. ?  
 question wh<sub>NOM</sub> AUX wh<sub>ACC</sub> pay  
 ‘The question is who is going to pay whom.’ (Google)

It is not clear, however, how the results of this test should be interpreted, especially when other contexts are considered with respect to whether they allow being split with an auxiliary:

- (38) Janek i Marysia będą biegli.  
 Janek and Marysia AUX<sub>3,PL</sub> run<sub>3,PL.M1</sub>  
 ‘Janek and Marysia will run.’

- (39) \*Janek będą i Marysia biegli.  
 Janek AUX<sub>3,PL</sub> and Marysia run<sub>3,PL.M1</sub>

Taking these examples into consideration, it seems to be the case that coordinate phrases in Polish do not allow being split by an auxiliary in general. Lexico-semantic coordination also features a coordinate phrase, so the fact that splitting these with an auxiliary results in ungrammaticality is expected and it is caused by reasons independent of whether such constructions are monoclausal or multiclausal.

### 3.5 Overt Pronouns

Kazenin 2001 advances an argument in favour of the monoclausal analysis on the basis of coreference effects with overt pronouns:

- (40) [Kogo<sub>i</sub> Petja izbil] i [za što Petja ego<sub>i</sub>/??pro<sub>i</sub> izbil]?  
 whom Peter beat and for what Peter him beat  
 ‘Whom did Peter beat and what for did Peter beat him?’  
 (Kazenin 2001: ex. (50))

- (41) \*Kogo<sub>i</sub> i za što Petja ego<sub>i</sub> izbil?  
 whom and for what Peter him beat  
 (Kazenin 2001: ex. (52))

Kazenin 2001 notes that under a coordination of two questions such as in (40) the *wh*-word in the first clause (*Kogo*) may be coreferential with an implicit pronominal (*pro*) or an overt pronoun (*ego*) in the second clause. By contrast, an overt pronoun coreferential with one of the conjuncts cannot be used under lexico-semantic coordination, as shown in (41). According to Kazenin 2001, this suggests that the structure of such examples is monoclausal – unlike in (40), *Kogo* and *za što* belong to the same clause, which makes it impossible to use an overt pronoun as the object of *izbil* – this argument position is already filled by *Kogo*.

This test is applicable to Polish and the facts are similar to Russian:

- (42) Kogo i za co Piotr (\*go) zbil?  
 who<sub>ACC</sub> and for what Piotr he<sub>ACC</sub> beat  
 ‘Who did Piotr beat and what did Piotr beat him for?’

However, this test does not exclude multiclausal analyses (see § 2): under  $T_1$  and  $CGY_1$  null pronouns could be claimed to block the use of lexical pronouns, while under  $T_2$  there is a multiple *wh*-question in the second clause and  $CGY_2$  uses multidominance with two multiple *wh*-questions. On the other hand, there seems to be no constructive evidence which would support using these multiclausal analyses instead of the monoclausal one.

### 3.6 Left Branch Extraction (LBE)

Tomaszewicz 2011b proposes a syntactic argument in support of the multiclausal analysis based on the unavailability of a certain type of extraction in this environment. The argument is that while left branch extraction is grammatical with multiple questions (see (43)) which are monoclausal, Tomaszewicz 2011b claims that such extraction is ungrammatical when lexico-semantic coordination is involved (compare (44)) and attributes this alleged contrast in grammaticality to the fact that the structure of lexico-semantic coordination is multiclausal.

- (43) Jaki kto kupił samochód swojej żonie?  
 which<sub>ACC</sub> wh<sub>NOM</sub> bought car<sub>ACC</sub> SELF<sub>DAT</sub> wife<sub>DAT</sub>  
 (Tomaszewicz 2011b: ex. (27a))
- (44) \*Jaki i kto kupił samochód swojej żonie?  
 which<sub>ACC</sub> and wh<sub>NOM</sub> bought car<sub>ACC</sub> SELF<sub>DAT</sub> wife<sub>DAT</sub>  
 (Tomaszewicz 2011b: ex. (27b))

However, judgements in Tomaszewicz 2011b are dubious – counter-examples may be found in the literature discussing similar phenomena:

- (45) Jakie i skąd zdobywał informacje?  
 what<sub>ACC</sub> and from where obtained information<sub>ACC</sub>  
 ‘What information and where from did he obtain?’  
 (Kallas 1993: p. 141, ex. (108))

Moreover, numerous attested examples may be found:

- (46) Jakie i kto miał rzucane kłody pod nogi?  
 what<sub>ACC</sub> and who<sub>NOM</sub> had thrown logs<sub>ACC</sub> under legs  
 ‘Who has been put what obstacles in their way?’ (NKJP)
- (47) Czy wiadomo jaki i kto będzie grał  
 PART known what<sub>ACC</sub> and who<sub>NOM</sub> AUX play  
 szwarzcharakter?  
 villain<sub>ACC</sub>  
 ‘Do we know who is going to play which villain?’ (NKJP)
- (48) Jakie i kto podjął w tej sprawie działania?  
 what<sub>ACC</sub> and who<sub>NOM</sub> took in this matter actions<sub>ACC</sub>  
 ‘Who took what action in this matter?’ (Google)
- (49) Jakie i kto może ponieść konsekwencje?  
 what<sub>ACC</sub> and who<sub>NOM</sub> can bear consequences<sub>ACC</sub>  
 ‘Who can suffer what consequences?’ (Google)

Since the examples listed above provide rich counterevidence to the judgements of Tomaszewicz 2011b, the conclusion drawn on the basis of her judgements does not hold – there is no contrast in grammaticality between LBE with multiple *wh*-questions and under lexico-semantic coordination. Therefore there is no reason to claim that the latter is multiclausal.

It must be noted, however, that undermining the argument of Tomaszewicz 2011b does not provide strong, constructive evidence in support of the monoclausal representation of lexico-semantic coordination: LBE is possible in this environment, whatever the representation.

### 3.7 Stranding

An argument in favour of the monoclausal analysis advanced by Paperno 2012 is based on the fact that there is a Russian pronoun, namely *čto*, which requires adjectival modifiers to appear in a non-agreeing genitive case form, unlike other nominals which take adjectival modifiers fully agreeing in case. Paperno 2012 offers a test based on the phenomenon of stranding, a kind of “partial *wh*-movement”, and illustrates it using the following example from Russian:

- (50) Čto i komu on xorošego sdelal?  
 what<sub>ACC</sub> and who<sub>DAT</sub> he good<sub>GEN</sub> did  
 ‘What good did he do, and to whom?’ (Paperno 2012: ex. (49))

Paperno 2012 argues that the structure of this example must be non-elliptical (monoclausal) due to the fact that the adjectival modifier *xorošego* must depend on – and hence belong to the same clause as – the first conjunct, *Čto*. He provides the following example to demonstrate that the modifier cannot occur in the non-agreeing genitive form on its own:

- (51) \*Komu on xorošego sdelal?  
 who<sub>DAT</sub> he good<sub>GEN</sub> did  
 ‘To whom did he do good?’ (Paperno 2012: ex. (50))

Furthermore, Paperno 2012 shows that this argument is immune to those multiclausal analyses which claim that there is ellipsis in one of the conjuncts coupled with the use of an indefinite pronoun (as in T<sub>1</sub>, for instance):

- (52) Komu on čto-libo xorošee sdelal?  
 who<sub>DAT</sub> he something good<sub>ACC</sub> did  
 ‘To whom did he do something good?’ (Paperno 2012: ex. (51))

This example features an indefinite pronoun, *čto-libo*, which triggers full modifier agreement. The multiclausal analysis of Tomaszewicz 2011a features an indefinite pronoun represented as *something*, see (5) and (7) which correspond to (4). However, assuming that the indefinite pronoun of Tomaszewicz 2011a behaves in the same way as *čto-libo*, (50) could not be an instance of ellipsis of an indefinite pronoun – if this was the case, the modifier would be expected to appear in the agreeing form, as in (52).

The facts in Polish are similar: while Polish adjectival modifiers usually fully agree in case with their nominal heads, there are pronominal forms ((non-)agreement depends on the case of the head) which require the modifier to appear in a non-agreeing case, namely genitive. The following example is analogous to the one provided in (50):

- (53) Co i komu ona ciekawego/\*ciekawe powiedziała?  
 what<sub>ACC</sub> and who<sub>DAT</sub> she interesting<sub>GEN/ACC</sub> said  
 ‘What interesting did she say, and to whom?’

There are, however, certain differences with respect to facts from Russian as Polish has more pronouns which display the (non-)agreement pattern shown above for the interrogative *co*. This class also includes elements such as *coś* (indefinite), *cokolwiek* (any-type pronoun) and *nic* (n-word). Since, unlike in Russian, such (non-)agreement is possible in Polish with

the indefinite pronoun (*coś*, see (54)), this test is vulnerable to claims that one of the conjuncts features an implicit indefinite pronoun (as in T<sub>1</sub>).

- (54) \*(Coś) ciekawego się stało.  
 something<sub>NOM</sub> interesting<sub>GEN</sub> REFL happened  
 ‘Something interesting happened.’

While this test does not provide a definite argument against multiclausal analyses (it seems that at least T<sub>2</sub> and CGY<sub>2</sub> would be technically able to account for such data), there, again, seems to be no motivation to use such accounts instead of the monoclausal analysis.

### 3.8 Governing Numerals

Examples such as the following might provide new evidence supporting monoclausal analyses of lexico-semantic coordination:<sup>3</sup>

- (55) Kto, ile i kiedy dostał unijnych dotacji?  
 who how much<sub>ACC</sub> and when got EU subsidies<sub>GEN</sub>  
 ‘Who got how much EU subsidies and when?’ (NKJP)
- (56) Nie wiem w ogóle, ile i kiedy dostanę pieniędzy na  
 NEG know at all how much<sub>ACC</sub> and when get<sub>FUT</sub> money<sub>GEN</sub> for  
 naszą działalność.  
 our operation  
 ‘I have no idea how much money I will get for our operation and when.’ (NKJP)

Both examples provided above contain governing numeral forms – the distinctive feature of such forms is that they assign genitive case to the accompanying nominal: the head numeral *ile* is marked for accusative case (structural case assigned by the verb) while its nominal object bears genitive case: *dotacji* in (55) and *pieniędzy* in (56).

This feature of governing numerals makes it difficult to argue that ellipsis is at work in such examples because their hypothetical multiclausal base sentences would lack identity across clauses, as shown in (57), a multiclausal paraphrase of (56):

<sup>3</sup> This argument is, however, similar to modifier stranding presented in § 3.7.

- (57) Nie wiem w ogóle, ile pieniędzy dostanę i kiedy  
 NEG know at all how much<sub>ACC</sub> money<sub>GEN</sub> get<sub>FUT</sub> and when  
 dostanę pieniądze.  
 get<sub>FUT</sub> money<sub>ACC</sub>  
 ‘I have no idea how much money I will get and when I will get the  
 money.’

(57) shows that ellipsis analyses which postulate deletion under identity in the first clause (such as  $T_1$ ) are impossible in such cases due to the fact that the case found in the second clause (accusative *pieniądze* required by the verb as structural case in this context) does not match the case found in the first clause (genitive *pieniędzy* required by the numeral head *ile*). If the example using lexico-semantic coordination, (56), were multiclausal, the genitive *pieniędzy* would be unexpected as the numeral (*ile*) would be placed in the first clause, while the verb in the second clause requires an object marked for accusative case (*pieniądze*, as in (57)).

Finally, though theoretically the verb *DOSTAĆ* ‘get’ can assign genitive case (as a realisation of structural case) to its object under the partitive reading, it does not seem to be an option in (55)–(56). Such interference can be eliminated by using predicates where such a reading is unavailable, as in the example below with *ROZWIĄZAĆ* ‘solve’:

- (58) Ile i kto rozwiązał zadań?  
 how many<sub>ACC</sub> and wh<sub>NOM</sub> solved tasks<sub>GEN</sub>  
 ‘How many tasks did who solve?’

While other multiclausal accounts such as  $T_2$  and  $CGY_2$  could probably handle such data technically, there seems to be no reason which would justify adopting these accounts instead of the monoclausal analysis.

### 3.9 Coordination with *Yes/No* Question Particle

Polish *yes/no* question particle *CZY* can be coordinated with *wh*-words:

- (59) Tytuł brzmiał prosto i uczciwie: "Czy i jaki jest Bóg"  
 title sounded simply and honestly PART and what is God  
 ‘The title sounded simple and honest: “Does God exist and what is  
 he like?”’ (NKJP)

- (60) Nie wiemy wreszcie, czy i co kto chowa w  
 NEG know besides PART and what<sub>ACC</sub> who<sub>NOM</sub> hides in  
 rękawie.  
 sleeve  
 ‘Besides, we don’t know if they got something up their sleeves and  
 who keeps what up their sleeve.’ (NKJP)

While such examples are common and their grammaticality is rather uncontroversial, it is worth noting that removing the conjunction results in ungrammaticality:<sup>4</sup>

- (61) \*Czy co kto chowa w rękawie?  
 PART what<sub>ACC</sub> who<sub>NOM</sub> hides in sleeve

This suggests that the question particle *CZY* cannot be used with *wh*-words as dependents of the same predicate.<sup>5</sup> However, sentences where *wh*-words depend on a different predicate are grammatical, as shown below:

- (62) Czy wiesz, co jesz?  
 PART know what<sub>ACC</sub> eat  
 ‘Do you know what you are eating?’ (NKJP)
- (63) Kto wie, czy Abraham nie był czarny?  
 who<sub>NOM</sub> knows PART Abraham NEG was black  
 ‘Who knows whether Abraham was not black?’ (NKJP)

In these examples the question particle *CZY* and *wh*-words belong to distinct clauses. In (62) *Czy* is placed in the main clause where *WIEDZIEĆ* ‘know’ is the main verb, while *co* is the object of *JESĆ* ‘eat’ in the subordinate clause. By contrast, in (63) *Kto* is the subject of the main verb (*WIEDZIEĆ*), while the question particle *czy* belongs to the subordinate clause (featuring *BYĆ* ‘be’). As a result, these examples satisfy the requirement that there be no *wh*-words in the clause which contains *CZY*.

If this constraint is accepted, it follows that the structure of lexico-semantic coordination featuring *CZY* as one of the conjuncts cannot be

<sup>4</sup> (61) can be judged as grammatical under the reading where *co* and *kto* are interpreted as indefinite pronouns (existential). This, however, does not affect the presented argument, since it is concerned with the interpretation where these are *wh*-words.

<sup>5</sup> This observation was also made by Tomaszewicz 2011a: “In Polish the clause-initial marker *czy* cannot co-occur with *wh*-phrases, yet it is allowed in Coordinated-WHs, which provides evidence for the clausal character of the conjuncts.”

monoclausal. However, this constraint is satisfied under multiclausal analyses where *CZY* and *wh*-words never belong to the same clause (these include  $T_1$  and  $CGY_1$  discussed in § 2).

Lexico-semantic coordination with *CZY* is a special case due to the fact that removing the conjunction in other lexico-semantic environments does not lead to ungrammaticality – the result of such an operation with coordinated *wh*-words is a monoclausal structure, a multiple question. However, this is not possible with *CZY*, which provides the only constructive argument in favour of adopting a multiclausal analysis – it is at the same time the only environment where the monoclausal analysis is not appropriate.

#### 4 Conclusion

This paper provided a critical review of 9 selected arguments arguing for monoclausal or multiclausal representation of lexico-semantic coordination applied to Polish. It showed that while there is evidence suggesting that structures with *CZY* should be analysed as multiclausal in Polish (see § 3.9), there is no evidence supporting such an analysis when *CZY* is not involved in such coordination. On the other hand, while it was demonstrated that some multiclausal analyses could not account for some phenomena (see § 3.7, § 3.8 and § 3.9), there is no evidence which would make it possible to reject the remaining multiclausal accounts.

Some multiclausal analyses use ellipsis (such as  $T_1$  and  $T_2$ ), but it is possible to argue against them since they postulate ellipsis under identity. However, if the identity requirement is abandoned, ellipsis becomes an extremely very powerful operation, which is starkly visible when considering phenomena such as gapping – for instance, the head of a clause may be removed and there seems to be no requirement of strict identity of verb forms (singular *lubi* vs plural *lubią*); besides, the dependent of the gapped clause may bear different case than in the full clause (accusative *Marysię* vs genitive *Marysi*, triggered by negation):

- (64) Janek *lubi* Marysię, a jego rodzice *nie* (*lubią* Marysi).  
 Janek like<sub>3,SG</sub> Marysia<sub>ACC</sub> and his parents NEG like<sub>3,PL</sub> Mary<sub>GEN</sub>  
 ‘Janek likes Marysia, but his parents don’t (like Marysia).’

The example featuring gapping serves to show that multiclausal analyses assuming ellipsis can be saved by stipulating the use of extra devices (such as the use of implicit pronouns) to account for relevant data.

While both analyses, monoclausal and multiclausal, are available in theory, it seems preferable to choose the more economic and simple analysis if there is no reason to do otherwise. As a consequence, the monoclausal analysis emerges as the default analysis – it does not require the use of implicit pronouns and coindexing, it does not use ellipsis mechanisms which are hard to justify in other syntactic contexts, it does not require multidominance. The multiclausal analysis seems to be motivated only for cases when one of the conjuncts is the *yes/no* question particle CZY.

Such a split analysis of Polish lexico-semantic coordination (formalised in Lexical-Functional Grammar (LFG, Bresnan 1982, Dalrymple 2001)) is presented in Patejuk and Przepiórkowski 2012: the multiclausal analysis is only used for coordination with CZY as one of the conjuncts, while the monoclausal analysis is used elsewhere.

## References

- Bresnan, Joan (ed.). 1982. *The Mental Representation of Grammatical Relations*. MIT Press Series on Cognitive Theory and Mental Representation, Cambridge, MA: The MIT Press.
- Bîlbîie, Gabriela and Gazdik, Anna. 2012. Wh-coordination in Hungarian and Romanian multiple questions. *Empirical Issues in Syntax and Semantics* 9 pages 19–36.
- Ćitko, Barbara and Gračanin-Yüksek, Martina. 2013. Towards a new typology of coordinated *wh*-questions. *Journal of Linguistics* 49, 1–32.
- Dalrymple, Mary. 2001. *Lexical-Functional Grammar*. Academic Press.
- Kallas, Krystyna. 1993. *Składnia współczesnych polskich konstrukcji współrzędnych*. Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika.
- Kazenin, Konstantin I. 2001. On coordination of Wh-phrases in Russian, MS, University of Tübingen and Moscow State University.
- Paperno, Denis. 2012. *Semantics and Syntax of Non-Standard Coordination*. Ph.D. dissertation, University of California.
- Patejuk, Agnieszka and Przepiórkowski, Adam. 2012. Lexico-semantic coordination in Polish. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'12 Conference*, pages 461–478, Stanford, CA: CSLI Publications.

- Patejuk, Agnieszka and Przepiórkowski, Adam. 2014. Lexico-semantic coordination in Polish: a critical review of tests for determining representation. In Małgorzata Gębka-Wolak, Joanna Kamper-Warejko and Andrzej Moroz (eds.), *Leksyka języków słowiańskich w badaniach synchronicznych i diachronicznych*, pages 119–134, Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, in print.
- Przepiórkowski, Adam, Bańko, Mirosław, Górski, Rafał L. and Lewandowska-Tomaszczyk, Barbara (eds.). 2012. *Narodowy Korpus Języka Polskiego [Eng.: National Corpus of Polish]*. Warsaw: Wydawnictwo Naukowe PWN.
- Przepiórkowski, Adam, Bańko, Mirosław, Górski, Rafał L., Lewandowska-Tomaszczyk, Barbara, Łaziński, Marek and Pęzik, Piotr. 2011. National Corpus of Polish. In Zygmunt Vetulani (ed.), *Proceedings of the 5th Language & Technology Conference: Human Language Technologies as a Challenge for Computer Science and Linguistics*, pages 259–263, Poznań, Poland.
- Rațiu, Dafina. 2009. Coordinatated Questions vs. Matching Questions in Romanian. Abstract of a talk presented at ConSOLE XVII, Nova Gorica, 16–18 January 2009; <http://sabotin.ung.si/~jezik/console/abstracts/012.pdf>.
- Tomaszewicz, Barbara. 2011a. Against Spurious Coordination in Multiple Wh-questions. In *Proceedings of the 28th West Coast Conference on Formal Linguistics*, pages 186–195, Somerville, MA: Cascadilla Press.
- Tomaszewicz, Barbara. 2011b. Wh & Wh: Syntactic and Semantic Arguments for Clausal Coordination. In *Formal Approaches to Slavic Linguistics 19*, pages 131–145, University of Michigan, Ann Arbor: Michigan Slavic Publications.

## **The Absolute Basis of Middles and the Status of vP and UTAH\***

*Ljiljana Progovac*  
Wayne State University

UTAH (Uniformity of Theta Assignment Hypothesis) has been a useful hypothesis, but it has been proven wrong in various ways. The theoretical question raised in this paper is whether or not there is an alternative universal organizing principle that can replace UTAH and that can help guide syntactic derivations. Here I explore such a principle, which I call FASA (First Argument, Second Argument), and propose that it is general enough to be equally applicable to nominative-accusative (nom-acc) and ergative-absolutive (erg-abs) patterns, as well as equally applicable to verb phrases and noun phrases. This proposal takes inspiration in the description and analysis of an absolutive language, Tongan, by e.g. Tchekhoff (1973, 1979). The empirical focus of the paper is on Serbian “middle” *se* constructions, showing that they have an absolutive basis, and that they can best be accommodated by FASA.

### **1 What is Wrong with UTAH?**

UTAH (Uniformity of Theta Assignment Hypothesis) is an influential hypothesis widely adopted/assumed in formal syntax, claiming that specific theta roles are associated with specific syntactic positions (see e.g. Baker 1988). For example, a theme is considered to be Merged in the

---

\* Many thanks to Robert Henderson for very useful discussions of ergativity, as well as of the semantics of theta roles. I am also grateful to Johanna Nichols for her comments and suggestions, as well as to an anonymous reviewer for good observations.

VP (First Merge), while an agent is considered to be Merged in the specifier of the vP shell (Second Merge in transitive structures). UTAH has played a useful role in sorting through the data of various languages, but now it seems clear that this hypothesis cannot be maintained, and that an alternative is needed.

### 1.1 *Non-Agents in vP*

The first type of problem faced by UTAH is on the vP end, and it concerns non-agent subjects being Merged in vP. Even when syntacticians say they adopt UTAH, they often generate theta roles other than agents in the light vP, or just remain silent with respect to where these non-agent subjects in transitive constructions should be generated (e.g. Chomsky 1995; Adger 2003):

- |     |                                       |               |
|-----|---------------------------------------|---------------|
| (1) | John saw/understood the problem.      | (experiencer) |
| (2) | John kicked the ball.                 | (agent)       |
| (3) | The hammer/the wind broke the window. | (instrument?) |
| (4) | The escalation frightened John.       | (theme?)      |

It would not be wise to propose that the subjects in each of (1-4) are generated in distinct syntactic positions. Syntax provides fewer slots for generating arguments than we can postulate semantic/thematic roles. It seems firmly established that (mono-)transitive structures in English have two structural positions for Merging the internal and external arguments, the first inside the VP and the second inside the vP.<sup>1</sup> On the other hand, not only are theta roles other than agents common as transitive subjects, but the differentiation among these roles is often not clear-cut.

To take just one example, one wonders if the role of *the wind* in (3) is that of an agent, an instrument, or something else? Linguists can argue and disagree over this, but regardless of what they conclude, *the hammer* and *the wind* in (3), as well as *John* in (2), are generated in the same syntactic position, such as the specifier of vP. My proposal here will build

---

<sup>1</sup> In this paper, I am not considering di-transitive verbs, that is, those verbs which seem to take more than two arguments, as illustrated in (i). They require some special explanation on any account, and I leave them aside for now.

(i) John gave Mary a book.

on this observation that syntax can provide two structural positions for Merging arguments, but that these positions are not associated with specific theta roles.

### 1.2 *Agents in VP? Absolutives and Nominals*

In addition to the problems on the vP end, there are problems for UTAH on the VP end as well, and this has to do with agents Merging inside VPs. This will be illustrated in this section with absolutives in an ergative language, Tongan, and with nominals in English and Serbian, while Section 3 focuses on *se* middles in Serbian in this respect.

When it comes to absolutives in intransitive constructions, especially in languages which show both morphological and syntactic ergativity (e.g. Tongan and Dyirbal), there is no structural differentiation between the agent and the theme in the absolutive roles, and it would be wrong to analyze absolutive agents as syntactically distinct from absolutive themes. In other words, it would be erroneous to analyze the two interpretations in (5) from Tongan as involving two distinct syntactic structures.

- (5) Oku ui 'a Mele (Tchekhoff 1973: 283)  
PRES call ABS Mary  
 'Mary calls.' / 'Mary is called.'

As Tchekhoff (1973: 283) characterizes the pattern in (5), *Mary* is neither an agent nor a theme, and the two translations in (5) just reflect a nom-acc bias. In his own words, in (5) "Mary is the only determiner [i.e. argument, L.P.], and the whole utterance gives us only the following information: present tense, verb *call*, *Mary*. And we don't know whether Mary was the agent of the calling, or the recipient of it. Nothing in the Tongan original informs us as to this particular point. Therefore, we cannot qualify an '*a* determiner as anything more than 'first determiner.' It is first because it can appear alone, without the '*e* determiner [ergative, as in (7) below, L.P.]. This latter's function is an explicit agent-function; it remains constant in all its uses."

If so, then (5) should be analyzed semantically as in (6), where *Mary* is just an underspecified participant in the event of calling (see also Dowty 1991 for the idea of proto-roles).

- (6)  $\lambda x \exists e [C(e) \wedge \text{Participant}(x,e)]$   
 $\exists e [C(e) \wedge \text{Participant}(Mary,e)]$

What happens in (7) is that the addition of the (ergative-marked) agent renders the internal argument *Mary* as non-agent, due to the exhaustivity of theta-roles, that is, due to a prohibition against having two agents. But, basically, the ergative phrase can be analyzed as added to the absolutive (VP) layer. If so, then the semantics for (7) should be as given in (8).

- (7) Oku ui 'e Sione 'a Mele  
PRES call ERG John ABS Mary  
 'John calls Mary.'
- (8)  $\lambda y \lambda x \exists e [C(e) \wedge \text{Agent}(x,e) \wedge \text{Participant}(y,e)]$   
 $\exists e [C(e) \wedge \text{Agent}(John,e) \wedge \text{Participant}(Mary,e)]$

In (7) the ergative is interpreted as an agent, given that *call* takes an agent and a theme. However, in other cases the ergative is interpreted as e.g. an experiencer, as in the example (9) below from Dyrirbal, given that the verb *see* takes an experiencer and a theme. In this case the absolutive argument is interpreted as non-experiencer, and thus as a theme. What the experiencer in (9) and the agent in (7) have in common is that they are both second arguments. What absolutives in (5), (7), and (9) have in common is that they are first arguments. These patterns thus reflect a grammar based on first vs. second argument, rather than a grammar based on theta roles.

Given this analysis, one can observe clear continuity between (5) and (7), as the role of Participant is there in both. In other words, the First Merge starts with an underspecified role, both in intransitive and transitive structures, and then, optionally, Second Merge adds a more specified role, a higher role, such as agent or experiencer. This is in line with some more recent proposals as well, such as Otsuka (2011), who has proposed that intransitive subjects in ergative languages (absolutives), whether unaccusative or unergative, are generated VP-internally, in violation of UTAH. Likewise, Wiltschko (2006) has argued that in Halkomelem Salish, an ergative language, intransitive subjects are generated VP-internally, whereas transitive subjects are generated outside VP.

Dyirbal (Australian language spoken in northeast Queensland) is another ergative language which, like Tongan, exhibits syntactic ergativity, in the sense that the absolutive role, even when the ergative-marked argument is present, continues to behave in a subject-like fashion, coordinating with intransitive absolutives (see e.g. Dukes 1998; Aldridge 2008), as illustrated below (Dixon 1994: 155):

- (9) nguma yabu-nggu buran banagan<sup>y</sup>u  
 father<sub>ABS</sub> mother<sub>ERG</sub> saw return  
 ‘Mother saw father and (father) returned.’

This clearly contrasts with English (10) below, in which a comparable coordinated structure yields the opposite result for the missing argument. However, if we were to coordinate a passive sentence and an active sentence in English (11), we would get the pattern comparable to the one in (9) from Dyirbal:

- (10) Mother saw father and (mother) returned home.  
 (11) Father was seen by mother, and (father) returned home.

In (9) and (11), there is only one *structural* argument (call it First Argument), as the optional ergative or *by*-phrase can be seen as an adjunct, not affecting coordination possibilities. This is the sense in which the ergative phrase can be likened to the passive *by*-phrase (see e.g. Nash 1996 and Alexidou 2001). The *by*-phrase in (11), just like the ergative phrase in (9), is not the true, structural subject, but only the “logical” subject. According to Nash (1995: 119), nom-acc languages have external arguments in vP, while ergative languages may treat their (ergative) agents as adjuncts. Similarly, ergative is treated as an inherent case by e.g. Woolford (1997, 2006); Legate (2008); and Massam (2000, 2001).<sup>2</sup>

It is also of interest that nominals (in nom-acc languages) have been analyzed as absolutive-like, as discussed in Alexidou (2001) (see also Picallo 1991; Bottari 1992; Alexiadou and Stavrou 1998). In the

---

<sup>2</sup> In contrast, Otsuka (2011) treats ergative as structural, rather than inherent case, based on Levin and Massam (1985); Bobaljik (1993); and Laka (1993). It may well be that ergative languages can vary in this respect.

examples below from English and Serbian, the only argument can be analyzed as not syntactically specified as theme or agent (12, 14, 15). Also, the addition of the optional *by*-phrase (13, 14, 16) has the same disambiguating effect as does the addition of the ergative argument in the erg-abs sentences in Tongan, as discussed above (see also Comrie 1978).<sup>3</sup>

- (12) John's examination was unnecessary.  
 (13) John's examination by the psychiatrist was unnecessary.  
 (14) John's portrayal (by the media/of the media) was unfair.  
 (15) kritikovanje studenata  
       criticizing students<sub>GEN</sub>  
       'criticizing of/by the students'  
 (16) kritikovanje studenata od strane profesora  
       criticizing students<sub>GEN</sub> by professors<sub>GEN</sub>

Both the Tongan examples and the nominals in English and Serbian above can thus be analyzed as essentially intransitive, with only one structural argument (First Argument) exhibiting an underspecified thematic role. The external argument (Second Argument) is not structural in these cases, but can be expressed optionally, as an ergative or *by*-phrase. In other words, my proposal here is that the First Argument is thematically underspecified, and that it is only the addition of the external argument (Second Argument), whether as an adjunct or as a vP specifier, that disambiguates the role of the First Argument.<sup>4</sup> The next

---

<sup>3</sup> There is typically a preference for the only argument to be interpreted as a theme, if possible, but what these data show is that it is not necessary to do so. The preference may be stronger with some sentences, as in the following example offered by the reviewer (i). However, (ii) seems to freely allow the agent interpretation of the genitive noun phrase, suggesting that the interpretations can be manipulated by pragmatic context. For my argument to go through, it is enough that the agent interpretation of the genitive argument is in principle possible.

- (i) Uništavanje insekata bilo je surovo.  
       destroying insects<sub>GEN</sub> been was cruel  
 (ii) Uništavanje/razaranje njihove vojske bilo je surovo.  
       destroying their army<sub>GEN</sub> been was cruel

<sup>4</sup> For the lack of neutral terminology, here I am referring to the agent/experiencer introduced in a *by*-phrase or in an ergative phrase as an argument of the verb, even though syntactically speaking these may be attached as (optional) adjuncts.

section establishes that the absolutive-like pattern also underlies “middle” *se* constructions in Serbian. As will be argued, they also do not observe UTAH, and are analyzed in this paper as Merging their only argument as an underspecified (absolutive-like) thematic role (First Merge).

## 2 The Absolutive Basis of *se* Middles

The goal of this section is to provide a unified syntax and semantics for *se* middles in Serbian, which is only possible by embracing underspecification, and by giving up UTAH. This analysis establishes a common ground between absolutives in e.g. Tongan and *se* middles in e.g. Serbian.<sup>5</sup> My proposal in Section 2.1 is to generate the only argument (First Argument) of *se* middles in VP, even when interpreted as an agent (as per FASA of Section 3, and contra UTAH). Importantly, the reflexive interpretation will follow for free from this analysis, and it will follow without a need to treat *se* as a reflexive pronoun. The problems for an alternative UTAH-based analysis are discussed in Section 2.2.

### 2.1 *Se Middles in Serbian: Some Surprising Data and Generalizations*

Consider now some well-known examples with *se* in Serbian:

- (17) Ljudi se briju.  
 people SE shave
- (18) Jabuke se jedu.  
 apples SE eat

One is tempted to analyze these as distinct constructions, so that (17) involves a reflexive pronoun (*se*), and *people* Merging in vP as an external argument, and so that (18) involves a passive-like middle, with *apples* Merging as an internal argument in VP, and *se* Merging in vP. This would be exactly the wrong way of approaching these data. This kind of differentiation comes to mind only because the pragmatics foregrounds these two respective readings. When one considers the data

---

<sup>5</sup> Interestingly, Otsuka 2011 has postulated a null *se* pronoun in Tongan, even though Tongan does not have any (overt) reflexive pronouns. I interpret this as his observation that there is structural similarity between Tongan absolutives and *se* middles in this respect, although I would certainly not posit any null pronouns in Tongan.

where pragmatics does not skew the readings so much, a different analysis suggests itself, involving underspecification/vagueness:

- (19) *Deca se tuku.*  
 children <sub>SE</sub> hit  
 ‘The children are hitting each other.’  
 ?‘The children are hitting themselves.’  
 ‘The children are hitting somebody (else.)’  
 ‘One spans children.’

Here, almost any interpretation involving *children* as a participant in the event of *hitting* is possible, and this calls for the underspecified semantics, along the lines in (20).

- (20)  $\lambda x \exists e [H(e) \wedge \text{Participant}(x,e)]$   
 $\exists e [H(e) \wedge \text{Participant}(\text{Children},e)]$

In support of the analysis in (20), various readings can overlap (be available at the same time), which is the hallmark of vagueness/underspecification, as opposed to structural ambiguity (e.g. Kempson 1977). The example in (19) can easily be interpreted as the children hitting themselves, each other, and others all at the same time. This kind of overlap does not seem possible with the clearly transitive structures below:

- (21) *Deca tuku (sama) sebe.*  
 ‘The children are hitting themselves.’  
 (22) *Deca tuku jedni druge.*  
 ‘The children are hitting each other.’  
 (23) *Deca tuku nekoga.*  
 ‘The children are hitting somebody.’

As was the case with the absolute arguments in the previous section, *deca* (First Argument) in (19) should be generated by First Merge in the VP regardless of the interpretation, rather than associating the sentence with a multitude of distinct syntactic structures.

From this underspecified (absolute-like) semantics follow all the readings above (for additional readings, see below), including the

reflexive reading, in which the expressed NP can be both an external argument and an internal argument *at the same time*, again, the hallmark of vagueness. Unlike *sebe* in (21), *se* is not a reflexive pronoun. I would like to submit that these *se* middles, due to their interpretation along the lines in (20), are particularly suited for expressing reflexive readings without using a reflexive pronoun, which is why such structures are often referred to as “reflexives,” and why *se* is often confused with a reflexive.

The following examples show comparable underspecification, with even some additional surprising readings:

- (24) Deca se ljube.  
 children SE kiss  
 ‘The children are kissing each other.’  
 ?‘The children are kissing themselves.’  
 ‘The children are kissing someone (else).’  
 ‘One kisses children.’
- (25)  $\exists e [K(e) \wedge \text{Participant}(\text{children}, e)]$
- (26) Pas se ujeda.  
 dog SE bites  
 ‘The dog bites (someone).’  
 ‘?The dog is biting itself.’  
 ‘??One bites dogs.’
- (27)  $\exists e [B(e) \wedge \text{Participant}(\text{dog}, e)]$
- (28) Marko se udara!  
 Marko SE hits  
 ‘Marko is hitting me/us.’  
 ‘Marko is hitting somebody.’  
 ‘Marko is hitting himself.’
- (29)  $\exists e [H(e) \wedge \text{Participant}(\text{Marko}, e)]$

Remarkably, if uttered with a sense of urgency, the most probable interpretation in (28) involves the most salient discourse participant, the speaker, even though there is no morpheme or word corresponding to 1sg at all! The semantics proposed in (29) and elsewhere can easily accommodate the “me” readings, but such readings pose quite a challenge for alternative accounts, including the one discussed in Section 3.2.

This “me” interpretation is hardly just a quirk of Serbian. As reported in e.g. Kański (1986: 195), a similar fact is observed in Polish imperatives with the “reflexive” clitic *się* (30), as replicated in Serbian (31).

- (30) Nie pchaj się pan!  
 not push<sub>SE</sub> man  
 ‘Stop pushing me, man!’
- (31) Ne guraj se!  
 not push<sub>SE</sub>  
 ‘Don’t push me/us.’  
 ‘Don’t push people./Don’t push around.’

As mentioned in Rivero and Milojević-Sheppard (2003), many have noted that reflexive clitics may denote a speaker or a hearer, but that the phenomenon is not well understood. These facts follow straightforwardly from the proposed semantics in e.g. (29), which only specifies that *Marko* or *pan* are involved in the events of hitting or pushing, respectively, and this can certainly include the interpretation on which the speaker is the one hit or pushed.

## 2.2. *An Alternative UTAH-based Analysis of se Middles*

The solution that always comes to mind first is the following: why not treat *se* as an argument and generate it in VP vs. vP, depending on interpretation, following UTAH? Why not just impose order on argument structure? Rivero and Milojević-Sheppard (2003) offer such an alternative analysis of *se* middles, although for Slovenian and Polish, in which *se* is treated as an argument. According to Rivero and Milojević-Sheppard (2003: 123) “the analysis in which indefinites are defective pronouns triggering movement in the syntax relates to the semantic proposal where they contain an existential quantifier and a human variable... Thus we unify the idea that defective pronouns/*se*-anaphors must trigger movement in the syntax and the view in dynamic semantics that indefinites are inherently quantificational.”

Rivero and Milojević-Sheppard do note that sometimes the interpretation of *se* is not existential/indefinite, and this includes, but is not limited to, the speaker/hearer interpretations given in (30). Their suggestion is that the unexpected “me” readings in e.g. (30) can be

derived from their indefinite pronoun analysis of *sig*, by allowing the reading along the lines of: “Do not push someone (who happens to be me).” One problem with this suggestion is that *someone* and its equivalents in other languages do not typically receive “me” interpretations, and certainly not as naturally as *se* middles do, especially those in (30-31).

But more generally, this kind of analysis does not extend to Serbian. As this paper argues, there is overwhelming evidence against it. First of all, such an analysis, which treats *se* as an argument generated either in VP or vP, would introduce multiple syntactic and/or lexical ambiguities in the absence of any morphosyntactic evidence for distinct structures. The analysis in this paper does not posit ambiguities (different structures and/or different lexical items for different interpretations), but instead posits one single structure for each example discussed above, which happens to be underspecified, and thus vague, but not ambiguous.

Second, the deep descriptive generalization would be missed, and the unified account of *se* lost. Third, we would have no explanation for why *se*, when generated VP internally as per UTAH, sometimes means *self*, sometimes *each other*, sometimes *someone else*, and other times *me* or *us*. One would not want to warrant several dictionary entries for *se* in Serbian, including those with an interpretation *me* or *us*, especially not if a unified analysis is available. Fourth, a syntactic approach which generates *se* in distinct syntactic positions cannot explain overlaps in interpretation involving multiple readings at the same time, such as those discussed in the previous section.

Last but not least, the following coordination tests show that the arguments in Serbian *se* middles can coordinate regardless of their theta roles. The following examples, in which *se* is shared, show that reflexive and null object readings are not distinct, as they can coordinate. If these readings were to involve two distinct lexical entries for *se* (*somebody* vs. *self* in (32)), then this coordination option would not be available.<sup>6</sup> It is important to point out here that *znojiti se* is a verb that can only be used with *se*, so *se* must be shared between the two verbs in (32).

---

<sup>6</sup> This example also, naturally, allows the reading on which both actions are reflexive, that is, the reading on which Jovan is kissing himself, and sweating (himself). This is the only reading that this alternative analysis should allow, and yet this reading is not even preferred, as it is pragmatically odd.

- (32) Jovan se ljubi i brije/znoji.  
 Jovan<sub>SE</sub> kisses and shaves/sweats  
 ‘Jovan is kissing (somebody) and shaving/sweating (himself).’

The following examples show that reflexive and reciprocal readings are not distinct either. Again, if *se* were to involve two separate lexical entries in the two conjuncts in (33) (*each other* vs. *self*), then the coordination below should not be possible, contrary to fact:

- (33) Oni se dodaju loptom i znoje.  
 they<sub>SE</sub> pass ball<sub>INST</sub> and sweat  
 ‘They are passing the ball to each other and sweating (themselves).’

Even more dramatically, passive and reflexive readings can also coordinate. Again, *radovati se* is a verb that must occur either with an object or with *se*, which means that *se* in (34) must be shared. Likewise, *smejati se* must occur with *se*, which again indicates that *se* in (35) must be shared between the two verbs.

- (34) ?U susednoj sobi, bebe se doje i raduju.  
 in next room babies<sub>SE</sub> breastfeed and rejoice  
 ‘Next door, babies are being breastfed and are rejoicing (themselves).’
- (35) ?U susednoj sobi, bebe se prskaju vodom i smeju.  
 in next room babies<sub>SE</sub> sprinkle water<sub>INST</sub> and laugh  
 ‘Next door, babies are being sprinkled with water and are laughing.’

If *babies* in the first (passive-like) conjunct of (34) were generated in VP as an object and *se* as an agent argument in vP, then this structure should not be able to coordinate with the second conjunct, in which *se* would be a reflexive pronoun on this analysis, presumably generated in VP, and *babies* would be an agent. Similar considerations hold of (35).<sup>7</sup>

<sup>7</sup> The reviewer agrees with the judgments in the text, but brings up the example in (i) to suggest that these kinds of mixed-and-matched interpretations are not always available:

- (i) Deca se ljube i udaraju.  
 children<sub>se</sub> kiss and hit

On the other hand, these coordinations pose no difficulties for the analysis proposed in this paper: if *se* middles are analyzed as involving one single First Argument, generated in VP with an underspecified theta role, then these coordinations should be grammatical, subject only to pragmatic possibilities and probabilities. In this view, *se* is not an argument, but rather just a grammatical marker, as discussed further in Section 3.

### 3 The Proposal: First Argument, Second Argument (FASA)

Given the discussion in the previous sections, my proposal is to replace UTAH with the following principle, which is inspired by the characterization of the absolutive and ergative arguments in Tongan by Tchekhoff (1973, 1979):<sup>8</sup>

(36) First Argument, Second Argument (FASA)

A) If there is only one argument (First Argument), it will be generated in the VP, and assigned a(n underspecified)

Participant role:

$\lambda x \exists e [E(e) \wedge \text{Participant}(x,e)]$

B) Only the presence of a higher Second Argument (e.g. agent/experiencer) will render the First Argument (Participant) as a non-agent/non-experiencer, as well as lead to the projection of a vP in nom-acc patterns:

$\lambda y \lambda x \exists e [E(e) \wedge \text{Agent/Experiencer}(x,e) \wedge \text{Participant}(y,e)]$

---

For the reviewer, only those interpretations are available in (i) in which “deca” is either the theme argument, or the agent argument of both verbs. Again, there is no doubt that many examples will favor, or strongly favor, one vs. another interpretation, as is also the case with the initial examples in (17) and (18).

In order to get the other interpretations, including the mixed-and-matched interpretations, one needs to construct specific contexts, such as using “babies” in (34) and (35) in the text. Such contexts, by excluding certain interpretations pragmatically, make the other (less likely) readings shine. This strategy is similar to the way inverse scope readings, typically hard to get, are made to shine by excluding the more likely interpretations pragmatically.

<sup>8</sup> This is not incompatible with Borer’s (1994) fully configurational approach to argument linking, according to which there is no lexical distinction between subjects and objects in the VP; such distinctions can only be made with the help of functional projections such as vP.

Tchekhoff (283) refers to the First Argument as First Determiner, as per the quote in Section 2.2. This suggests that the number of arguments, not just their nature, determines where they are Merged. Some appeal to a hierarchy of theta-roles will be needed (e.g. Fillmore 1968), which would prevent themes from being generated higher than agents/experiencers, but such a stipulation is also an integral part of UTAH.

Needless to say, this proposal has far-reaching consequences for the structure of the vP/VP shell. One important advantage of this proposal is that it can provide a common foundation for building structure (i) in both nominal and verbal domains, (ii) in both nom-acc and erg-abs patterns, and (iii) both across erg-abs and nom-acc languages. FASA works particularly well for absolutive phrases in syntactically ergative languages (e.g. Tongan), as well as for *se* middles in e.g. Serbian, as per the proposed (absolutive-like) analysis. It also carries over to unaccusatives (e.g. *fall*, *collapse*), as they also have only one argument generated in VP; for example Chomsky (1995) and Kratzer (1996, 2000) analyze unaccusatives as projecting only VP, but not a vP shell.

However, the question now arises as to how to analyze unergatives and transitives with implicit objects. Consider first the verbs that seem to have an implicit object:

- (37) Ovi psi ujedaju.  
'These dogs bite (someone).'
- (38) Marko udara.  
'Marko is hitting (somebody).'
- (39) #Jabuke jedu.  
'The apples are eating (something).'

What is relevant here is that without *se*, theta-underspecification vanishes in Serbian, resulting in implicit object interpretations, identical to the ones available in English translations. These are clearly nom-acc patterns, and the question now is whether FASA can extend to those as well. If these structures have only one argument, First Argument, then they should be equally underspecified, and allow the same type of vagueness found with Tongan absolutives and Serbian middles. One possibility is to say that there are two types of grammars at work, erg-abs grammar, which in Serbian gets flagged by the grammatical word *se*, and the nom-acc grammar, which has the familiar VP/vP properties, and is

perhaps governed by some version of UTAH. I do not explore this analysis here, but just mention that it is in principle possible.

Here, I briefly explore the strong version of FASA, the one in which all structure, including transitive structures in nom-acc patterns, are subject to FASA. According to this approach, the examples in (37-39) would have to be analyzed as having null objects, forcing a projection of vP, and the generation of the only overt argument in vP. This would be the only way under FASA to insure that the subjects in these sentences cannot be interpreted as internal arguments. This would then help differentiate between absolutive patterns (including *se* middles), which cannot have null objects, or objects of any kind, and nom-acc patterns, which are obsessed with objects, and which posit (null) objects wherever possible, perhaps even with unergatives.

Unergatives are considered in Minimalism to be intransitive verbs whose only argument is an agent (e.g. *complain, laugh, sigh*), generated in the vP. However, FASA would not allow Merging an argument in the vP without first Merging an internal argument (First Argument) in the VP. It is conceivable that the (implied) cognate objects of unergatives are syntactically present in the VP, forcing the overt argument to be agent-like, generated in vP, as in (40) below. Alternatively, unergatives can be derived *à la* Hale and Kayser (1993, 2002), by an incorporation of a cognate noun into a light v, as in (41). In both cases, there would be an object in VP, allowing *Peter* to be generated in vP.

(40) ?Peter sighed a sigh of relief.

(41) Peter made a sigh.

I leave the resolution of this issue for future research, but if these structures are indeed to be analyzed as having null objects at some level, then this would follow from FASA.

Finally, the question arises regarding the semantic and syntactic contribution of *se* in this analysis. To put it bluntly, what is *se*, if not a reflexive, and if not an argument? One possibility is along the lines of Franks (1995) and Progovac (2005), where *se* is analyzed as an expletive pronoun. More recently, Progovac (2013, 2014, 2015) has analyzed *se* as a proto-transitive marker, a transitional stage between absolutive-like structures and accusativity, but still more on the absolutive side, at least in Serbian. Synchronically, perhaps *se* can be seen as flagging erg-abs

nature of syntactic structures, structures which can have no (accusative) objects. According to e.g. Kemmer (1994: 181), “the reflexive and the middle can be situated as semantic categories intermediate in transitivity between one-participant and two-participant events.”<sup>9</sup>

This analysis also sheds some light on the typical presence of *se* with dative subjects, which can be likened to ergatives:

- (42) Meni \*(se) pije kafa.  
 me<sub>DAT</sub> se drinks coffee<sub>NOM</sub>  
 ‘I feel like drinking coffee.’

Nominative marking on the “object” is like absolutive, being also the case of intransitive subjects, while dative introduces an external argument, akin to an ergative. According to Nash (1996: 171), ergative subjects, like dative subjects, cannot co-occur with structural accusative, but instead appear with absolutive/nominative “objects.” As noted in Trask (1979: 398), the ergative case is often identical to the genitive, dative, or locative.

#### 4 Conclusion

It is worth pursuing a unified analysis of *se* and its various uses in Serbian, but also in other languages. Some advantages of the proposed (absolutive-like) analysis of the only argument in *se* middles are the following. First, it captures the vagueness and overlaps in meaning in a straightforward fashion. Second, it derives reflexive interpretations (for free), without a need for treating *se* as a reflexive pronoun. Third, it establishes deep connections and common ground between erg-abs and nom-acc patterns. This analysis rejects UTAH, but UTAH is independently shown to be too strong.

On the theoretical side, I propose to replace UTAH with a principle such as FASA (First Argument, Second Argument), which can provide better guidance as to how to build syntactic structure across languages and structures. FASA is comparable in scope to the role of UTAH, but it

---

<sup>9</sup> Kemmer (1994: 184) also points out that middle systems are quite widespread, being found in a large number of genetically and areally divergent languages.

has better empirical coverage. In addition, FASA may be able to provide universal foundation (continuity) in building all structure: (i) in both nominal and verbal domains; (ii) in both nom-acc and erg-abs patterns; (iii) in both active and middle domains; (iv) in both intransitive and transitive domains.

As an added bonus, this approach provides the right kind of scaffolding for a gradualist approach to the evolution of syntax. According to Progovac (2013, 2014, 2015), the initial stage of proto-syntax was an intransitive two-word stage, which was absolutive-like. From there, languages diverged in their expression of transitivity in several different directions, including erg-abs and nom-acc. The postulated absolutive-like proto-layer is the most robust layer of structure, which still provides a necessary foundation for building more complex structures across languages and constructions. In this picture, middles in general can be seen as intermediate structures, straddling the boundary between transitivity and intransitivity, subjecthood and objecthood, passive and active.

## References

- Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.
- Aldridge, Edith. 2008. Generative approaches to ergativity. *Language and Linguistics Compass* 2/5: 966-995.
- Alexiadou, Artemis. 2001. *Functional Structure in Nominals: Nominalization and Ergativity*. Amsterdam: John Benjamins.
- Alexiadou, Artemis, and Melita Stavrou. 1998. On derived nominals in Greek. In B. Joseph, G. Horrocks, and I. Philippaki-Warbuton (eds.), *Themes in Greek Linguistics II*, 101-129. Amsterdam: John Benjamins.
- Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: Chicago University Press.
- Bobaljik, Jonathan. 1993. On ergativity and ergative unergatives. In C. Phillips (ed.), *Papers on Case and Agreement II. MIT Working Papers in Linguistics* 19, 45-88. Cambridge, MA: MITWPL.

- Borer, Hagit. 1994. The projection of arguments. In E. Benedicto and J. Runner (eds.), *University of Massachusetts Occasional Papers in Linguistics* 17, 19-47. Amherst: GLSA.
- Bottari, Piero. 1992. Romance passive nominals. *Geneva Generative Papers* 0.0: 66-80.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Comrie, Bernard. 1978. Ergativity. In W. P. Lehmann (ed.), *Syntactic Typology: Studies in the Phenomenology of Language*, 329-394. Austin: University of Texas Press.
- Dixon, Robert M. W. 1994. *Ergativity*. Cambridge Studies in Linguistics 69. Cambridge: Cambridge University Press.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67.3: 547-619.
- Dukes, Michael. 1998. Evidence for grammatical functions in Tongan. In M. Butt and T. Holloway King (eds.), *Proceedings of the LFG 98 Conference, the University of Queensland, Brisbane*, 144-161. Stanford: CSLI Publications.
- Fillmore, Charles. 1968. The case for case. In E. Bach and R. T. Harms (eds.), *Universals in Linguistic Theory*, 1-8. New York: Holt, Rinehart, and Winston.
- Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. Oxford: Oxford University Press.
- Hale, Ken, and Samuel Jay Kayser. 1993. On argument structure and lexical expression of syntactic relations. In K. Hale and J. Kayser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Brombert*, 53-110. Cambridge, MA: MIT Press.
- Hale, Ken, and Samuel Jay Kayser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge, MA: MIT Press.
- Kański, Zbigniew. 1986. *Arbitrary Reference and Reflexivity: A Generative Study of the Polish Pronoun się and its English Equivalents*. Katowice: Uniwersytet Śląski.
- Kemmer, Suzanne. 1994. Middle voice, transitivity, and the elaboration of events. In B. Fox and P. J. Hopper (eds.), *Voice: Form and Function*, 179-230. Amsterdam: John Benjamins.
- Kempson, Ruth M. 1977. *Semantic Theory*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

- Kratzer, Angelika. 1996. Severing the external argument from its verb. In J. Rooryck and L. Zaring (eds.), *Phrase Structure and the Lexicon*, 109–137. Dordrecht: Kluwer.
- Kratzer, Angelika. 2000. Building statives. In L. Conathan et al. (eds.), *Berkeley Linguistic Society* 26: 385–399.
- Laka, Itziar. 1993. Unergatives that assign ergative, unaccusatives that assign accusative. In J. Bobaljik and C. Phillips (eds.), *Papers on Case and Agreement I. MIT Working Papers in Linguistics* 18, 149–172. Cambridge, MA: MITWPL.
- Legate, Julie A. 2008. Morphological and abstract case. *Linguistic Inquiry* 39: 55–101.
- Levin, Juliette, and Diane Massam. 1985. Surface ergativity: case/theta relations reexamined. In S. Berman, J-W. Choe, and J. McDonough, *Proceedings of NELS 15*, 286–301. Amherst, MA: GLSA Publications.
- Massam, Diane. 2000. VSO and VOS: aspects of Niuean word order. In A. Carnie and E. Guilfoyle (eds.), *The Syntax of Verb Initial Languages*, 97–116. Oxford: Oxford University Press.
- Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19:153–197.
- Nash, Lea. 1995. *Argument Scope and Case Marking in SOV and in Ergative Languages: The Case of Georgian*. Ph.D. Dissertation, Université Paris 8.
- Nash, Lea. 1996. The internal ergative subject hypothesis. In K. Kusumoto et al. (eds.), *Proceedings of NELS 26*, 195–210.
- Otsuka, Yuko. 2011. PRO versus null *se*: case, tense, and empty categories in Tongan. *Syntax* 14.3: 265–296.
- Picallo, M. Carme. 1991. Nominals and nominalizations in Catalan. *Probus* 3: 279–316.
- Progovac, Ljiljana. 2005. *A Syntax of Serbian: Clausal Architecture*. Bloomington, IN: Slavica Publishers.
- Progovac, Ljiljana. 2013. Rigid syntax, rigid sense: absolutes/unaccusatives as evolutionary precursors. *Proceedings of FASL*, Bloomington, IN, 246–259. Ann Arbor: Michigan Slavic Publications.
- Progovac, Ljiljana. 2014. Degrees of complexity in syntax: a view from evolution. In F. J. Newmeyer and L. B. Preston (eds.), *Measuring Linguistic Complexity*, 83–102. Oxford: Oxford University Press.

- Progovac, Ljiljana. 2015. *Evolutionary Syntax*. In Press. Oxford: Oxford University Press.
- Rivero, Maria Luisa, and Milena Milojević-Sheppard. 2003. Indefinite reflexive clitics in Slavic: Polish and Slovenian. *Natural Language and Linguistic Theory* 21: 89–155.
- Tchekhoff, Claude. 1973. Some verbal patterns in Tongan. *The Journal of the Polynesian Society* 82.3: 281-292.
- Tchekhoff, Claude. 1979. From ergative to accusative in Tongan: An Example of Synchronic Dynamics. In F. Plank (ed.), *Ergativity: Towards a Theory of Grammatical Relations*, 407-418. London: Academic Press.
- Trask, Robert L. 1979. On the origins of ergativity. In F. Plank (ed.), *Ergativity: Towards a Theory of Grammatical Relations*, 385-404. London: Academic Press.
- Wiltschko, Martina. 2006. Inlocatives in Halkomelem Salish. In M. Kiyota, J. Thompson, and N. Yamane-Tanaka (eds.), *Papers from the Forty First Conference on Salish and Neighbouring Languages. University of British Columbia Working Papers in Linguistics* 18: 286-310.
- Woolford, Ellen. 1997. Four-way case systems: ergative, nominative, objective, and accusative. *Natural Language and Linguistic Theory* 15: 181-227.
- Woolford, Ellen. 2006. Lexical case, inherent case, and argument structure. *Linguistic Inquiry* 37: 111-130.

progovac@wayne.edu

## A Weakly Compositional Analysis of Distance Distributivity in Polish

*Adam Przepiórkowski*

Institute of Computer Science of the Polish Academy of Sciences

### 1 Introduction

Distance distributivity is a phenomenon where a distributive element such as *each* occurs at some structural distance from the nominal phrase that restricts it, as in *The boys have two apples each*.<sup>1</sup> In this sentence, the distributive element *each* occurs in the object position, while its restriction, *the boys*, is the subject of the sentence. This should be contrasted with the determiner uses of *each*, as in *Each boy has two apples*, where *each* combines directly with its restriction, as other ad-nominal quantifiers do.

There are various terminological conventions in the literature, e.g., Choe 1987 calls such uses of *each* “anti-quantifiers”, and Safir and Stowell 1988 call them “binominal”. Both terms are suboptimal: much subsequent literature attempts to describe such distributive elements (DEs) as more-or-less ordinary quantifiers (not as special “anti-quantifiers”) and it is clear now that DEs in other languages, including German and Polish, do not need two nominal expressions (*the boys* and *two apples* above) but – as shown by Moltmann 1991, 1997 – may quantify over events expressed by verbal constituents (hence, they are not “binominal”). In this paper we adopt the terminology of Zimmermann 2002, who introduced the term

---

<sup>1</sup> I am grateful for comments from the audience of FASL 23 and, especially, to the two anonymous reviewers. Also the acknowledgements of the accompanying papers Przepiórkowski 2014a,b carry over here. Needless to say, all remaining errors are my own.

“distance distributivity” (DD), and call the nominal phrase DE attaches to (*two apples* above) “distributive share” (DS), and the phrase expressing the set of entities which restrict DE (*the boys* above) – “distributive key” (DK).

It is clear that DD is not a completely uniform phenomenon and constraints on structural relations between DSs (and, hence, DEs which attach to them) and DKs differ across languages. Zimmermann 2002, working within the transformational paradigm of late 1990s, explains these differences in terms of inherent features of DEs and distinguishes two classes of DEs: those that have determiner features and, hence, must be c-commanded by a DP for these features to be licensed, and those that do not have such determiner features. The c-command requirement prevents the former from occurring in the (underlying) subject position, as in the unacceptable English \**One student each gave presents to the teachers* (Safir and Stowell 1988: 436, (26a)), while no such restriction is observed in case of the German DE *jeweils* or the Polish DE *po*:<sup>2</sup>

- (1) *Jeweils ein Offizier begleitete die Ballerinen nach Haus.*  
 DISTR one officer<sub>NOM</sub> accompanied the ballerinas<sub>ACC</sub> to home  
 (German)  
 ‘Each ballerina was accompanied home by one officer.’  
 (Zimmermann 2002: 27, (16))
- (2) *Z drzew spadło po jabłku.* (Polish)  
 from trees fell DISTR apple<sub>LOC</sub>  
 ‘An apple fell from each tree.’ (Łojasiewicz 1979: 154)

To the best of our knowledge, Zimmermann 2002 remains the only comprehensive syntactico-semantic analysis of DD of the kind observed in German and Polish. The aim of this paper is to show that Polish data do not comfortably fit the account of Zimmermann 2002 (§ 2) and to introduce a construction which that analysis cannot account for (§ 3). Due to lack of space, an alternative analysis is only sketched here (§ 4), but it is presented in gory technical detail in accompanying papers (Przepiórkowski 2014a,b).

<sup>2</sup> Unlike the binominal *each*, which always follows the DS, *jeweils* usually precedes the DS (Zimmermann 2002: § III.5.3), while *po* always occurs immediately before it.

## 2 Zimmermann 2002

Various problems, both empirical and theoretical, with earlier accounts of DD such as Choe 1987, Safir and Stowell 1988, Moltmann 1991, 1997 and Link 1998, are discussed and criticised in Zimmermann 2002, so here we only refer to Zimmermann's approach.

While Zimmermann 2002 remains the most comprehensive account of distance distributivity in German and cross-linguistically, it is not without problems. First, as noted by Dotlačil 2012, Zimmermann's assumption that the relation between DS and DK is expressed by a syntactic constituent (e.g., *have* in *Each boy have two apples*) does not always hold. For example, in *Alex and Sasha visited the capitals of three states each* there is no constituent corresponding exactly to *visited the capitals of*.

Second, the careful reader of footnotes will note that Polish (and Slavic in general) fits rather uncomfortably into Zimmermann's account.<sup>3</sup> In particular, it seems unexpected on Zimmermann's analysis that the *po* DE obligatorily precedes DS in Slavic. While a cross-linguistically valid analysis is highly desired, we feel that it should be guided by more detailed investigations into particular languages.

Third, although Zimmermann (2002) seeks to provide an account not relying on LF movement (and gives good arguments against the LF-based analysis of Safir and Stowell 1988), he acknowledges that his analysis must assume such covert movement for some occurrences of *jewels*, including (1) above (see his § 2.4.2 in ch. V, pp. 271ff.).

Fourth, in the course of providing the details of the syntactico-semantic analysis of DD across languages, Zimmermann (2002) is forced to introduce some non-standard mechanisms and make a number of assumptions contradicting the majority view in the framework hosting the analysis. One such mechanism is the "Type-Triggered  $\lambda$ -Abstraction" (p. 219), a very specific composition rule supplementing the more run-of-the-mill (Bittner 1994, Heim and Kratzer 1998) "Index-Triggered  $\lambda$ -Abstraction" and triggered in some contexts as "a last resort mechanism that only applies if all else fails". Among unusual assumptions there is also one about head

<sup>3</sup> See, e.g., fn. 86 on p. 131, fn. 87 on p. 132 (together with fn. 76 on p. 119). Also aspects of Korean seem problematic, e.g., fn. 83 on p. 134, fn. 98 on p. 143, fn. 21 on p. 276, as well as main text on p. 140.

movement out of adjuncts (fn. 76 on p. 119, but see also fn. 87 on p. 132), and another about event binding within VP (p. 226). Moreover, while putting much emphasis on the compositionality of the proposed analysis, some of its elements are not fully compositional, e.g., the context-driven insertion of various restrictions into the representation of the German DE *jewels*, e.g., in (176) on p. 232 (with the introduced relation  $\in$ ), in (184) on p. 234 (with the relation  $\subseteq$ ), and in (219) on p. 247 (with a new set variable). Such *ad hoc* mechanisms result in rather different representations of similar sentences (e.g., (171e) on p. 230 vs. (177) on p. 232).

Fifth, despite all this additional machinery, there are attested constructions that – as far as we see – cannot be handled in the approach of Zimmermann 2002. We introduce one such construction below.

### 3 Inverse Linking Distance Distributivity Construction

There is a construction problematic for previous analyses of DD, bearing certain resemblance to the inverse linking construction discussed in May 1985: 68ff. and Heim and Kratzer 1998: § 8.6, among others. In this construction – exemplified with the Polish sentence (3) (whose schematic syntactic structure is given in (4)) and the corresponding German sentence (5) – the distributive key is syntactically embedded within the distributive share:

- (3) Przybyło po 3 przedstawicieli 25 krajów. (Polish)  
 arrive<sub>PAST</sub> DISTR 3 representatives 25<sub>GEN</sub> countries<sub>GEN</sub>  
 ‘3 representatives arrived from each of 25 countries.’
- (4) Przybyło [po [3 [przedstawicieli [25 krajów]]]].
- (5) Jeweils 3 Abgeordnete aus 25 Ländern trafen ein. (German)  
 DISTR 3 representatives from 25 countries arrived  
 ‘3 representatives arrived from each of 25 countries.’  
 (Malte Zimmermann, p.c.)

The structure given in (4) is not controversial. The Polish DE *po* is analysed as – or simply assumed to be – a preposition (Łojasiewicz 1979, Franks 1995) which combines with the following nominal phrase.<sup>4</sup> Numerals are

<sup>4</sup> While there are reasons to postulate more than one DE *po* in Polish, they are all best analysed as heads (Przepiórkowski 2006, 2010, 2013, Przepiórkowski and Patejuk 2013),

also analysed as heads of numeral phrases in Polish on the basis of substitution tests and case assignment (Saloni and Świdziński 1998, Przepiórkowski 1999; but see also Franks 1995). In any case, whether the numeral phrase *3 przedstawicieli...* ‘three representatives...’ is taken to be headed by the numeral or by the noun, *25 krajów* ‘25 countries’ is an argument of *przedstawicieli* ‘representatives’, so – at least at the surface – it must be contained in the maximal projection of this noun.<sup>5</sup> Hence, the DK *25 krajów* ‘25 countries’ is contained within the DS *3 przedstawicieli 25 krajów* ‘3 representatives of 25 countries’.

Note that although (3) is a constructed example, analogous attested examples may easily be found in the National Corpus of Polish (NKJP; Przepiórkowski et al. 2012; <http://nkjp.pl/>) and in the Internet, e.g. (constraining our search to the same relational noun):<sup>6</sup>

- (6) ...proponował po dwóch przedstawicieli miasta  
 proposed<sub>M1</sub> DISTR two<sub>ACC.M1</sub> representatives<sub>ACC.M1</sub> city<sub>GEN.N</sub>  
 i ComArchu...  
 and ComArch<sub>GEN.M3</sub>  
 ‘...he proposed two representatives for each of the city and  
 ComArch.’ (NKJP)
- (7) W skład jury wchodzi po 2  
 into make-up<sub>ACC</sub> jury<sub>GEN</sub> enters DISTR two<sub>ACC.M1</sub>  
 przedstawicieli organizatorów konkursu.  
 representatives<sub>ACC.M1</sub> organisers<sub>GEN.M1</sub> competition<sub>GEN.M3</sub>  
 ‘2 representatives of each of the organisers of the competition belong  
 to / constitute the jury.’  
 ([http://zporuszcza.polaniec.pl/index\\_pliki/bezpieczna\\_skola.pdf](http://zporuszcza.polaniec.pl/index_pliki/bezpieczna_skola.pdf))

so treating them all as prepositions is a reasonable first approximation. Note that they differ from the prefix *po-* (Bogusławski 1993, <http://pinon.sdf-eu.org/covers/dpp.html>), which has a related but different distributive meaning.

<sup>5</sup> Note that this is an island, presumably also for covert movement:

- (i) \*Czego przybyło po 3 przedstawicieli?  
 what<sub>GEN</sub> arrive<sub>PAST</sub> DISTR 3 representatives

<sup>6</sup> In the glosses, M1 stands for the human-masculine gender and M3 – for human-inanimate, assuming the 5 Polish genders proposed in Mańczak 1956. Other morphosyntactic symbols follow the Leipzig Glossing Rules.

In the examples above, the DK is an argument of the relational noun *przedstawiciel* ‘representative’, which heads the DS. Examples of DK adjuncts to heads of DSs are also easy to find, but they are not discussed here, as they do not pose a particular problem for Zimmermann’s analysis.

Let us attempt to analyse such constructions. The cross-linguistic denotation of DEs proposed by Zimmermann 2002: 122 is given below:

$$(8) \quad \llbracket \text{DE} \rrbracket = \lambda P. \forall z [(z \in Z_i) \rightarrow \exists x [P(x) \wedge R_j(z, x)]]$$

In this representation,  $P$  stands for the property expressed by the DS. For example, in *The boys have two apples each*,  $P$  would be the property of being a set of two apples; let us schematically represent this property as  $\lambda x. 2apples(x)$ . Given the representation of *each* in (8), *two apples each* receive the following representation (via functional application):

$$(9) \quad \llbracket \text{two apples each} \rrbracket = \forall z [(z \in Z_i) \rightarrow \exists x [2apples(x) \wedge R_j(z, x)]]$$

$Z_i$  and  $R_j$  are variables which are coindexed with, respectively, the DK (*the boys*) and the relation between the DK and the DS (*have*). Via the “Index-Triggered  $\lambda$ -Abstraction” (Zimmermann 2002: 217), when the phrase *two apples each* with the representation in (9) is a constituent-tree sister of a node with index  $j$ , expressing a 2-place relation such as  $\lambda z \lambda x. have(z, x)$ , (9) can be transformed to (10) below and then be applied to the *have*-relation to render (11) for the verbal phrase (VP) *have two apples each*.

$$(10) \quad \llbracket \text{two apples each} \rrbracket = \lambda R_j. \forall z [(z \in Z_i) \rightarrow \exists x [2apples(x) \wedge R_j(z, x)]]$$

$$(11) \quad \llbracket \text{have two apples each} \rrbracket = \forall z [(z \in Z_i) \rightarrow \exists x [2apples(x) \wedge have(z, x)]]$$

Similarly, when the VP is a sister to *the boys* indexed with  $i$ ,  $\lambda$ -abstraction is licensed again (see (12)) resulting in a function that can be applied to the meaning of *the boys*, giving the meaning of the sentence in (13).

$$(12) \quad \llbracket \text{have two apples each} \rrbracket = \lambda Z_i. \forall z [(z \in Z_i) \rightarrow \exists x [2apples(x) \wedge have(z, x)]]$$

$$(13) \quad \llbracket \text{the boys have two apples each} \rrbracket = \forall z [(z \in \llbracket \text{the boys} \rrbracket) \rightarrow \exists x [2apples(x) \wedge have(z, x)]]$$

Returning to (3), its analogous desired representation is given in (14):<sup>7</sup>

<sup>7</sup> We ignore here the event variable introduced by the verb and bound via existential closure at the end of the derivation, but see below.

$$(14) \llbracket \text{przybyło po 3 przedstawicieli 25 krajów} \rrbracket = \\ \forall z[(z \in \llbracket 25 \text{ countries} \rrbracket) \rightarrow \exists x[3\text{representatives}(x, z) \wedge \text{arrived}(x)]]$$

How can it be derived, assuming the representation of the DE *po* as in (8)? For the sake of the argument, let us give as much leeway to Zimmermann's approach as possible and assume that any kind of LF-movement is allowed, even in violation of island constraints.

The DE must first combine with its sister – either the DS or a trace resulting from movement. But since traces are of the semantic type  $\langle e \rangle$ , and DE expects a property of type  $\langle e, t \rangle$ , no movement of the whole DS is possible.<sup>8</sup> On the other hand, we have to assume the LF-movement of *25 krajów* ‘25 countries’; otherwise, if the whole *3 przedstawicieli 25 krajów* ‘3 representatives of 25 countries’ is consumed as *P*, there would be no DK to subsequently provide the meaning of  $Z_i$ . So the only way to proceed with the analysis is to assume the following schematic structure at LF:

$$(15) [25 \text{ krajów}]_i; [\text{przybyło [po 3 przedstawicieli } t_i]]$$

The DE *po* expects a property, so let us assume the representation of the argument of *po* as in (16) and the result of its combination with the representation of *po* given in (8) – as in (17):

$$(16) \llbracket 3 \text{ przedstawicieli } t_i \rrbracket = \lambda x.3\text{representatives}(x, z_i)$$

$$(17) \llbracket \text{po 3 przedstawicieli } t_i \rrbracket = \\ \forall z[(z \in Z_i) \rightarrow \exists x[3\text{representatives}(x, z_i) \wedge R_j(z, x)]]$$

This representation is already getting incoherent, as it now involves two variables coindexed with *25 krajów* ‘countries’ –  $z_i$  of type  $\langle e \rangle$  and  $Z_i$  of type  $\langle e, t \rangle$ . Obviously, instead of the variable  $z_i$ , the second argument of *3representatives* should be the variable  $z$  bound by the universal quantifier. Hence, the existential closure over  $z_i$  in (16) would not help here either. Even if this problem could somehow be solved, there is no binary relation that could provide the meaning of the binary  $R_j$  – *przybyło* ‘arrived’ is a unary predicate.<sup>9</sup>

<sup>8</sup> Also, the “Index-Triggered  $\lambda$ -Abstraction” is not applicable in this configuration.

<sup>9</sup> Zimmermann 2002: 226, fn. 67, considers the possibility of a *family* of denotations for DE, with  $R_j$  of different arities greater or equal to 2. Perhaps this idea could be extended even further, to  $R_j$  of arity 1, but this would not solve the problem of the incoherent

Since there is no constituent in the representation of (3) that expresses a binary relation needed to provide the denotation of  $R_j$ , let us attempt to analyse this sentence in a way analogous to the German (18), which also involves a unary predicate (the idiomatic *keep watch*).

- (18) *Jeweils zwei Jungen standen Wache.* (German)  
 DISTR two boys<sub>NOM</sub> kept guard  
 ‘Two boys kept watch at a time.’ (Zimmermann 2002: 249)

Here, the distribution is over events; the target denotation can be paraphrased as “for all elements  $z$  of a contextually salient set (of events)  $Z_i$ , there is a set of two boys  $x$ , and an event  $e$ , such that the elements of  $x$  kept watch in  $e$ , and event  $e$  is related to event  $z$  by a temporal, causal, subpart, or other contextual relation  $R$ ” (Zimmermann 2002: 261):

- (19)  $\llbracket \text{jeweils zwei Jungen standen Wache} \rrbracket =$   
 $\forall z[(z \in Z_i) \rightarrow \exists x[2\text{boys}(x) \wedge \exists e[\text{kept\_watch}(x, e) \wedge R(e, z)]]]$

In order to derive this representation, Zimmermann 2002: 259 assumes the standard representation of *jeweils* in (8), which gives rise to the following representation of *jeweils zwei Jungen*:

- (20)  $\llbracket \text{jeweils zwei Jungen} \rrbracket = \forall z[(z \in Z_i) \rightarrow \exists x[2\text{boys}(x) \wedge R_j(z, x)]]]$

The meaning of the verbal component  $t_1$  *standen Wache*, with  $t_1$  representing the trace of the subject *jeweils zwei Jungen*, is less obvious (here, after applying  $\lambda$ -abstractions):

- (21)  $\llbracket t_1 \text{ standen Wache} \rrbracket = \lambda x_1 \lambda e_i. \exists e[\text{kept\_watch}(x_1, e) \wedge R(e, e_i)]$

The variable  $x_1$  in (21) represents the subject of the predicate, while  $R$  represents a contextually given relation between the event  $e$  predicated by the verb and an event  $e_i$  in the preceding discourse (Zimmermann 2002: 260). This way the denotation of *standen Wache* ‘stood guard’ is a 2-place predicate, as expected. With this representation of the verbal predicate, the result of  $\lambda$ -abstraction of  $R_j$  in (20) applied to (21) is (19) above.

How can this analysis be carried over to (3)? First of all, let us assume the representation of *przybyło* ‘arrived’ analogous to that in (21):

---

representation in (17). Also, the analysis proposed below does not have to assume a family of representations of the DE *po*.

$$(22) \llbracket t_1 \text{ przybyło} \rrbracket = \lambda e_i \lambda x_1. \exists e [\text{arrived}(x_1, e) \wedge R(e, e_i)]$$

Assuming  $\lambda$ -abstraction over  $R_j$  in (17) above and subsequent application of the resulting function to the denotation in (22), the denotation in (23) below results. Combining this representation with that of *25 krajów* ‘25 countries’, we get (24).

$$(23) \llbracket \text{przybyło po 3 przedstawicieli } t_i \rrbracket = \\ \forall z [(z \in Z_i) \rightarrow \exists x [3\text{representatives}(x, z_i) \wedge \exists e [\text{arrived}(x, e) \wedge R(e, z)]]]$$

$$(24) \llbracket \text{przybyło po 3 przedstawicieli 25 krajów} \rrbracket = \\ \forall z [(z \in \llbracket 25 \text{ krajów} \rrbracket) \rightarrow \\ \exists x [3\text{representatives}(x, z_i) \wedge \exists e [\text{arrived}(x, e) \wedge R(e, z)]]]$$

This representation is close to the correct one but – unfortunately – it again contains the free variable  $z_i$  which should really be bound by the universal quantifier  $\forall z$ .

Note that the target representation in (14) *could* be derived from the DE denotation in (8), but such a derivation would violate Zimmermann’s basic assumptions about constituency and surface compositionality: the DE *po* would first have to combine with *przybyło* ‘arrived’ in (25) rendering the denotation in (26), then with *3 przedstawicieli* ‘3 representatives’ treated as a binary relation  $\lambda z \lambda x. 3\text{representatives}(x, z)$ , resulting in (27), and then with *25 krajów* ‘25 countries’, resulting in (28):

$$(25) \llbracket \text{przybyło} \rrbracket = \lambda x. \text{arrived}(x)$$

$$(26) \llbracket \text{przybyło po} \rrbracket = \forall z [(z \in Z_i) \rightarrow \exists x [\text{arrived}(x)] \wedge R_j(x, z)]$$

$$(27) \llbracket \text{przybyło po trzech przedstawicieli} \rrbracket = \\ \forall z [(z \in Z_i) \rightarrow \exists x [\text{arrived}(x) \wedge 3\text{representatives}(x, z)]]$$

$$(28) \llbracket \text{przybyło po trzech przedstawicieli 25 krajów} \rrbracket = \\ \forall z [(z \in \llbracket 25 \text{ countries} \rrbracket) \rightarrow \exists x [\text{arrived}(x) \wedge 3\text{representatives}(x, z)]]$$

In summary, whether treating (3) as an instance of distribution over entities (25 countries) or over events (arrivals), we do not see a way to derive an acceptable meaning of this sentence, given the approach of Zimmermann 2002. This, combined with the reservations expressed in § 2, calls for a new approach to distance distributivity in Polish; such an approach is sketched below.

#### 4 An Outline of an Alternative Account

The main idea of an alternative account, more fully described in Przepiórkowski 2014a, from which this section draws heavily, is this: the semantic impact of *po* activates only once the distributive share combines semantically with the verb and creates a property. For example, in case of (3), the meaning of *przybyło 3 przedstawicieli*, ‘ $\lambda Y$ . 3 representatives of  $Y$  arrived’, is derived first. Then, the meaning of *po* combines with this property, let us call it  $S$ , holding of some set  $Y$ , and produces a new property, which is just like  $S$  but holds of each element of  $Y$  individually: ‘ $\lambda Y$ . for each element  $y$  of  $Y$ , 3 representatives of  $y$  arrived’. Finally, this new property combines with the distributive key *25 krajów* ‘25 countries’, resulting in the meaning: ‘for each of 25 countries, 3 representatives arrived’.

This idea relies on the possibility to combine the meaning of *po* with the property ‘ $\lambda Y$ . 3 representatives of  $Y$  arrived’ expressed by *przybyło 3 przedstawicieli*, rather than with the meaning of the syntactic sister of *po*. It would be difficult to implement this idea in a framework that understands compositionality narrowly, as in these two recent formulations:

- *The meaning of a complex expression functionally depends on the meanings of its immediate parts and the way in which they are combined.* (Zimmermann 2012: 82)
- *The meaning of a complex expression is determined by its immediate structure and the meanings of its immediate constituents.* (Szabó 2012: 79)

This is the usual understanding of compositionality – unquestioned in transformational approaches – but it is not the only one. In fact, as discussed in detail in Janssen 2012 and Szabó 2012, the provenance of this – originally massively ambiguous – principle is murky (it should probably *not* be attributed to Frege, but rather to his student, Carnap 1947), there are no strong fundamental – as opposed to methodological – arguments for adopting it, and the reasons for its widespread use are mostly technical.

As noted already in 1987 (see the reprint, Halvorsen 1995: 295), compositionality should be replaced in constraint-based theories by *systematicity*, a method of automatic derivation of utterance interpretations from the lexical information and any rules of the interpretation scheme.<sup>10</sup> The

<sup>10</sup> This emphasis on the meanings of utterances rather than the meanings of arbitrary

alternative analysis of DD in Polish is couched in just such a constraint-based theory, namely, Lexical Functional Grammar (LFG; Bresnan 2001, Dalrymple 2001), coupled with a resource-based approach to meaning composition, namely, Glue Semantics (Dalrymple 1999, 2001). The latter explicitly adopts this weak notion of compositionality, *where the meaning of a sentence depends on the meanings of its words and the way these are combined, but where syntactic structure and lexical semantics may not fully specify either* (Crouch and van Genabith 1999: 122).

In traditional approaches to compositionality (e.g., Heim and Kratzer 1998), meanings combine when they are expressed by siblings in a constituency tree. By contrast, in LFG + Glue, meanings combine based on f(unctional)-structures, rather than on c(onstituent)-structures, and meaning representations are paired with glue formulae specifying how these meanings combine with which other meanings. Any pair consisting of a meaning representation and a glue formula is called a *meaning constructor*.

For example, the glue part of the meaning constructor for various forms of *yawn* is:

$$(29) (\uparrow \text{SUBJ})_{\sigma} \multimap \uparrow_{\sigma}$$

As usual in LFG, the up arrow  $\uparrow$  in a lexical entry denotes the f-structure of the word,  $(\uparrow \text{SUBJ})$  denotes the f-structure of the subject of this word, and  $\sigma$  is a function from f-structures to s(emantic)-structures. In effect, (29) says that, by consuming the s-structure corresponding to the subject of *yawn*, we may produce the s-structure corresponding to *yawn* and, hence, to the whole clause headed by *yawn* (in LFG heads normally share their f-structure with their projections).

This mode of composition remains true regardless of specific tree configurations. For example, when *yawn* is a complement of a control verb, its covert subject is never realised in the c(onstituent)-structure, according to standard LFG analyses, but it is still present in its f-structure, as the value of the SUBJ attribute, so (29) is still relevant.

The other part of the meaning constructor is a formula in any language that allows application and abstraction, e.g., the language of the first-order predicate logic with lambda calculus. For example, the meaning of *David*

---

syntactic components is strongly related to the principle of *contextuality*, a postulate that does deserve to be called *Frege's principle*; see Janssen 2012 for discussion.

can be defined as a logical constant, *David*, and the meaning of *yawned* can be defined as usual, as  $\lambda X.yawn(X)$  (ignoring event variables, semantic roles, tense and aspect, etc.). In complete meaning constructors, the meaning part is separated from the glue part by the uninterpreted colon character (:), so the complete meaning constructors for *David* and *yawned* are as in the second lines of the following lexical entries:

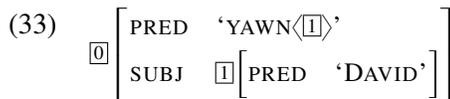
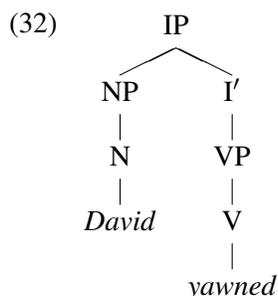
$$(30) \quad David \quad N \quad (\uparrow \text{ PRED}) = \text{'DAVID'}$$

$$David : \uparrow_{\sigma}$$

$$(31) \quad yawned \quad V \quad (\uparrow \text{ PRED}) = \text{'YAWN<SUBJ>'}$$

$$\lambda X.yawn(X) : (\uparrow \text{ SUBJ})_{\sigma} \multimap \uparrow_{\sigma}$$

According to these lexical entries and standard LFG constituency rules, *David yawned* receives the c-structure displayed in (32) and the f-structure in (33); moreover, given this f-structure, meaning constructors are instantiated as in (34):



$$(34) \quad \mathbf{[David]} \quad David : \boxed{1}_{\sigma}$$

$$\mathbf{[yawned]} \quad \lambda X.yawn(X) : \boxed{1}_{\sigma} \multimap \boxed{0}_{\sigma}$$

Now, using one of the proof rules of Glue Semantics, namely, the Implication Elimination rule in (35), and performing the usual  $\beta$ -reduction, the meaning of *David yawned* may be derived from the meaning constructors in (34) as shown in (36):

$$(35) \quad \frac{a : A \quad f : A \multimap B}{f(a) : B} \multimap_{\mathcal{E}}$$

$$(36) \quad \frac{David : \boxed{1}_{\sigma} \quad \lambda X.yawn(X) : \boxed{1}_{\sigma} \multimap \boxed{0}_{\sigma}}{yawn(David) : \boxed{0}_{\sigma}} \multimap_{\mathcal{E}}$$

Since both meaning resources introduced by lexical items,  $\boxed{1}_\sigma$  and  $\boxed{1}_\sigma \multimap \boxed{0}_\sigma$ , were consumed in this proof, and the only meaning resource produced,  $\boxed{0}_\sigma$ , corresponds to the f-structure of the whole sentence, this is a valid proof that the meaning side of the whole sentence is *yawn(David)*.

Obviously, we cannot do justice to Glue Semantics within the confines of this paper; the above is only meant to make the analysis below more accessible to motivated readers not familiar with this approach. The best introduction to Glue Semantics may still be found in the classical LFG textbook of Dalrymple 2001, on which the above exposition is based.

Let us now return to the problematic distance distributivity construction exemplified by (3), repeated below:

- (3) Przybyło po 3 przedstawicieli 25 krajów.  
 arrive<sub>PAST</sub> DISTR 3 representatives 25<sub>GEN</sub> countries<sub>GEN</sub>  
 ‘3 representatives arrived from each of 25 countries.’

The lexical entry for *przybyło* ‘arrived’ matches that of *yawned* given in (31) above (note that we ignore the event variable again, solely for reasons of simplicity):

- (37) *przybyło* V  $(\uparrow \text{PRED}) = \text{‘ARRIVE<SUBJ>’}$   
 $\lambda X. \text{arrive}(X) : (\uparrow \text{SUBJ})_\sigma \multimap \uparrow_\sigma$

The meaning constructors of common nouns are a little less obvious:

- (38) *krajów* N  $(\uparrow \text{PRED}) = \text{‘COUNTRIES’}$   
 $\lambda X. \text{country}^s(X) \wedge |X| > 1 : (\uparrow_\sigma \text{VAR}) \multimap (\uparrow_\sigma \text{RESTR})$

First, we follow Dotlačil 2012 and earlier work on treating type *e* objects as sets, and properties – as sets of such sets. For example, *country<sup>s</sup>* is the property of being a non-empty set of countries – either a singleton or a set of higher cardinality (the superscript *s* indicates the possible plural) – and  $\lambda X. |X| > 1 \wedge \text{country}^s(X)$  is the property of being a set of at least two countries. On this view, the standard inclusion relation  $\subseteq$  is defined on type *e* objects. Second, the glue side shows that semantic structures may have some internal structure: s-structures of common nouns, which are of type  $\langle e, t \rangle$ , have the attributes VAR and RESTR, representing a variable (of type *e*) and a restriction on that variable (of type *t*); cf. Dalrymple 2001: 250–253.

Entries of relational nouns are just like those of common nouns, but they add a specification of an internal argument:

$$(39) \quad \textit{przedstawiciele} \quad \text{N} \quad (\uparrow \text{PRED}) = \text{'REPRESENTATIVES<OBJ>'} \\ \lambda Y. \lambda X \textit{representative}^s(X, Y) \wedge |X| > 1 : \\ (\uparrow \text{OBJ})_\sigma \multimap [(\uparrow_\sigma \text{VAR}) \multimap (\uparrow_\sigma \text{RESTR})]$$

The meaning constructor of (39) differs from that of (38) and other non-relational nouns in the additional requirement of the semantic resource corresponding to the argument of the noun.

Further, simplifying somewhat, we treat cardinals as existential quantifiers:

$$(40) \quad 3 \quad \text{Num} \quad (\uparrow \text{SPEC}) = 3 \\ \lambda R. \lambda S. \textit{exists}(Y, |Y| = 3 \wedge R(Y), S(Y)) : \\ [(\uparrow_\sigma \text{VAR}) \multimap (\uparrow_\sigma \text{RESTR})] \multimap [\forall H. [\uparrow_\sigma \multimap H] \multimap H]$$

$$(41) \quad 25 \quad \text{Num} \quad (\uparrow \text{SPEC}) = 25 \\ \lambda R. \lambda S. \textit{exists}(Y, |Y| = 25 \wedge R(Y), S(Y)) : \\ [(\uparrow_\sigma \text{VAR}) \multimap (\uparrow_\sigma \text{RESTR})] \multimap [\forall H. [\uparrow_\sigma \multimap H] \multimap H]$$

As common in LFG and Glue Semantics, generalised quantifiers are represented here as *pair quantifiers*, that is, as relations between an individual and two propositions involving that individual, so that *Someone yawned* has the basic representation  $\textit{exists}(X, \textit{person}(X), \textit{yawn}(X))$  (Dalrymple 2001: 227). In our setup, cardinality is additionally specified, so – for example – *Two people yawned* will have the following representation:  $\textit{exists}(X, \textit{person}^s(X) \wedge |X| = 2, \textit{yawn}(X))$ .

Finally, we assume the following lexical entry of *po*:

$$(42) \quad \textit{po} \quad \text{P} \quad (\uparrow \text{PRED}) = \text{'PO<OBJ>'} \\ (\uparrow \text{OBJ})_\sigma = \uparrow_\sigma \\ \lambda S. \lambda Z. \textit{all}(X, |X| = 1 \wedge X \subset Z, S(X)) : \\ \forall G, H. [G \multimap H] \multimap [G \multimap H]$$

Observe that *po* is analysed as a preposition here (but see fn. 4). The import of the second line,  $(\uparrow \text{OBJ})_\sigma = \uparrow_\sigma$ , will be explained below. The third line – the meaning part of the meaning constructor – says that *po* takes a property *S* and returns a property that holds of *Z* if and only if *S* holds of all singleton (proper) subsets of *Z*. Finally, the glue part in the fourth line says that *po*

is an identity function on semantic resources corresponding to properties: it consumes any resource  $[G \multimap H]$  in order to produce the same resource. Since  $G$  and  $H$  may be any semantic resources (of appropriate types), this analysis is much too permissive as it stands – it is appropriately constrained in Przepiórkowski 2014b.

We do not present here syntactic rules which serve to build the constituency structure of the running example, as they are trivial and of secondary importance to the current analysis. Crucially, we assume that these rules – together with the lexical entries above – lead to the following functional structure for the complete sentence in (3):

$$(43) \left[ \begin{array}{l} \text{[0]} \left[ \begin{array}{l} \text{PRED 'ARRIVED' } \langle \text{[1]} \rangle \\ \text{SUBJ [1]} \left[ \begin{array}{l} \text{PRED 'PO' } \langle \text{[2]} \rangle \\ \text{OBJ [2]} \left[ \begin{array}{l} \text{SPEC '3'} \\ \text{PRED 'REPRESENTATIVE' } \langle \text{[3]} \rangle \\ \text{OBJ [3]} \left[ \begin{array}{l} \text{SPEC '25'} \\ \text{PRED 'COUNTRY'} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right]$$

While there are syntactic reasons to assume that numerals take the following NPs as their arguments, we simplify here by treating the numeral and the following noun as co-heads. Hence, both the lexical entry for *krajów* in (38) and the lexical entry for 25 in (41) contribute to the innermost feature structure in (43), marked with the label [3]. In other words, the  $\uparrow$  variable in these lexical entries instantiates to [3], so the meaning constructors instantiate, respectively, to:

$$(44) \text{ [countries]} \\ \lambda X. \text{country}^s(X) \wedge |X| > 1 : ([3]_{\sigma} \text{VAR}) \multimap ([3]_{\sigma} \text{RESTR})$$

$$(45) \text{ [25]} \\ \lambda R. \lambda S. \text{exists}(X, |X| = 25 \wedge R(X), S(X)) : \\ [([3]_{\sigma} \text{VAR}) \multimap ([3]_{\sigma} \text{RESTR})] \multimap [\forall H. [3]_{\sigma} \multimap H] \multimap H]$$

Using the Implication Elimination rule in (35), and performing the usual  $\beta$ -reduction, these meanings combine to:<sup>11</sup>

<sup>11</sup> In (35), substitute “ $([3]_{\sigma} \text{VAR}) \multimap ([3]_{\sigma} \text{RESTR})$ ” for  $A$ , “ $\forall H. [3]_{\sigma} \multimap H$ ” for  $B$ , “ $\lambda X. \text{country}^s(X) \wedge |X| > 1$ ” for  $a$  and “ $\lambda R. \lambda S. \text{exists}(X, |X| = 25 \wedge R(X), S(X))$ ” for  $f$ .

(46) **[25-countries]**

$$\lambda S.exists(X, |X| = 25 \wedge country^s(X), S(X)) : \forall H. [\boxminus_{\sigma} \multimap H] \multimap H$$

Similarly, lexical entries (39) (for *przedstawiciele*) and (40) (for *3*) contribute to the construction of f-structure  $\boxminus$ , so  $\uparrow$  in those entries instantiates to  $\boxminus$  and, hence,  $(\uparrow \text{ OBJ})$  instantiates to  $\boxminus$ :

(47) **[representatives]**

$$\lambda Y. \lambda X. representative^s(X, Y) \wedge |X| > 1 : \\ \boxminus_{\sigma} \multimap [(\boxminus_{\sigma} \text{ VAR}) \multimap (\boxminus_{\sigma} \text{ RESTR})]$$

(48) **[3]**

$$\lambda R. \lambda S.exists(X, |X| = 3 \wedge R(X), S(X)) : \\ [(\boxminus_{\sigma} \text{ VAR}) \multimap (\boxminus_{\sigma} \text{ RESTR})] \multimap [\forall H. [\boxminus_{\sigma} \multimap H] \multimap H]$$

(49) **[3-representatives]**

$$\lambda Y. \lambda S.exists(X, |X| = 3 \wedge representative^s(X, Y), S(X)) : \\ \forall H. \boxminus_{\sigma} \multimap [(\boxminus_{\sigma} \multimap H) \multimap H]$$

In fact, in order to derive **[3-representatives]** from **[3]** and **[representatives]**, another standard proof rule is needed, Implication Introduction (Dalrymple 2001: 236, Asudeh 2012: 79), which we will not cite here for lack of space. Instead we note that the proof captures the intuition behind the function composition in Categorical Grammar (cf., e.g., Steedman 2000: 40), where functions  $X/Y$  and  $Y/Z$  may compose into  $X/Z$ .

Given the f-structure (43), the meaning constructor of *przybyło* in (37) instantiates to **[arrived]**, as  $\uparrow$  instantiates to  $\boxminus$  and, hence,  $(\uparrow \text{ SUBJ})$  – to  $\boxminus$ :

(50) **[arrived]**

$$\lambda X.arrived(X) : \boxminus_{\sigma} \multimap \boxminus_{\sigma}$$

Finally, the meaning constructor of *po* in (42) contains no  $\uparrow$  symbols, only variables  $G$  and  $H$  matching any (appropriately typed) resource, but there is another line in this lexical entry,  $(\uparrow \text{ OBJ})_{\sigma} = \uparrow_{\sigma}$ , which – given (43) – instantiates to  $\boxminus_{\sigma} = \boxminus_{\sigma}$ . The intuition behind this meaning constructor and this constraint is that *po* makes no semantic impact where it occurs – it equates its semantic resource  $\boxminus_{\sigma}$  with that of its argument  $\boxminus_{\sigma}$  – but it contributes the distributive meaning constructor which activates elsewhere in the semantic derivation.

Given that  $\boxed{2}_\sigma = \boxed{1}_\sigma$ , and substituting  $\boxed{0}_\sigma$  for  $H$  in the meaning constructor **[3-representatives]** (49), this constructor may be combined with **[arrived]** (50) (again, via a few proof steps, including Implication Introduction), rendering:

(51) **[arrived-3-representatives]**

$$\lambda Y. \text{exists}(X, |X| = 3 \wedge \text{representative}^s(X, Y), \text{arrived}(X)) : \boxed{3}_\sigma \multimap \boxed{0}_\sigma$$

After substituting  $G$  and  $H$  with, respectively,  $\boxed{3}_\sigma$  and  $\boxed{0}_\sigma$  in the meaning constructor for  $po$  in (42), **[arrived-3-representatives]** combines with this meaning constructor directly, resulting in:

(52) **[distr-arrived-3-representatives]**

$$\lambda Z. \text{all}(Y, |Y| = 1 \wedge Y \subset Z, \text{exists}(X, |X| = 3 \wedge \text{representative}^s(X, Y), \text{arrived}(X))) : \boxed{3}_\sigma \multimap \boxed{0}_\sigma$$

Finally, substituting  $\boxed{0}_\sigma$  for  $H$  in the meaning constructor for the quantifier phrase *25 krajów*, given in (46), it combines directly with the above meaning constructor (52), rendering the intended meaning of the whole functional structure  $\boxed{0}$ :

(53) **[25-countries-distr-arrived-3-representatives]**

$$\begin{aligned} &\text{exists}(Z, |Z| = 25 \wedge \text{country}^s(Z), \\ &\quad \text{all}(Y, |Y| = 1 \wedge Y \subset Z, \\ &\quad \quad \text{exists}(X, |X| = 3 \wedge \text{representative}^s(X, Y), \text{arrived}(X)))) : \boxed{0}_\sigma \end{aligned}$$

## 5 Conclusion

One of the first influential analyses of distance distributivity, Choe 1987, is not compositional. Further work – of which Zimmermann 2002 is a premiere example – tried to provide compositional analyses of the phenomenon at the syntax-semantics interface. While it remains the most comprehensive analysis of DD of the kind also observed in Slavic languages, it is not without problems and limitations, discussed in § 2 and § 3. The alternative analysis, outlined in § 4, is compositional in a rather weak sense, but it is systematic: the meaning of an utterance is derived from the meanings of lexical items and the way they combine. Even if not all technical details of the presented analysis are transparent to readers not previously exposed to LFG and Glue Semantics, it should be clear that the advantage of this relaxed approach to compositionality is a much simpler

syntax: no *ad hoc* (covert movement, etc.) rules are needed to account for the semantic complexity. Instead, the complexity resides exactly where it should: in the lexical entries of semantically complex items.

### References

- Asudeh, Ash. 2012. *The Logic of Pronominal Resumption*. Oxford: Oxford University Press.
- Bittner, Maria. 1994. Cross-linguistic Semantics. *Linguistics and Philosophy* 17, 53–108.
- Bogusławski, Andrzej. 1993. Polskie *po-* dystrybutywne i sprawa granic słowotwórstwa. *Zeszyty Naukowe — Prace Językoznawcze* 17–18, 61–68, uniwersytet Gdański.
- Bresnan, Joan. 2001. *Lexical-Functional Syntax*. Blackwell Textbooks in Linguistics, Malden, MA: Blackwell.
- Carnap, Rudolph. 1947. *Meaning and Necessity*. Chicago: University of Chicago Press.
- Choe, Jae-Woong. 1987. *Anti-quantifiers and a Theory of Distributivity*. Ph. D. dissertation, University of Massachusetts, Amherst.
- Crouch, Richard and van Genabith, Josef. 1999. Context Change, Underspecification, and the Structure of Glue Language Derivations. In Dalrymple (1999), pages 117–189.
- Dalrymple, Mary (ed.). 1999. *Semantics and Syntax in Lexical Functional Grammar: The Resource Logic Approach*. Cambridge, MA: The MIT Press.
- Dalrymple, Mary. 2001. *Lexical Functional Grammar*. San Diego, CA: Academic Press.
- Dotlačil, Jakub. 2012. Binominal *each* as an Anaphoric Determiner: Compositional Analysis. In Ana Aguilar Guevara, Anna Chernilovskaya and Rick Nouwen (eds.), *Proceedings of Sinn und Bedeutung 16: Volume 1*, volume 16 of *MIT Working Papers in Linguistics*, pages 211–224.
- Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. New York: Oxford University Press.
- Halvorsen, Per-Kristian. 1987. Situation Semantics and Semantic Interpretation in Constraint-Based Grammars. Technical Report CSLI-TR-101, Stanford University, Stanford, CA.

- Halvorsen, Per-Kristian. 1995. Situation Semantics and Semantic Interpretation in Constraint-Based Grammars. In Mary Dalrymple, Ronald M. Kaplan, John T. Maxwell III and Annie Zaenen (eds.), *Formal Issues in Lexical-Functional Grammar*, CSLI Lecture Notes, No. 47, pages 293–309, Stanford, CA: CSLI Publications.
- Heim, Irene and Kratzer, Angelika. 1998. *Semantics in Generative Grammar*. Blackwell Textbooks in Linguistics, Malden, MA: Blackwell.
- Janssen, Theo M. V. 2012. Compositionality: Its Historic Context. In Werning et al. (2012), pages 19–46.
- Link, Godehard. 1998. *Algebraic Semantics in Language and Philosophy*. CSLI Lecture Notes, No. 74, Stanford, CA: CSLI Publications.
- Łojasiewicz, Anna. 1979. O budowie wyrażeń z przyimkiem *po* dystrybutywnym. *Polonica* V, 153–160.
- Mańczak, Witold. 1956. Ile jest rodzajów w polskim? *Język Polski* XXXVI(2), 116–121.
- May, Robert. 1985. *Logical Form: Its Structure and Derivation*, volume 12 of *Linguistic Inquiry Monographs*. Cambridge, MA: The MIT Press.
- Moltmann, Friederike. 1991. On the syntax and semantics of binary distributive quantifiers. In *Proceedings of the North East Linguistic Society (NELS)*, Vol. 22, pages 279–292.
- Moltmann, Friederike. 1997. *Parts and Wholes in Semantics*. Oxford: Oxford University Press.
- Przepiórkowski, Adam. 1999. *Case Assignment and the Complement-Adjunct Dichotomy: A Non-Configurational Constraint-Based Approach*. Ph.D. dissertation, Universität Tübingen, Tübingen.
- Przepiórkowski, Adam. 2006. O dystrybutywnym PO i liczebnikach jedynkowych. *Polonica* XXVI–XXVII, 171–178.
- Przepiórkowski, Adam. 2010. Towards a Construction Grammar Account of the Distributive PO in Polish. *Études Cognitives / Cognitive Studies* 10, 163–176.
- Przepiórkowski, Adam. 2013. The Syntax of Distance Distributivity in Polish: Preserving Generalisations with Weak Heads. In Stefan Müller (ed.), *Proceedings of the HPSG 2013 Conference*, pages 161–181, Stanford, CA: CSLI Publications.
- Przepiórkowski, Adam. 2014a. Distance Distributivity in Polish: Towards a Glue Semantics Approach. In Christopher Piñón (ed.), *Empirical Is-*

- sues in Syntax and Semantics 10 (CSSP 2013 Proceedings)*, pages 107–124.
- Przepiórkowski, Adam. 2014b. Syntactic and Semantic Constraints in a Glue Semantics Approach to Distance Distributivity. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'14 Conference*, pages 482–502, Stanford, CA: CSLI Publications.
- Przepiórkowski, Adam and Patejuk, Agnieszka. 2013. The Syntax of Distance Distributivity in Polish: Weak Heads in LFG via Restriction. In Miriam Butt and Tracy Holloway King (eds.), *The Proceedings of the LFG'13 Conference*, pages 482–502, Stanford, CA: CSLI Publications.
- Przepiórkowski, Adam, Bańko, Mirosław, Górski, Rafał L. and Lewandowska-Tomaszczyk, Barbara (eds.). 2012. *Narodowy Korpus Języka Polskiego*. Warsaw: Wydawnictwo Naukowe PWN.
- Safir, Ken and Stowell, Tim. 1988. Binominal *each*. In *Proceedings of the North East Linguistic Society (NELS), Vol. 18*, pages 426–450.
- Saloni, Zygmunt and Świdziński, Marek. 1998. *Składnia współczesnego języka polskiego*. Warsaw: Wydawnictwo Naukowe PWN, fourth edition.
- Steedman, Mark. 2000. *The Syntactic Process*. Cambridge, MA: The MIT Press.
- Szabó, Zoltán Gendler. 2012. The Case for Compositionality. In Werning et al. (2012), pages 64–80.
- Werning, Markus, Hinzen, Wolfram and Machery, Edouard (eds.). 2012. *The Oxford Handbook of Compositionality*. Oxford: Oxford University Press.
- Zimmermann, Malte. 2002. 'Boys buying two sausages each' – *On the Syntax and Semantics of Distance Distributivity*. Ph. D. dissertation, Universiteit van Amsterdam.
- Zimmermann, Thomas Ede. 2012. Compositionality Problems and How to Solve Them. In Werning et al. (2012), pages 81–106.

## **The Inferential Future in Bulgarian: An Evidential Modal Proposal\***

*María Luisa Rivero*  
University of Ottawa

*Vesela Simeonova*  
University of Ottawa

In Bulgarian, *šte*, morphologically glossed as **FUT**, is used with a prospective reading, (1a) and (1b), and may also signal an inference based on indirect evidence made at Speech Time, (2), like epistemic *will* in *He will (must) be in Toronto right now*. Such a presumptive reading is mentioned in descriptive grammars (Nitsolova 2008, Pašov 2005, Scatton 1983, a.o.), but has not been discussed in the recent formal literature on Bulgarian evidentials which mainly concerns the ‘Renarrative Mood’ (Arregui, Rivero & Salanova 2014, Izvorski 1997, Koev 2011, 2014, Rivero & Slavkov 2014, Sauerland & Schenner 2007, 2013, Smirnova 2013a,b, a.o.). The aim of the present study is to examine the inferential *šte* within the views of formal syntax and semantics.

We argue that *šte*, as a marker of presumptive meaning, is an evidential modal fit for deductions, not reports. It takes a tensed complement with the time of the depicted event as past or present, but not future. Inferential

---

\* Research partially subsidized by the Social Sciences and Humanities Research Council of Canada (Grant 410-2009-0828 to M. L. Rivero). We are very grateful to the reviewers, to Galia Dukova-Zheleva and her family and to Nikolay Slavkov for discussion and judgments. We also thank the audience at FASL 23 and Igor Yanovich. Translations are approximate, with some adopting suggestions by one reviewer. Our claims apply to Bulgarian, not to the modals in the English translations.

*šte* often shares form with prospective *šte*, but we argue in §2 that the two should be formally differentiated. We compare inferential *šte* to epistemic modals in §3, and propose that it lacks a fixed quantificational force. In sum, inferential *šte* is an evidential for inferences that participates in a dedicated morpho-syntactic system not shared by prospective *šte*, so it cannot be viewed as a (purely pragmatic) ‘evidential strategy’ (Aikhenvald 2004), parasitic on the operator at the source of prospective *šte*.

## 1 Introducing Prospective *šte* and Inferential *šte*

Let us introduce Bulgarian affirmative future constructions, which always contain *šte* and are thus periphrastic.<sup>1</sup>

### 1.1 Prospective *šte*

Patterns (1a) and (1b) illustrate future readings we call ‘prospective’ which are forward-shifted with Event Time following Speech Time. They also illustrate that *šte* combines with imperfective and perfective verbs, *piša* and *napiša* in (1b) respectively.

- (1) Context: The instructor in your class asks about your final paper:  
 (1a). You reply with (1b), pointing to the title of an article.  
 a. Gotov li **šte** ti e doklada skoro?  
     Ready Q FUT you<sub>DAT</sub> be<sub>PRES.3SG</sub> paper.the soon  
     ‘Will your paper be ready soon?’

<sup>1</sup> In syntax and morphology, Bulgarian futures differ from East/West Slavic futures. When perfective, West/East Slavic futures bear present morphology, e.g. (ia), and are ungrammatical with the auxiliaries of imperfective futures, e.g. (ic). By contrast, all Bulgarian (affirmative) futures display *šte*, and present perfective verbs are ungrammatical in main clauses, e.g. (i.d).

- (i) a. Naš poezd **ot-pravit-sya** v 10 časov. (Russian)  
     Our train PR-leave<sub>PRES.PRF.3SG</sub> at 10 o'clock  
     b. Našijat vlak **šte za-mine** v 10 časa. (Bulgarian)  
     Our train FUT Pr-leave<sub>PRES.PRF.3SG</sub> at 10 o'clock  
     ‘Our train will leave at 10 o'clock.’  
     c. \*Naš poezd **budet ot-pravit-sya** v 10 časov. (Russian)  
     d. \*Našijat vlak **za-mine** v 10 časa. (Bulgarian)

- b. **Šte** piša, (š**te** piša)  
 FUT write<sub>PRES.IMP.1SG</sub> (FUT write<sub>PRES.IMP.1SG</sub>)  
 i š**te** go napiša.  
 and FUT it<sub>ACC</sub> PR.write<sub>PRES.PF.1SG</sub>  
 ‘I will write and write, and I will finish it.’

### 1.2 Inferential/Presumptive šte

Presumptive **šte** is an evidential modal indicating inferences not reports, as (2) and (3) illustrate.<sup>2</sup>

- (2) **Context:** Your friend asks you which one among 3 singers in a photo is the winner of a competition. You listen to a tape, and pointing to one singer you state:  
 Tazi **šte** (da) e pobeditelkata.  
 This FUT (da) be<sub>PRES.3SG</sub> winner<sub>1SG.FEM.the</sub>  
 ‘This one must be the winner.’
- (3) **Context:** You cannot see Ivan but hear noise next door. You state:  
 Ivan **šte** (da) piše pismo  
 Ivan FUT (da) write<sub>PRES.IMP.3SG</sub> letter  
 v sasednata staja v momenta.  
 in neighbor.the room in moment.the  
 ‘Ivan **must be writing** a letter in the room next door right now.’

Inferential **šte** is felicitous when the evidence is indirect, as in (2) and (3), and infelicitous when direct, as in (4):

- (4) **Context:** You look into the next room, identify the person there as Ivan, and his action as one of writing a letter. You state:  
 #Ivan **šte** (da) piše pismo.  
 Ivan FUT (da) write<sub>PRES.IMP.3SG</sub> letter  
 # ‘Ivan **must be writing** a letter.’

<sup>2</sup> Since Inferential **šte** lacks a reportative reading, it clearly contrasts with the evidential of the Renarrative Mood also known as *preizkazno naklonenie* ‘discourse mood’ (Andrejčin 1977), *énonciation médiatisée* ‘mediated enunciation’ (Guentchéva 1996), Perfect of Evidentiality (Izvorski 1997), *vid na izkazvaneto* ‘discourse aspect’ (Kučarov 1998: 413), and Indirect (Koev 2011, 2014). In (2) and (3), *da* is optional, but it may be obligatory in other contexts, which is a topic beyond the scope of this paper.

We define direct/indirect evidence in terms of propositions (Matthewson 2011, a.o.). It is direct if the event depicted by P in [M [<sub>P</sub> *Ivan write a letter*]] is ‘seen’ as it occurs. Indirect evidence concerns incomplete propositions, for example doubts on the agent’s identity (Ivan or Peter?), the activity (Writing or reading?) or results (A letter or a book?).

Inferential *šte* participates in the two-way orientation of modals. (a) It is anchored to Speech Time, and it signals a present inference when it is in main clauses. (b) But the inference may concern present or past events. With present complement verbs, (2) and (3), inferences are about present events. With present perfect, (5) and (6), or imperfect complement verbs (7), inferences are about the past. In §2, we argue that inferential *šte* does not depict events that extend into the future, which is in contrast with prospective *šte*.

- (5) Context: You wonder why Ivan has never gone to Paris. Since his mom lives there, you suppose that she has often told him to visit.

You state:

Tja **šte** (da) mu e **kazvala**  
 she FUT (da) he<sub>DAT</sub> be<sub>PRES.3SG</sub> tell<sub>PP.IMP</sub>  
 mnogo pāti da ja poseti.  
 many times da she<sub>ACC</sub> visit<sub>PRES.3SG</sub>  
 ‘She **must have told** him to visit her many times.’

- (6) Ivan **šte** (da) e **iztärpjal**  
 Ivan FUT (da) be<sub>PRES.3SG</sub> endure<sub>PP.PRF</sub>

mного през vojната.  
 a.lot during war.the  
 ‘Ivan **must have endured** a lot during the war.’

- (7) Context: You went to a party but have forgotten the name of a guy you met there. You state:

Maj Ivan **šte** da **beše**.  
 Maybe Ivan FUT da be<sub>IMPERF.3SG</sub>  
 ‘It seems like it was Ivan.’

Aspect is encoded in the verbs that complement *šte*. Present perfects with imperfective participles signal ongoing/repetitive events: *kazvala* (5). Perfective participles describe episodic/resultative events: *iztärpjal* (6).

In sum, the evaluation time of a modal claim that contains a main clause inferential *šte* is NOW (in Condoravdi’s (2000) terms, the ‘temporal

perspective' is fixed). The time of the depicted event can either coincide with, or precede, Speech Time (in Condoravdi's (2000) terms, the 'temporal orientation' may vary), but in §2 we see that it cannot be future.

Inferential *šte* always remains invariable, in contrast with future auxiliaries. Prospective *šte* does not overtly encode tense/person/number in (1), but we argue in §2 that it shares the characteristics of the inflected future auxiliary of past future and past future perfects.

In (8), we sketch a (simplified) syntactic structure for inferential *šte*.

(8) [MP [M *šte*] [TP [Tense] [AspectP [Aspect] [VP V]]]]

Based on Rivero (1994), a.o., *šte* heads a Modal Phrase (MP), which dominates both the Tense Phrase (TP) and Person/Number if they are independent of T. TP scopes over Aspect Phrase (AspP) for Viewpoint. Inferential *šte* above T does not inflect for tense/ person/ number.

## 2 Distinguishing between Inferential *šte* and Prospective *šte*

There has been a long debate around forms such as English *will*, which display epistemic and prospective readings. Do they share common semantics disambiguated in context (Lyons 1977, a.o.), or do they represent two temporal/modal operators (Hornstein 1990, a.o.)? Here we argue that in Bulgarian, inferential *šte* must be differentiated from prospective *šte* in syntax and semantics, i.e. that the contrast is grammaticalized.

Bulgarian constructions with prospective and inferential readings may overlap in form, as (9) and (10) illustrate (our glosses and translations).

(9) Kato se sreštnete s nego sled edna sedmitsa,  
 When REFL meet<sub>PRES.2SG</sub> with him after one week,  
 toj **šte** e **razbral** istinata. (Pašov 2005)  
 he FUT be<sub>PRES.3SG</sub> learn<sub>PP.PRF</sub> truth.the  
 'When you meet with him in one week, he will have learned the truth.'



Given the above contrast, we can compare (9) with (13) as a prospective. Likewise, (14) corresponds to (10), with the form of an inferential and an unambiguous epistemic reading.

- (13) Kato se sreštnete s nego sled edna sedmitsa,  
 When REFL meet<sub>PRES.2SG</sub> with him after one week,  
 toj **njama** da e **razbral** istinata.  
 he NEG+FUT da be<sub>PRES.3SG</sub> learn<sub>PP.PRF</sub> truth.the  
 ‘When you meet with him in one week, he will not have learned the truth (at some future time from the time of utterance).’
- (14) Nespokoen e nešto – **ne šte** e  
 Uneasy be<sub>PRES.3SG</sub> something – NEG FUT be<sub>PRES.3SG</sub>  
**razbral** istinata.  
 learn<sub>PP.PRF</sub> truth.the  
 ‘He is somewhat uneasy (at present) – it must be that he has not learned the truth (at some past time before the time of utterance).’

Negation, then, supports the hypothesis that Bulgarian grammaticalizes the contrast between inferentials and prospectives, thus arguing against their unification. The above patterns also show that inferentials specialize in locating the description of events in the past or present. Patterns like (14) lack readings that extend into the future. Constructions that extend into the future such as (13) should thus be viewed as ‘predictive’, not ‘inferential’.

In sum, the grammar of Bulgarian grammaticalizes prospectives and inferentials. Inferential (*ne*) *šte* specializes for epistemic information, with actual/realis-like readings that speak of (possible) present/past events, not future events. Prospective *šte* and *njama* display readings that could be dubbed non-actual/irrealis/predictive, as they speak of events that may extend indefinitely into the future.

## 2.2 Tense, Person, and Number Inflections

In (1), prospective *šte* does not overtly inflect. However, we earlier suggested that this form should be paired with the future auxiliary of past futures and past future perfects, which is inflected in Bulgarian.<sup>3</sup> By

<sup>3</sup> Prospective *šte* was still overtly inflected for person/number in the 19<sup>th</sup> century and could be negated with *ne*, now obsolete but recognizable as literary or poetic. By contrast,

contrast, we mentioned that inferential *šte* is invariable. Let us motivate this proposed second difference between inferentials and prospectives. We illustrate past futures in (15a)-(15c), and past future perfects in (16).

- (15) a. **Štjah** da napiša kniga utre /včera.  
 FUT<sub>IMPERF.1SG</sub> da PR.write<sub>PRES.1SG</sub> book tomorrow/yesterday  
 i. ‘I was going to write a book tomorrow.’  
 ii. ‘I would have written a book yesterday.’
- b. Ivan **šteše** da plati mnogo pari.  
 Ivan FUT<sub>IMPERF.3SG</sub> da pay<sub>PRES.3SG</sub> much money  
 ‘Ivan {was going to pay/would have paid} a lot of money.’
- c. Utre Ivan **šteše** da xodi na gosti  
 tomorrow Ivan FUT<sub>IMPERF.3SG</sub> da go<sub>PRES.3SG</sub> on visit  
 na majka si.  
 at mother his  
 ‘Tomorrow Ivan was going to go on a visit to his  
 mother.’ (adapted from Rivero and Slavkov 2014)
- (16) Do 17 časa včera **štjah**  
 By 17 hour yesterday FUT<sub>IMPERF.1SG</sub>  
 da sām napisala knjigata.  
 da be<sub>1SG</sub> PR.write<sub>PP.PRF</sub> book.the  
 ‘By 5 o’clock yesterday I would have written the book.’

In morphology and syntax, past futures (15a)-(15c) and past future perfect (16) contain a future auxiliary inflected for the imperfect tense, person, and number. The differences are encoded in the complement. Past future complements display present verbs: *xodi* in (15c). Past future perfect complements contain present perfects with an auxiliary and a past participle with aspect: *sām napisala* in (16).

As to the interpretation, past futures and past future perfects display several (complex) meanings, which we do not survey. So-called past futures, for instance, may project into the past or the future in relation to Speech Time, (15a)<sup>4</sup> (or be used for present events, not illustrated).

---

inferential *šte* has always been invariable. Interested readers are referred to Scatton (1983) for a complete inventory and basic descriptions of Bulgarian tenses.

<sup>4</sup> Sentence (15a) illustrates that the Bulgarian past future auxiliary can project into the past without perfect *have* (i.e. a present perfect complement); thus, it differs from English *would*, which can only project into the past with perfect *have*: *Yesterday I would have written the*

Readings in past futures and past future perfects fall within the non-actual/irrealis category (in Condoravdi's terms (2000), 'metaphysical' and not epistemic). That is, (15)-(16) bring to mind (implicit) *if*-clauses and intentions: *I intended to have finished the book by 5 o'clock* for (16).

Negation is the factor that unifies the above inflected future auxiliary with prospective *šte*, and distinguishes it from inferential *šte*. In parallel to (plain) *šte*-prospectives, past futures and past future perfects negate with *njama* which is inflected (imperfect/person/number) (17a)-(17c).

- (17) a. Ivan **njamaše** da plati mnogo pari.  
 Ivan NEG+FUT<sub>IMPERF.3SG</sub> da pay<sub>PRES.3SG</sub> much money  
 'Ivan would not pay a lot of money.'
- b. Utre Ivan **njamaše** da xodi  
 Tomorrow Ivan NEG+FUT<sub>IMPERF.3SG</sub> da go<sub>PRES.3SG</sub>  
 na gosti na majka si.  
 on visit at mother his  
 'Tomorrow Ivan would not/was not going to go on a visit to his mother.'
- c. Do 17 časa včera Ivan  
 By 17 hour yesterday Ivan  
**njamaše** da e napisal knjigata.  
 NEG+FUT<sub>IMPERF.1SG</sub> da be<sub>PRES.1SG</sub> PR.write<sub>PP.PRF</sub> book.the  
 'By 5 o'clock yesterday Ivan would not have written the book.'

In sum, prospective *šte* and the inflected future auxiliary of past futures and past future perfects pattern together. By contrast, inferential (*ne*) *šte* may also depict past events, as in (5)-(7) and (11), but it remains invariable. In conclusion, prospectives inflect while inferentials do not.<sup>5</sup>

---

*book.* We derive this difference from the properties of the Bulgarian present in the *da*-complement in (15a)-(15c). In such syntactic environments, Bulgarian presents are relative tenses with a temporal reference that depends on the main clause, not Speech Time. The embedded presents in (15a)-(15c), then, are not deictic, and may function as 'pasts' when the main clause auxiliary is also understood as a (counterfactual) past, with an interpretive result equivalent to English *would have*.

<sup>5</sup> The syntactic structure (8) for inferentials may not be suitable for Bulgarian prospectives. Due to their inflectional properties, prospectives could be in T or lower, but are unlikely to be in M; in view of §3, they resemble root modals, not epistemics.

2.3 *Conditionals*

Conditionals also support the idea that inflected *šteše* patterns with prospective *šte*, unlike inferential *šte*. First consider contrary-to-fact conditionals with a past perfect in the antecedent clause, and a future auxiliary in the imperfect in the consequent clause, as in (18a) and (18b).

- (18) a. Ako Ivan beše kupil tazi kăšta minalata godina  
 If Ivan be<sub>IMPERF.3SG</sub> bought this house last year  
 toj **šteše** da {e platil/ plati} mnogo pari.  
 he FUT<sub>IMPERF.3SG</sub> da {be<sub>PRES.3SG</sub> paid/ pay<sub>PRES.3SG</sub>} much money  
 ‘If Ivan had bought this house last year (but he did not), he would have paid a lot of money (at that past time).’
- b. Ako Ivan beše kupil tazi kăšta utre  
 If Ivan be<sub>IMPERF.3SG</sub> bought this house tomorrow  
 toj **šteše** da {e platil/ plati} mnogo pari.  
 he FUT<sub>IMPERF.3SG</sub> da {be<sub>PRES.3SG</sub> paid/ pay<sub>PRES.3SG</sub>} much money  
 ‘If Ivan had bought this house tomorrow (but he already bought it), he would have paid a lot of money (at that future time).’

Counterfactuals may project into the past or future. Past (18a) is felicitous if the speaker knows both that Ivan did not buy a house last year when prices were high, and that house prices came down. Future (18b) is felicitous as a comment on what could have happened at some future time if instead of buying the house Ivan purchased, he had waited to buy. Both (18a) and (18b) speak of events that did/will not take place.

A second conditional with a future marker in (19) parallels Greek constructions Iatridou (2000) labels as ‘future less vivid conditionals’ which contemplate future possibilities. The antecedent has an imperfect verb, and the consequent the imperfect auxiliary of counterfactuals.

- (19) Ako Ivan **kupeše** tazi kăšta utre,  
 If Ivan buy<sub>IMPERF.3SG</sub> this house tomorrow,  
 toj **šteše** da plati mnogo pari.  
 he FUT<sub>IMPERF.3SG</sub> da pay<sub>PRES.3SG</sub> much money  
 ‘If Ivan bought/were to buy this house tomorrow (an open possibility), he would pay a lot of money.’

A conditional with *šte* in antecedent and consequent clauses is (20).

- (20) Ako **šte** idvaš utre, az sašto **šte** doida.  
 If FUT go<sub>PRES.2SG</sub> tomorrow, I also FUT go<sub>PRES.1SG</sub>  
 ‘If you (will) go tomorrow, I will also go.’

We do not discuss the readings of the above conditionals, which depend on the form of both the antecedent and the complement of the auxiliary. We concentrate on negation, which formally unifies the three types: they are negated with *njama*. In counterfactuals (21a) and (21b) and the ‘future less vivid conditional’ (21c), the negative auxiliary is inflected. ‘Bare’ *njama* in (21d) is not overtly inflected.

- (21) a. Ako Ivan beše kupil тази কাঁšta minalata godina,  
 If Ivan be<sub>IMPERF.3SG</sub> bought this house last year  
 toj **njamaše** da e platil mnogo pari  
 he NEG+FUT<sub>IMPERF.3SG</sub> da be<sub>PRES.3SG</sub> paid much money  
 ‘If Ivan had bought this house last year, he would NOT have paid a lot of money.’
- b. Ako Ivan beše kupil тази কাঁšta utre,  
 If Ivan be<sub>IMPERF.3SG</sub> bought this house tomorrow  
 toj **njamaše** da e platil /plati  
 he NEG+FUT<sub>IMPERF.3SG</sub> da be<sub>PRES.3SG</sub> paid /pay<sub>PRES.3SG</sub>  
 mnogo pari.  
 much money  
 ‘If Ivan had bought this house tomorrow, he would NOT have paid a lot of money.’
- c. Ako Ivan kupeše тази কাঁšta utre,  
 If Ivan buy<sub>IMPERF.3SG</sub> this house tomorrow,  
 toj **njamaše** da plati mnogo pari.  
 he FUT<sub>IMPERF.3SG</sub> da pay<sub>PRES.3SG</sub> much money  
 ‘If Ivan bought this house tomorrow, he would not pay a lot of money.’
- d. Ako **njama** da idvaš utre,  
 If NEG+FUT da go<sub>PRES.2SG</sub> tomorrow,  
 az **njama** da doida.  
 I NEG+FUT da go<sub>PRES.1SG</sub>  
 ‘If you do not go tomorrow, I will not go.’

Inferential *šte* may also appear in consequent clauses in conditionals and speak of past events, (22a). Hence, it partially resembles classical counterfactuals such as (18a). However, inferential *šte* is negated with *ne*, so (22b) displays the form and meaning of an epistemic.

- (22) a. Ako Ivan e kupil тази кăшта миналата година,  
 If Ivan be<sub>PRES.3SG</sub> bought this house last year,  
 toj **šte** da e platil mnogo pari.  
 he FUT da be<sub>PRES.3SG</sub> paid much money  
 ‘If Ivan bought this house last year, he must have paid a lot of money.’
- b. Ako Ivan e kupil тази кăшта миналата година,  
 If Ivan be<sub>PRES.3SG</sub> bought this house last year ,  
 toj ne **šte** da e platil mnogo pari.  
 he NEG FUT da be<sub>PRES.3SG</sub> paid much money‘  
 ‘If Ivan bought this house last year, he must/will NOT have paid a lot of money.’

In sum, negation formally divides conditionals. Conditionals with inferentials contrast with conditionals with (a) counterfactuals, (b) less vivid futures, and (c) ordinary futures, which all pattern together. Bulgarian distinguishes between epistemic (*ne šte*) and prospective *šte /njama*. Pace Pašov (2005), we conclude that inferentials and prospectives may often overlap in form, but represent two different paradigms. In Bulgarian, then, prospectives and inferentials are grammaticalized, and prospectives specialize for future events.#

### 3 Comparing Inferential *šte* and Epistemic Modals

Bulgarian has two modals with epistemic and root readings: *trjabva* ‘must’ and *može* ‘may, can’. When they overtly inflect for tense (imperfect), person, and number, they are restricted to root readings, (23a)-(23b), but they remain invariable under epistemic readings, (25), etc.

- (23) a. Ivan **trjabvaše** da otide do pazara.  
 Ivan must<sub>IMPERF.3SG</sub> da go<sub>PRES.3SG</sub> to market  
 ‘Ivan had the obligation to go to the market.’

- b. Predi **možeh** da b jagam burzo no veče ne.  
 Before can<sub>IMPF.1SG</sub> da run<sub>PRES.1SG</sub> fast but already no  
 ‘Before I was able to run fast but not anymore.’

We next show that inferential *šte* and epistemic *trjabva* and *može* share four similarities. However, we preliminarily suggest that they also differ: *trjabva* is universal, *može* is existential, and inferential *šte* is a degree expression.

### 3.1 Similarities

Inferential *šte*, epistemic *trjabva* ‘must’, and epistemic *može* ‘may’ are invariable (no tense, person, number inflection). They take parallel complements, (24)-(26). All three embed under parallel propositional attitude verbs. In such contexts, they are anchored to main clauses in ways familiar in the literature on epistemics, (27).

- (24) Ivan **trjabva** /**može** da **piše** pismo.  
 Ivan must/may da write<sub>PRES.IMP.3SG</sub> letter  
 ‘Ivan **must/may be writing** a letter.’  
 (compare with (3): Ivan **šte** (da) **piše** pismo.)
- (25) Az **trjabva** /**može** da **sām** mu **kazvala** mnogo pati.  
 I must/may da be<sub>PRES.1SG</sub> he<sub>DAT</sub> tell<sub>PP.IMP</sub> many times  
 ‘I **must/may have told** him many times.’  
 (see (5): Tja **šte** (da) mu **e kazvala** mnogo pati.)
- (26) Ivan **trjabva** /**može** da **e** **iztārpjal** mnogo prez vojnata.  
 Ivan must/may da be<sub>PRES.3SG</sub> endured a.lot during war.the  
 ‘Ivan **must/may have endured** a lot during the war.’  
 (see (6): Ivan **šte** (da) **e iztārpjal** mnogo prez vojnata.)
- (27) Context: Yesterday, we were watching a crime movie: a woman’s body was discovered. We now discuss the identity of the killer, and you state Mary’s opinion at the time:  
 (Včera) Maria misleše če Ivan **šte** /**trjabva**/**može**  
 Yesterday Mary think<sub>IMPF.3SG</sub> that Ivan FUT/ must /may  
 da ja **e** **ubil**.  
 da she<sub>ACC</sub> be<sub>PRES.3SG</sub> kill<sub>PP.PRF.MASC</sub>  
 ‘(Yesterday) Mary thought that Ivan must/may have killed her.’

Thus, inferential *šte* is an evidential with formal modal properties, not the properties often assigned in the literature to illocutionary markers (see Faller 2002, a.o.).

### 3.2 *A Suggested Difference: Quantificational ‘Flavor’*

Often, inferential *šte* is reminiscent of universal modals including *must*, but there are both declarative and interrogative contexts where it seems closer to *može* ‘may’, as the comparison of (28) and (29) suggests. In our view, inferential *šte* is a variable force modal, one without fixed quantificational force, as we argue next when we identify some of its characteristics (on variable force modals see Deal 2011, Kratzer 2012, Lassiter 2010, Rullmann, Matthewson & Davis 2008, Yalcin 2007, a.o.).

(28) No zašto **šte** (da) gi e ubil (včera)?  
 But why FUT (da) they<sub>ACC</sub> be<sub>PRES.3SG</sub> kill<sub>PP</sub> (yesterday)  
 ‘But why would/should/may he have killed them (yesterday)?’

(29) No zašto **može/ #trjabva** da gi e ubil?  
 But why may/ #must da they<sub>ACC</sub> be<sub>PRES.3SG</sub> kill<sub>PP</sub>  
 ‘But why may/ #must he have killed them?’

To motivate the force variability of *šte*, and its distinction from *trjabva* ‘must’ and *može* ‘may’, we are inspired by Kratzer’s general theory of modality, in particular notions such as ‘at least as good a possibility of’ and ‘better possibility’, which holds when *p* is at least as good a possibility as *q* but not vice versa (Kratzer 2012:41). In our view, inferential *šte* identifies an option that is better than some other option, but not necessarily the best option. Thus, the gradability of *šte* shines through in comparing possibilities, where this modal participates in patterns that are in principle excluded for a modal we consider universal, namely *trjabva* ‘must’, as we show next. To develop our argument, we recall the scenario in (27), adding more than one suspect to the discussion of possible killers. First note contrasts between *može* ‘may’ (30), and *trjabva* ‘must’ (31).

- (30) **Može** da e bil Ivan, ili **može**  
 May da be<sub>PRES.3SG</sub> be<sub>PP</sub> Ivan or may  
 da e bil Boris.  
 da be<sub>PRES.3SG</sub> be<sub>PP</sub> Boris  
 ‘It may have been Ivan, or it may have been Boris.’
- (31) \***Trjabva** da e bil Ivan, ili /no  
 Must da be<sub>PRES.3SG</sub> be<sub>PP</sub> Ivan or /but  
**trjabva** da e bil Boris.  
 must da be<sub>PRES.3SG</sub> be<sub>PP</sub> Boris  
 \*‘It must have been Ivan, or/but it must have been Boris.’

Sentence (30) is fine but (31) is not felicitous because a true necessity modal like *trjabva* ‘must’ needs to report on an option that is better than all other options in all accessible worlds. In other words, in comparing two options *p* and *q*, (31) states that each one of them is the best, i.e. better than **every other option**. Now consider inferential *šte* in comparisons with either *može* ‘may’, (32), or *trjabva* ‘must’, (33). These sentences are both felicitous, and their different readings serve to highlight the flexibility/gradability we attribute to inferential *šte*.

- (32) **Može** da e bil Ivan, ili/no **šte**  
 May da be<sub>PRES.3SG</sub> be<sub>PP</sub> Ivan or/but FUT  
 da e bil Boris.  
 da be<sub>PRES.3SG</sub> be<sub>PP</sub> Boris  
 ‘It could have been Ivan, but it is more likely that it was Boris.’
- (33) **Trjabva** da e bil Ivan, ili / \*no  
 Must da be<sub>PRES.3SG</sub> be<sub>PP</sub> Ivan, or/\*but  
**šte** da e bil Boris.  
 FUT da be<sub>PRES.3SG</sub> be<sub>PP</sub> Boris  
 ‘It must have been Ivan, but it could also have been Boris.’

On the one hand, both Ivan and Boris are possible options in (32), but Boris is the better or more likely option – the suspect with the more dubious alibi, for instance. On the other hand, (33) opposes the best to a ‘better’ or less likely option (a better alibi) without a clash. Crucially, Bulgarian (33), then, differs from (31), which constitutes an attempt to contrast two ‘best’ options. Finally, (34) involves a comparison with two *šte*, and is not felicitous. We suggest that its infelicity derives from setting

up two options that are equal or ‘undefined’ as to which one is to be chosen as better or more likely.

- (34) #**Šte** da e bil Ivan, ili / no  
 FUT da be<sub>PRES.3SG</sub> be<sub>PP</sub> Ivan or/ but  
**šte** da e bil Boris.  
 FUT da be<sub>PRES.3SG</sub> be<sub>PP</sub> Boris  
 #‘It must have been Ivan, or it must have been Boris.’

The comparison with existential *može* in (32), then, increases the ‘strength’ of inferential *šte*, which goes on to identify the better/more likely option (the suspect with a bad alibi). A comparison with the universal modal in (33) weakens *šte*, which then goes on to identify the less preferred/less likely option (the suspect with the better alibi). Both *trjabva* and *može* offer the compositional means to provide appropriate but nevertheless different standards of comparison.

The above situation suggests that a variable force modal is one that can associate with flexible rankings in comparisons – something that fixed universal modals cannot do. A variable force modal, then, need not be equated with the fixed force modals. Therefore, contra the first impression, *šte* is not a universal modal in cases where only one suspect seems to be involved, (35). This sentence is equally felicitous if Ivan is the most likely suspect out of 10 potential suspects, or if there is no other possible suspect.

- (35) Ivan **šte** da e bil.  
 Ivan FUT da be<sub>PRES.3SG</sub> be<sub>PP</sub>  
 ‘It must (degree modal) have been Ivan.’

Inferential *šte* brings to mind expressions with a hidden degree structure such as *tall* (Kennedy & MacNally 2005). We understand sentences such as *Mary is a tall lady* by providing some scale of tallness where Mary is above average without the need of being the tallest (universal). Similarly, we suggest that inferential *šte* in (35) brings to mind a scale of suspects where the chances of Ivan being the killer are better than, say, those of the average possible suspect in a pool of contextually relevant possible suspects. On this view, the universal-like reading of inferential *šte* is a consequence of its comparative properties. To conclude, inferential *šte* is

a **degree** modal without fixed quantificational force, which should not be identified with *trjabva* or with *može*.

#### 4 Conclusions

Our views on inferential *šte* in Bulgarian impinge on long debated issues concerning futures, modals, and evidentials in both general linguistics and Balkan linguistics. We conclude by relating our proposals on *šte* to some of those issues within the framework of recent theoretical views.

We argued in favor of a grammaticalized distinction between inferential *šte* and prospective *šte* in modern Bulgarian. Thus, we joined the long debate on the unity/diversity of futures, opting for a position where inferential and ‘ordinary’ futures are not unified in Bulgarian. This is in contrast with, for instance, some recent views on other languages in the Balkans including Greek (see Giannakidou & Mari 2013) and Rumanian (see Mihoc 2012).

We touched indirectly on the traditional debate about whether ordinary futures are modal or temporal. We concluded that in Bulgarian both inferential and ordinary futures are modal, but must be nevertheless distinguished from one another, which suggests that their modality may not be of the same type. Bulgarian ‘ordinary’ futures formally pattern with counterfactuals and ‘less vivid futures’, and so they are undoubtedly modal, but their agreement characteristics pair them with the types of modals Kratzer dubs circumstantial, i.e. they are not epistemic. We may then ask if the morphological connection with circumstantial modals as opposed to epistemics could also hide a semantic connection.

We have argued that inferential *šte* behaves like a ‘tenseless’ modal anchored to Speech Time, and takes tensed complements. By contrast, prospective *šte* should be paired to past future auxiliaries, which may project into the past ‘on their own’ (i.e. without a present perfect complement). Such an opposition between inferential and prospective markers may shed light on the proper characterization of modals for the present and those for the past, which display crosslinguistic variation (see Condoravdi 2002 on English and the effect of *have*, Giannakidou & Mari 2013 on Greek and Italian, Rivero 2014 on Spanish a.o). The distinctions in Bulgarian may also shed light on the much-debated topic of the relation between counterfactuals and inferentials.

We have shown that inferential *šte* has both evidential and modal properties, and that it cannot be regarded as an illocutionary operator. Thus inferential *šte* may shed additional light on ongoing debates on contrasts between evidentials with modal properties and those with illocutionary properties (Davis, Potts, & Speas 2007, Faller 2002, 2011, von Fintel & Gillies 2010, Matthewson 2011, Matthewson, Davis & Rullmann 2007).

We have sketched out a proposal that evidential *šte* is a degree expression with comparative properties that distinguish it both from traditional universal and existential modals. Thus, we have added it to the inventory of forms that participate in ongoing debate on the proper definition of gradable modals (Deal 2011, Kratzer 2012, Lassiter 2010, Rullmann, Matthewson & Davis 2008, Yalcin 2007, Yanovich 2013).

## References

- Aikhenvald, Alexandra. 2004. *Evidentiality*. Oxford: Oxford U Press.
- Andrejčin, Ljubomir. 1977. Glagol. In Andrejčin, L., Popov, K., & Stojanov, S. (Eds.), *Gramatika na bālgarskija ezik*. Sofia: Nauka i izkustvo, 190-296.
- Arregui, Ana, María Luisa Rivero & Andrés Salanova. 2014. Cross-linguistic variation in imperfectivity. *Natural Language and Linguistic Theory* 32, 307–362.
- Condoravdi, Cleo. 2002. Temporal interpretation of modals: modals for the present and for the past. In Beaver, David et al. (Eds.) *The Construction of Meaning*, Stanford, CA: CSLI Publications, 59–87.
- Davis, Christopher, Christopher Potts & Margaret Speas. 2007. The pragmatic values of evidential sentences. *SALT XVII*, 71-88.
- Deal, Amy R. 2011. Modals without scales. *Language* 87, 559-585.
- Faller, Martina 2002. *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. Ph.D. Dissertation, Stanford.
- Faller, Martina. 2011. A possible worlds semantics for Cuzco Quechua evidentials. *SALT 20*, 660-683.
- von Fintel, Kai & Anthony S. Gillies. 2010. Must... stay.... strong! *Natural Language Semantics* 18, 351-383.
- Giannakidou, Anastasia & Alda Mari, 2013. The future of Greek and Italian: an epistemic analysis. *Sinn und Bedeutung* 17.

- Guentchéva, Zlatka. 1996. *L'énonciation médiatisée*. Paris: Bibliothèque de l'information grammaticale.
- Hornstein, Norbert. 1990. *As Time Goes By: Tense and Universal Grammar*. Cambridge, Mass : MIT Press.
- Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31, 231–270.
- Izvorski, Roumyana. 1997. The present perfect as an epistemic modal. *SALT* 7.
- Kennedy, Chris, & Louise McNally. 2005. Scale structure and semantic typology of gradable predicates. *Language* 81, 345–81.
- Koev, Todor. 2011. Evidentiality and temporal distance learning. *SALT* 21, 95-114.
- Koev, Todor. 2014. *Evidentiality, Learning Events, and Spatiotemporal distance*. University of Stuttgart Ms.
- Kratzer, Angelika. 2012. *Modals and Conditionals: New and Revised Perspectives*. Oxford: Oxford University Press.
- Kučarov, Ivajlo. 1998. Morfologija. In: Bojadžiev, T., Kučarov, I. & Penčev, J. (Eds.), *Savremenena bālgarski ezik*. Sofia: P. Beron, 277-497.
- Lassiter, Daniel. 2010. Gradable epistemic modals, probability, and scale structure. *SALT* 20, 197-215.
- Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.
- Matthewson, Lisa. 2011. On apparently non-modal evidentials. *Empirical Issues in Syntax and Semantics* 8, 333–357.
- Matthewson, Lisa, Henry Davis & Hotze Rullmann. 2007. Evidentials as epistemic modals: evidence from St'át'imcets. *Linguistic Variation Yearbook* 7, 201–254.
- Mihoc, Teodora. 2012. The Romanian presumptive mood: inferential evidentiality and upper-end degree epistemic modality. Unpublished MA Memoire. University of Ottawa.
- Nitsolova, Ruselina. 2008. *Bālgarska Gramatika: Morfologija*. Sofia: St. Kliment Ohridski University Press.
- Pašov, Petăr. 2005. *Bālgarska Gramatika*. Sofia: Hermes.
- Rivero, María Luisa. 1994. Clause structure and V-movement in the languages of the Balkans. *Natural Language and Linguistic Theory* 12, 63-120.
- Rivero, María Luisa. 2014. Spanish inferential and mirative futures and conditionals. *Lingua* 151, 197- 215.

- Rivero, María Luisa & Nikolay Slavkov. 2014. Imperfect(ive) variation: the case of Bulgarian. *Lingua* 150, 232-277.
- Rullmann, Hotze, Lisa Matthewson & Henry Davis. 2008. Modals as distributive indefinites. *Natural Language Semantics* 16, 271-295.
- Sauerland, Uli & Mathias Schenner. 2007. Shifting evidentials in Bulgarian. *Sinn und Bedeutung* 11, 525-539.
- Sauerland, Uli & Mathias Schenner. 2013. On embedding and evidentiality in Bulgarian. *Contrastive Linguistics* XXXVIII, 2-3, 131-152.
- Scatton, Ernest. 1983. *A Reference Grammar of Modern Bulgarian*. Columbus: Slavica.
- Smirnova, Anastasia. 2013a. Evidentiality in Bulgarian. *Journal of Semantics* 30, 479-532.
- Smirnova, Anastasia. 2013b. The meaning of the Bulgarian and Turkish evidentials. *Contrastive Linguistics* XXXVIII, 2-3, 205-223.
- Yalcin, Seth. 2007. Epistemic modals. *Mind* 116.464, 983-1026.
- Yanovich, Igor. 2013. *Four Pieces of Modality, Context and Usage*. Ph.D. Dissertation, MIT, Cambridge, Mass.

mrivero@uottawa.ca  
vsime026@uottawa.ca

**Locus of Causation and *by itself* Phrases:  
A Case Study of Russian *sam po sebe*\***

*Katie Sardinha*  
University of California, Berkeley

The interpretation of *by itself* phrases has been used to argue for different and even competing theories of causal semantics, even within the same language (Chierchia 2004, Koontz-Garboden 2009, Levin and Rappaport-Hovav 1995, Schäfer 2007). Given the centrality of the claims at stake, it is important that we investigate the semantics of *by itself* phrases in particular languages as a prerequisite to relying on them as diagnostics for lexical semantic features. The goal of the present study is to provide a descriptive, empirically-driven generalization about the meaning of the *by itself* phrase in Russian, *sam po sebe*. In particular, I will argue that *sam po sebe* is used to both assert the presence of a cause and to profile a referent as a causal locus.

**1 Introducing *sam po sebe***

English *by itself* phrases are often ambiguous between two readings (Levin and Rappaport-Hovav 1995). This ambiguity is apparent in (1), which can mean either that Masha walked to school unaccompanied – the ‘alone’ reading – or that Masha walked to school unassisted – the ‘without outside

---

\* I would especially like to thank Darya Kavitskaya, Line Mikkelsen, Ryan Bochnak, Denis Paperno, the FASL-23 audience, members of the Fall 2013 Linguistics 137C class at UC Berkeley for their helpful and challenging questions, and my Russian consultants for their insightful comments. In what follows I use # to indicate a high level of semantic anomaly, ?? to indicate a lesser though still significant level of anomaly, ? to indicate slight semantic anomaly, and no marking to indicate semantic felicity.

help' reading. Of these two readings, only the latter is causal in nature.<sup>1</sup> Schäfer (2007) recognized that to account for examples like (2) with inanimate referents, we in fact need to adopt a broader paraphrase for the causal reading of *by itself*, namely 'without outside force'. In (2), only the 'without outside help/force' interpretation is available.

- (1) Masha walked to school (**all**) **by herself**.  
 a. Masha walked alone/unaccompanied: ✓  
 b. Masha walked without outside help or force: ✓
- (2) The alarm turned on (**all**) **by itself**.  
 a. The alarm turned on alone/unaccompanied: ✗  
 b. The alarm turned on without outside help or force: ✓

Unlike the English *by itself* phrase, Russian *sam po sebe* is unambiguous: it can only have the causally-relevant 'without outside help/force' reading. An example sentence containing *sam po sebe* is in (3), and possible and impossible readings of the sentence are summarized in (4).

- (3) Kompjuter vyključaetsja i vključaetsja **sam po sebe**.  
 computer turn.off<sub>3SG.REFL</sub> and turn.on<sub>3SG.REFL</sub> intens<sub>M</sub> prep self<sub>DAT</sub>  
 'The computer turns off and on **all by itself**.'
- (4) a. The computer turns off and on alone/unaccompanied: ✗  
 b. The computer turns off and on without outside help/force: ✓

The phrase *sam po sebe* consists of three lexical items. The first of these is the intensifier *sam* (König, Siemund, and Töpfer 2014), which agrees in gender and/or number with the referent it modifies.<sup>2</sup> The intensifier is followed by the preposition *po*, which has many uses in Russian and no single translation in English, overlapping in distribution with 'by', 'according to', 'along', 'around', 'about', or 'on', depending on context. The preposition *po* assigns dative case to the third lexical item *sebe*, which is a reflexive pronoun. My assumptions about these three meaning

<sup>1</sup> The presence of *all* in (1-2) appears to be significant for the interpretation of English *by itself* phrases, but I leave this issue aside here.

<sup>2</sup> Syntactically, *sam po sebe* phrases appear to only modify structural subjects and not objects, a property common to *by itself* phrases cross-linguistically (Schäfer 2007). See Comrie (1974) and Madariaga (2006) for discussion on the syntactic position of phrases similar to *sam po sebe* in Russian.

components of *sam po sebe* phrases are summarized in Table 1, along with glossing conventions (in square brackets). In what follows, I simply write *sam* to refer to the set including *sam*, *sama*, *samo*, and *sami*.

| LEXICAL ITEM               | GLOSS                                                      |
|----------------------------|------------------------------------------------------------|
| <i>sam</i> (masculine sg.) | ‘self’ intensifier [intens <sub>M</sub> ]                  |
| <i>sama</i> (feminine sg.) | [intens <sub>F</sub> ]                                     |
| <i>samo</i> (neuter sg.)   | [intens <sub>N</sub> ]                                     |
| <i>sami</i> (plural)       | [intens <sub>PL</sub> ]                                    |
| <i>Po</i>                  | ‘by, according to, along, around, about, on,’ [prep]       |
| <i>sebe</i>                | reflexive pronoun in dative case<br>[self <sub>DAT</sub> ] |

Table 1: Assumptions about the meaning components of *sam po sebe*

In this paper I will be concentrating on the interpretation of *sam po sebe* as a phrasal constituent, thereby relegating the task of providing a compositional semantic analysis of the phrase to future research. I will also be treating the intensifier *sam* as an obligatory component of the *by itself* phrase, and will set aside questions of how *sam po sebe* phrases differ from closely related phrases such as *sam soboj*.

The rest of this paper investigates the interpretation of *sam po sebe* phrases with verbs that differ in lexicalized causal properties. I begin in Section 2 by providing an overview of data used in the study; then in Section 3 I present the data and use it to state three empirical generalizations. In Section 4 I offer an analysis where *sam po sebe* is used both to assert the presence of a cause and to profile an argument as the locus of the causal event. I then explain how the analysis can account for my three generalizations, and discuss how and why the analysis differs from a previous proposal made in relation to other languages, the ‘no cause’ analysis (Schäfer 2007). After that, I briefly address the question of what *sam po sebe* modification can tell us about the causative alternation in Russian. Section 5 concludes.

## 2 Overview of the Data

In this section I provide an overview of the data used for this study, and motivate the set of verb classes I selected for investigation below.

### 2.1 Sources of Data

Three principle sources of data were used in this study: 1) a questionnaire; 2) the Russian National Corpus; and 3) internet data.

2.2.1 Questionnaire. The questionnaire consisted of 27 questions, and was completed in Russian by twelve native Russian speakers. These twelve participants included seven women and five men, aged 20 to 60. As there did not appear to be any obvious differences between the responses of speakers currently residing in Russia (six of the total) and speakers residing in the United States (five of the total), I simply pooled the results. The questionnaire consisted of grammatical sentences including the phrase *sam po sebe* along with instructions on how to judge the sentences as *xorošo* ('good'), *tak sebe* ('iffy') or *ploxo* ('bad') based on how meaningful and correct they sounded. Some sentences were also accompanied by explicit contexts, and speakers were asked to judge if the sentence containing *sam po sebe* could describe that context. Following each judgment, speakers were invited to provide comments concerning why they judged the sentence the way they did, how the sentences could be improved, and what additional situations the sentence could be used in: in the end, every question received comments from between three and seven speakers in total.

2.2.2 Online Corpus. The second source of data for this study was the online Russian National Corpus (hereafter RNC) at <http://www.ruscorpora.ru/en/>, accessed during the months of November and December of 2013. The examples cited in this study are taken from the spoken corpus only, and come from a pool of 370 contexts including the phrase *po sebe*, with or without *sam*. I restricted the dataset to the spoken corpus to keep the study a manageable size. As I have no reason to expect *sam po sebe* to be used differently in written versus spoken speech, this choice should not effect the findings.

2.2.3 Yandex.ru Search Engine. Additional examples of spontaneous uses of *sam po sebe* were taken from online forums and message boards accessed through yandex.ru. Examples obtained this way were later checked for grammaticality by native Russian speakers.

## 2.2 Verb Classes Surveyed

Given that interpretations of *by itself* phrases in other languages have been taken to diagnose lexical causal properties of verbs, the study here focused on testing the interpretation of *sam po sebe* in sentences with verbs that have particular relevance to causal semantics. In particular, I selected Russian equivalents of verbs which in English have been argued (Levin and Rappaport-Hovav 1995) to lexicalize externally-caused events (agent transitives), internally-caused events (bodily process verbs, verbs of emission), and acausal events (verbs of appearance, disappearance, and occurrence); as well, I looked at how *sam po sebe* modifies adjectival predicates (which are stative, and therefore acausal) and at causative-alternating verbs (discussed below). These verb classes are summarized in Table 2 alongside Russian examples.

| <i>Category</i>          | <i>Verb Class</i>                                                      | <i>Example</i>                                                                           |
|--------------------------|------------------------------------------------------------------------|------------------------------------------------------------------------------------------|
| <b>EXTERNALLY-CAUSED</b> | AGENT TRANSITIVES                                                      | narezat' 'slice'                                                                         |
| <b>INTERNALLY-CAUSED</b> | BODILY PROCESS VERBS<br>VERBS OF EMISSION                              | krovotočit' 'bleed'<br>taraxtet' 'rattle'                                                |
| <b>ACAUSAL</b>           | VERBS OF APPEARANCE,<br>DISAPPEARANCE, and<br>OCCURRENCE<br>ADJECTIVES | pojavit'sja 'appear'<br>propadat' 'disappear'<br>proisxodit' 'occur'<br>xorošo 'be good' |
| <b>???</b>               | CAUSATIVE<br>ALTERNATING                                               | otkryt'(sja) 'to open'                                                                   |

Table 2: Classification and Examples of Verbs Studied

In what follows, I assume the following definitions for the terms in Table 2, adapted from Levin and Rappaport-Hovav (1995): an *externally-caused* verb lexicalizes an eventuality that is brought about by a property, force, or agent that is construed as existing external to an argument that undergoes a change of state or position; an *internally-caused* verb

lexicalizes an event that is brought about by a property or force inherent to, or located within, an argument that undergoes a change of state or position; and an *acausal* verb lexicalizes a state of being or existence which is unspecified with regards to its causal genesis.

It is important to note that the classification of events in Table 2 as externally caused, internally caused, or acausal was established on the basis of English data and has not been established for Russian at the level of the entire lexicon. In particular, the classification of causative alternating verbs as externally-caused or internally-caused is an area of active cross-linguistic research (Levin and Rappaport-Hovav 1995 Schäfer 2007), including in Russian (Paducheva 2003). An example of the causative alternating verb in Russian *otkryt* 'to open' is illustrated in (6); verbs in this class occur with both transitive and intransitive alternants – the latter with reflexive morphology on the verb – and can occur with either agentive or non-agentive causer subjects.

- (6) a. Vanja / silnyj veter otkryl dver'.  
 Vanya / strong<sub>M.SG</sub> wind<sub>M</sub> open<sub>M.PST</sub> door<sub>F.ACC</sub>  
 '{Vanya / a strong wind} opened the door.'
- b. Dver' otkrylas'.  
 door<sub>F</sub> open<sub>F.PST.REFL</sub>  
 'The door opened.'

Paducheva (2003) provides empirical arguments for Russian alternating verbs being conceptually externally-caused verbs. I will return in 4.3 to the question of whether the interpretation of *sam po sebe* with transitive and intransitive alternants can be used to argue for one alternant being conceptually more 'basic' than the other.

### 3 The Interpretation of *sam po sebe* Phrases

In this section I illustrate how *sam po sebe* is interpreted in sentences containing verbs belonging to the classes identified in Section 2.2. We will see that *sam po sebe* sounds **redundant** and is judged infelicitous modifying an event containing a verb that lexicalizes either an external or internal cause, but sounds **informative**, and is accepted, with verbs lacking a lexically-specified cause. Although our focus will be on the default readings of sentences with *sam po sebe*, we will also see that default

readings of internally-caused verbs can be overridden when an external cause is present in context, making *sam po sebe* felicitous.

### 3.1 Agent Transitives

Agent transitives lexicalize events with an agentive subject – the external cause of the event – and an undergoer direct object. The subject must meet strict requirements of animacy and ability to complete the action denoted by the verb. With agent transitives, speakers consider *sam po sebe* phrases to sound redundant. Example (7) with *narezat* ‘to slice’ was rejected as infelicitous by 12/12 of my Russian consultants:

- (7) # Mixail narezal kartofel' **sam po sebe**.  
 M. slice<sub>3.SG.M</sub> potato intens<sub>M</sub> prep self<sub>DAT</sub>  
 ‘Michael sliced the potatoes **by himself**.’

Consultants’ comments (8) are helpful in articulating how redundancy is at the heart of why sentences like (7) are judged as infelicitous:

- (8) a. “*Kak eščë on mog narezat' kartofel'?*” (“How else could he cut the potatoes?”)  
 b. “*Po sebe - lišnee.*” (“‘po sebe’ is superfluous”.)  
 c. “‘*sam po sebe*’ ne nužen, Mixail i tak vpolne samostojatelen.” (“‘sam po sebe’ isn’t necessary, Michael is totally independent.”)

No consultant was able to volunteer a context where (7), as-is, could be felicitous. Instead, to get the intended reading of ‘without outside help’ with *narezat*, one consultant recommended removing *po sebe* as in (9):

- (9) Deti **sami** narezali kartošku.  
 children intens<sub>PL</sub> slice<sub>PST.PL</sub> potato<sub>ACC</sub>  
 ‘The children **themselves** cut the potatoes.’

The possibility of using the intensifier *sam* alone to get the intended meaning may play a role in blocking speakers’ attempts to come up with a context where sentences like (7) are felicitous. In any case, the reading of redundancy in sentences like (7) is robust.

### 3.2 *Bodily Process Verbs*

By default, bodily processes are conceived of as occurring inside a referent's body and as occurring naturally – that is, without any sort of intervention. This is the basis of their classification as verbs which lexicalize internally-caused events. As with agent transitives, speakers find *sam po sebe* phrases to sound redundant with these verbs. This is illustrated in (10) with *krovotočit* 'to bleed', which was judged 'good' by 4/12 consultants, 'iffy' by 2/12, and 'bad' by 6/12.

- (10)?? Ranka **sama po sebe** krovotočit, zaživat' ne xočet.  
 wound<sub>F</sub> intens<sub>F</sub> prep self<sub>DAT</sub> bleed<sub>3SG</sub> heal<sub>NFIN</sub> neg want<sub>3SG</sub>  
 'The wound is bleeding **all on its own**, it doesn't want to heal.'

Once again, consultants' comments in (11) establish that redundancy plays a significant role in making (10) infelicitous.

- (11) a. ['bad'] "*Vpolne ponjatno bylo by bez oborota 'sama po sebe', no s nim pojavljaetsja verojatnost' togo, čto ranke čto-to ili kto-to možet pomešat' zaživat'.*" ("It would make complete sense without 'sama po sebe', but with it there, it makes it sound likely that something or someone could be interfering with the healing.")  
 b. ['bad'] "*Rana v principe ne možet krovotočit's čej-libo pomošč'ju. Utočnenija takogo roda javljajutsja izlišnimi i ploxo zvučat.*" ("Wounds, in principle, cannot bleed with any kind of help. Refinements like this are unnecessary and sound bad.")

More specifically, the comments in (11) imply that (10) would be felicitous in a non-prototypical context where wounds were understood to somehow require outside forces to cause them to bleed; otherwise *sam po sebe* is redundant. In fact, consultants generally found it possible to use *sam po sebe* to modify events with bodily process verbs whenever the default semantics of the verb (as internally-caused) could be overridden by a context licensing the existence of an external cause. The sentence in (12) with *zasnut* 'to fall asleep' clearly illustrates this possibility; it contains an overt external cause (the singing). This example was judged as 'good' by 6/12, 'iffy' by 5/12, and 'bad' by 1/12, but the comments in (13) show

that peoples' judgments crucially depended on whether or not they accepted the overriding context.

- (12) Obyčno mne nado pet' malčiku, do togo kak  
 usually me<sub>DAT</sub> necessary sing<sub>NFIN</sub> boy<sub>DAT</sub> until dem<sub>GEN</sub> as  
 ...on zasypaet, a sevodnja on zasnul **sam**  
 he sleep<sub>3SG</sub> but today he fall.asleep<sub>M</sub> intens<sub>M</sub>  
 ...**po sebe**.  
 prep self<sub>DAT</sub>  
 'Usually I have to sing to the boy until he falls asleep, but  
 today he fell asleep **all on his own**.'
- (13) a. ['good'] "*Normal'noe opisanie, gde vtoraja situacija zasypanija reběnka protivopostavljaetsja pervoj imenno blagodarja oborotu 'sam po sebe'.*" ("This is an okay description, where the second situation concerning the sleeping child is opposed to the first owing primarily to the use of 'sam po sebe'.")  
 b. ['bad'] "*Značenie 'samostojatel'nosti' peredaetsja s pomoščju 'sam', no ne 'sam po sebe'.*" ("The meaning of 'independence' is given with the help of 'sam', but not 'sam po sebe'.")

Likewise, the sentence in (14) with *česat'sja* 'to itch, scratch' was judged by 5/12 as 'good', 1/12 as 'iffy', and 5/12 'bad'. The polarity of peoples' judgments related to different construals of the event.

- (14) ? Moi ruki češutsja **sami po sebe**.  
 my<sub>PL</sub> hand<sub>PL</sub> itch<sub>3.PL.REFL</sub> intens<sub>PL</sub> prep self<sub>DAT</sub>  
 'My hands are itching **all on their own**.'
- (15) a. ['good'] "*Esli ruki češutsja, to predpolagaetsja čto est' pričina (grjaznye, pocarapannye, i t.d.) – esli češutsja sami po sebe značit est' kontrast meždu ožidaemym i dejstvitel'nym, predloženie obosnovano.*" ("If hands are itching, its assumed that there's some reason for it (they're dirty, scratched, etc.) – and if they itch on their own it means that that there's a contrast between what we expect and what is really happening, so the usage [of *sami po sebe*] is licensed.")  
 b. ['bad'] "*Možno tak skazat', predpologaja, čto u vas net česotki ili allergii.*" ("It's possible to say that, assuming you don't have scabies or allergies.")

The final example in (16), found online, shows *sami po sebe* felicitously occurring with *krasnet'* 'to turn red, blush'. Prior to the occurrence of (16), a mother is discussing how her daughter keeps inexplicably flushing. At first she suspects allergies to be the cause, but later reasons this can't be the case. In (16), she is using *sam po sebe* to express the lack of any apparent external cause for the flushing.

- (16) V tom to i delo, što èto ne svjazano s užinom...  
 as a matter of fact comp this neg connected with dinner<sub>M,INST</sub>  
 ...ščeki krasnejut **sami po sebe.**  
 cheek<sub>PL</sub> turn.red<sub>3,PL</sub> intens<sub>PL</sub> prep self<sub>DAT</sub>  
 'As a matter of fact, this wasn't connected with the dinner [we  
 ate]. [Her] cheeks just turned red **on their own.**'  
 [<http://2006-2009.littleone.ru/archive/index.php/t-940183.html>]

The examples in this section show that *sam po sebe* can be used felicitously with bodily process verbs if a context is first established for the event being externally-caused. Otherwise, modification with *sam po sebe* sounds redundant with the default readings of these verbs.

### 3.3 *Verbs of Emission*

Verbs of emission encode events of sound, light, smell, or substance emission. The subject is by default the internal cause of the emission event, and so these are internally-caused eventualities. We might predict that *sam po sebe* behaves similarly with this class as with bodily process verbs and indeed, this is what we find: *sam po sebe* sounds redundant with these verbs, unless a context is established where an external cause is present. The sentence in (17) with *taraxtet'* 'to rattle' illustrates this interpretive pattern; it was judged 'good' by 3/12 consultants, 'iffy' by 6/12, and 'bad' by 3/12. Consultants' comments in (18) are illuminating.

- (17) ?? Nočju moj xolodil'nik taraxtit **sam po sebe.**  
 night<sub>INST</sub> my<sub>M</sub> fridge rattle<sub>3SG</sub> intens<sub>SM</sub> prep self<sub>DAT</sub>  
 'At night my fridge rattles **all on its own.**'

- (18) a. [‘iffy’] “*Sam po sebe – lišnee.*” (“‘sam po sebe’ is superfluous.”)
- b. [‘iffy’] “*Taraxtet’ – estestvennoe povedenie dlja xolodil’nika, ne trebujščee naružnogo impul’sa...ispol’zovanie budet obosnovano esli budet kontekst objasnjaet čto tvoj xolodil’nik obyčno nikogda ne taraxtit.*” (“Rattling is a natural behaviour for a refrigerator that does not require an external impulse...the use [of this sentence] would be justified in a context where its explained that your refrigerator usually doesn’t rattle.”)
- c. [‘iffy’] “*Neponjatno, čto podrazumevaetsja, ili dnëm xolodil’nik molčit, ili emu pomogajut taraxtet’.*” (“It isn’t clear what is being implied, either during the day the refrigerator is silent, or they are helping the refrigerator rattle.”)

Example (19) shows *sam po sebe* felicitously modifying a verb of light emission, *svetit’sja*. It is felicitous because the speaker first construes chemical glowing as potentially externally-caused.

- (20) U menja jest’ židkij fosfor i on svetitsja. Počemu by me<sub>GEN</sub> is liquid<sub>M.SG</sub> phosphorus and he glow<sub>3.SG.REFL</sub> why ...on imenno nakaplivaja svet svetitsja? Ili on voobščè he exactly accumulating<sub>F.SG</sub> light glow<sub>3.SG.REFL</sub> or he in.general ...**sam po sebe** svetitsja?  
intens<sub>M</sub> prep self<sub>DAT</sub> glow<sub>3.SG.REFL</sub>  
‘I have some liquid phosphorus and its glowing. Why exactly is the light it’s accumulating glowing out? Or does it generally just glow **on its own**? [<http://otvet.mail.ru/question/40944061>]

These examples show that *sam po sebe* can be used felicitously with verbs of emission in contexts construed as having an external cause. Otherwise by default, *sam po sebe* sounds redundant with these verbs.

### 3.4 Verbs of Appearance, Disappearance, and Occurrence

I am assuming that verbs of appearance, disappearance, and occurrence encode states of being which are not lexically-specified as externally or internally caused. Though they are lexically acausal, a cause may be specified in context. Unlike the verbs we have seen above, these verbs

readily accept modification with *sam po sebe* and *sam po sebe* sounds informative. The examples in (21)-(23) were found online.

- (21) Vsë suščee javljaetsja rezul'tatom samorazvitija.  
 everything existing be<sub>3SG.REFL</sub> result<sub>INST</sub> self.development  
 Mir pojavilsja **sam po sebe**, on xoroš i  
 world<sub>M</sub> appear<sub>3SG.REFL</sub> intens<sub>M</sub> prep self<sub>DAT</sub> he good and  
 soveršen, izmenjat' ego ne nado.  
 perfect change<sub>NFIN</sub> him neg necessary  
 'Everything that exists is the result of self-development. The  
 world appeared **all on its own**, its good and perfect, and its not  
 necessary to change it.'  
 [<http://rpp.nashaubeba.ru/docs/index-25004.html>]
- (22) **Sam po sebe** propadaet zvuk vxoda v  
 intens<sub>M</sub> prep self<sub>DAT</sub> disappear<sub>3SG.M</sub> sound<sub>M</sub> entrance<sub>GEN</sub> into  
 sistemu.  
 system<sub>ACC</sub>  
 'All on its own, the system log-in sound disappeared.'  
 [<http://forum.ubuntu.ru/index.php?topic=180026.0>]
- (23) Ničevo ne proisxodit **samo po sebe**. Bez novyx  
 nothing neg occur<sub>3SG.N</sub> intens<sub>N</sub> prep self<sub>DAT</sub> without new<sub>PL.GEN</sub>  
 ljudej žizn' Kompanii zamiraet.  
 people<sub>GEN.PL</sub> life<sub>F</sub> company<sub>GEN</sub> freeze<sub>3SG.F</sub>  
 'Nothing happens **all on its own**. Without new people, the life  
 of a Company freezes.' [<http://www.kapitalsugurta.uz/career/>]

In (21), the 'cause' of the world's appearance is being conceived of as originating from the properties inherent to the world itself, or at least not from factors external to it; in (22), a sound's disappearance is attributed to causal factors within an implicit argument (presumably some part of the computer system) and not outside of it; and in (23), a claim is being denied that things can happen 'all by themselves' – that is, without any external influence. Unlike the lexically-causal verbs we have seen so far, modification with *sam po sebe* is informative and natural with these lexically acausal verbs.

### 3.5 Adjectival Predicates

Adjectival predicates like *xorošij* ‘good’ are lexically acausal. They therefore help us see whether the pattern observed in 3.4, in which lexically acausal verbs allowed informative modification with *sam po sebe* phrases, holds more generally. Example (24) shows this to be the case: *sam po sebe* informatively modifies a sentence involving the adjectival predicate *xorošij* ‘good’. In this example, *sam po sebe* is being used to assert that the mirror possesses inherent properties that enable the state of its ‘goodness’.

- (24) Ono **samo po sebe** zerkalo xorošee. Bolšoe.  
 it<sub>N</sub> intens<sub>N</sub> prep self<sub>DAT</sub> mirror<sub>N.SG</sub> good<sub>N.SG</sub> big<sub>N.SG</sub>  
 ‘It is **on its own** a good mirror. Its big.’  
 [RNC: *Разговор знакомых // Из материалов Саратовского университета, 1988*]

As with acausal verbs in 3.4, *sam po sebe* can be used informatively with adjectival predicates to specify the cause of the state lexicalized by the adjective. Here, the cause is identified in some way with the referent modified, namely the mirror. We will return to discuss the nature of this identification below.

### 3.6 Causative-Alternation Verbs

Recall from (6) above that causative-alternating verbs can occur either transitively or intransitively. Here I show that *sam po sebe* is interpreted differently depending on which alternant is being modified. When *sam po sebe* occurs with the transitive alternant, it sounds redundant and is rejected; but when *sam po sebe* occurs with the intransitive alternant, the resulting sentence is accepted and judged as sounding informative.

Example (25) with the transitive alternant of *razbit* ‘is rejected and judged as sounding redundant.

- (25) #Vladimir razbil čašku **sam po sebe**.  
 Vladimir break<sub>PST.M</sub> cup<sub>ACC</sub> intens<sub>M</sub> prep self<sub>DAT</sub>  
*lit.* ‘Vladimir broke the cup **by himself (without outside help)**.’<sup>3</sup>

<sup>3</sup> This example was judged by two native Russian speakers, and was not part of the original questionnaire.

This judgment puts transitive alternants of causative alternating verbs broadly in the same category as agent transitives, bodily process verbs, and verbs of emission.

On the other hand, examples with *sam po sebe* and intransitive alternants of causative alternating verbs are well-attested. In (26), found online, the speaker is using *sam po sebe* to assert that some glass in his or her house broke without any apparent external cause.

- (26) Samo po sebe razbilos' steklo doma. Čto èto  
 intens<sub>N</sub> prep self<sub>DAT</sub> break<sub>PST.N.REFL</sub> glass at.home what this  
 značit?  
 mean<sub>3SG</sub>  
 'All on its own the glass at home broke. What does this mean?'  
 [Source: <http://otvet.mail.ru/question/44844512>]

In (27), the speaker is using *sam po sebe* to assert that one of the doors of his car opens without any apparent external cause.<sup>4</sup>

- (27) Otkryvaetsja dver' sama po sebe.  
 open<sub>3.SG.REFL</sub> door<sub>F</sub> intens<sub>F</sub> prep self<sub>DAT</sub>  
 'The door opens all on its own.' [<http://kiario4.ru/t655/>]

In (26)-(27), *sam po sebe* is being used to assert that the cause of breaking and the cause of opening are in some way local to the glass and the door, respectively, even while the exact nature of these causes remains mysterious. Significantly, there is no sense of redundancy in examples (26)-(27); this puts intransitive variants of causative alternating verbs in the same class as the lexically causal verbs we have seen.

### 3.7 Generalizations

The following three generalizations arise from the data in 3.1-3.6 above.

Generalization 1. *sam po sebe* is redundant when it modifies events with (i) agent transitives, (ii) bodily process verbs, (iii) verbs of emission, and (iv) transitive alternants of causative alternating verbs.

---

<sup>4</sup> The author of (27) goes on to tell the following story: recently he stopped at a stoplight, and even though his car doors were locked, one of the doors just clicked open. He had to get out of the car to close it. He then mentions that he is still not altogether sure how it could have happened.

Generalization 2. *sam po sebe* is informative when it occurs in sentences with (v) verbs of appearance, disappearance, and occurrence, (vi) adjectival predicates, and (vii) intransitive alternants of causative alternating verbs.

Generalization 3. Bodily process verbs and verbs of emission can felicitously occur with *sam po sebe* phrases in a context where the event is construed (non-prototypically) as externally-caused; *sam po sebe* modification is then informative.

In the next section I will propose an analysis, which explains these generalizations as deriving ultimately from lexical causal semantics.

#### 4 Analysis

In Section 4.1 I present the details of my *Causal Locus Analysis*, and explain how it derives the generalizations stated in 3.7. Then in 4.2 I contrast this analysis with the *No Cause Analysis* proposed for a different set of languages in Schäfer (2007). Finally in 4.3 I discuss how my findings bear upon the question of whether transitive or intransitive alternants of causative alternating verbs are conceptually ‘basic’.

##### 4.1 *Causal Locus Analysis*

According to the *Causal Locus Analysis*, modification with *sam po sebe* involves adding the following two assertions to an event description: 1) *sam po sebe* asserts that the event it is modifying has a cause; and 2) *sam po sebe* identifies a particular referent – namely its antecedent, the referent with which *sam* agrees – as the locus of this cause. By ‘locus of cause’, or ‘causal locus’, I mean simply the location in the world where the causing event occurred. In other words, I am proposing that speakers use *sam po sebe* to profile a particular referent as being the site of a causing event. This analysis is summarized semi-formally in (28):<sup>5</sup>

---

<sup>5</sup> Previous formal denotations of *by itself* phrases have been defined only for particular verb classes – for example, the denotation for Spanish *por sí solo* in Koontz-Garboden (2009) is defined only for change-of-state verbs. There are significant formal challenges in defining a denotation that works for all verb classes: in particular, a mechanism is needed to reliably pick out any event’s highest (subject) argument. I hope the reader will forgive me for leaving this problem unresolved here, and will find (28) sufficient for the exposition at hand.

- (28) *Causal Locus Analysis: sam po sebe* has two meaning components:
- $\lambda e.\exists e'$ [CAUSE( $e'$ ,  $e$ )] (existence of causing event)
  - $\lambda e.\lambda x$ .[CAUSAL.LOCUS( $x$ ,  $e$ )] (identification of causal locus)

Note that in asserting that a causing event exists and is located ‘at’ a particular referent, speakers need not be making a choice about whether the cause is located *internal* to the profiled argument or just *not external* to the profiled argument – the speaker could have either or both of these assertions in mind. This flexibility in interpretation is consistent with the data in (29), where sentence ‘a’ and ‘b’ are both judged to be possible alternative second statements within the discourse. Specifically, this example shows that speakers allow an assertion with *sam po sebe* to be followed up either with a statement that identifies the cause as internal to the profiled referent (the ‘a’ example) or as not external to the profiled referent (the ‘b’ example). These data show, therefore, that we need to allow for both of these possibilities in defining the meaning of *sam po sebe*.

- (29) Moja čaška razbilas’ **sama po sebe!**  
 my<sub>F</sub> cup<sub>F</sub> break<sub>F.PST.REFL</sub> intens<sub>F</sub> prep intens<sub>F</sub>  
 ‘My cup broke **all on its own!**’
- Možet byt’, ona byla ploxogo kačestva.  
 possible be<sub>NFIN</sub> she be<sub>F.PST</sub> bad<sub>F.GEN</sub> quality<sub>F.GEN</sub>  
 ‘Maybe it was bad quality.’
  - Nikogo ne bylo v kuxne ves’ den’, i vetra  
 nobody<sub>GEN</sub> neg be<sub>N.PST</sub> in kitchen<sub>PREP</sub> all day and wind  
 ne bylo.  
 neg be<sub>N.PST</sub>  
 ‘Noone was in the kitchen all day, and there was no wind.’

This flexibility in interpretation is also consistent with the fact that *sam po sebe* is often used in situations where nothing is known about the nature of a particular cause; all that is known is that the locus of the causal event is somehow ‘at’ the site of a particular referent.

How does this analysis explain the empirical generalizations outlined in Section 3.7? Recall that the generalizations related to how modification with *sam po sebe* is either redundant or informative, depending on which class a verb belongs to. Putting aside causative alternating verbs for a moment, I’ll now attempt an explanation.

Generalization 1 can be restated as follows: if a verb is lexically causative – that is, if it is an externally or internally caused verb – modification with *sam po sebe* is redundant. This follows from the first component of the analysis proposed above: since *sam po sebe* asserts the presence of a cause, it is redundant to modify an event using *sam po sebe* if a cause is already lexically present. In asserting the presence of a cause for an event, which already has a cause, modification with *sam po sebe* fails to add new information; thus speakers judge it to be redundant.

Generalization 2 can be restated as follows: if a verb is lexically acausal, modification with *sam po sebe* is informative. This is because *sam po sebe* asserts the presence of a cause which is not lexically present: the fact that a cause exists is always new information.

Generalization 3 is a little trickier. In a context where an internally-caused eventuality is being construed as externally-caused, we might expect *sam po sebe* to sound redundant, since a cause is present at some level of representation. The reason that *sam po sebe* is nevertheless informative in these instances is due to the second part of (28), namely, the identification of the profiled referent as the causal locus. While the existence of a cause is not new information, the locus of the cause is – therefore, *sam po sebe* is informative in these cases, and not redundant.

If it is possible to make *sam po sebe* work in context with internally-caused verbs, why then is it not possible to do the same with an externally-caused verb like *narezat* ‘to slice’? The answer may relate to the fact that while verbs like *narezat* lexicalize externally-caused eventualities, the agentive arguments that saturate these eventualities are themselves *internal* causes. That is to say that agents can be conceived of as having internal properties such as volition, goals, and intentions, which serve as internal causes that enable them to take part in macroevents in which they are external causes. For example, an agentive ‘slicer’ is both an external cause of a slicing event, and the possessor of certain inherent causal properties, which enable her willful participation in events in general. If this explanation is on the right track, then the reason *sam po sebe* sounds redundant with externally caused verbs is that they are already in some sense causal loci by virtue of their agenthood; *sam po sebe*, then, would simply be stating redundant information.

4.2 *The No-Cause Analysis*

Schäfer (2007) uses data from English, German, Greek, and Italian to argue that *by itself* phrases in these languages are used to deny the presence of a cause(r) for an event. While Schäfer’s arguments and analysis may hold up for the languages discussed there, Russian appears to crucially differ from these languages. Consider once again examples like (7) with *narezat’* ‘to slice’. If Russian *sam po sebe* was used to assert that an eventuality had no cause, we might expect (7) to be judged as *contradictory* as opposed to *redundant*, since modification with *sam po sebe* would in that case involve saying that an event, externally caused by Michael, has no cause. Moreover, English sentences such as ‘Michael sliced the potatoes (all) by himself’, unlike in Russian, are felicitous and can receive the ‘without outside help/force’ interpretation. The very fact that these Russian and English sentences differ suggests that Russian *sam po sebe* requires a language-specific analysis.

4.3 *Classification of Causative Alternating Verbs*

We saw above that the transitive alternant of causative alternating verbs patterns with externally-caused verbs, while the intransitive alternant patterns with acausal verbs. Table 3 shows a revised version of Table 2.

| <i>Category</i>          | <i>Verb Class</i>                                  | <i>Example</i>                                                         |
|--------------------------|----------------------------------------------------|------------------------------------------------------------------------|
| <b>EXTERNALLY-CAUSED</b> | AGENT TRANSITIVES                                  | narezat’ ‘slice’                                                       |
|                          | TRANSITIVE ALTERNANTS                              | otkryt’ ‘open <sub>TR</sub> ’                                          |
| <b>INTERNALLY-CAUSED</b> | BODILY PROCESS VERBS                               | krovotočit’ ‘bleed’                                                    |
|                          | VERBS OF EMISSION                                  | taraxtet’ ‘rattle’                                                     |
| <b>ACAUSAL</b>           | VERBS OF APPEARANCE, DISAPPEARANCE, and OCCURRENCE | pojavit’ sja ‘appear’<br>propadat’ ‘disappear’<br>projisxodit’ ‘occur’ |
|                          | ADJECTIVAL PREDICATES                              | xorošo ‘be good’                                                       |
|                          | INTRANS. ALTERNANTS                                | otkryt’ sja ‘open <sub>INTR</sub> ’                                    |
|                          |                                                    |                                                                        |

Table 3: Classification and Examples of Verbs Studied (Revised)

At this point it is worthwhile to consider whether the pattern outlined here with *sam po sebe* can be used to argue for whether one alternant of causative alternating verbs is conceptually more ‘basic’ than the other.

Paducheva (2003) argues that in Russia the transitive alternant is conceptually basic and that the intransitive alternant is derived via ‘adjunct

causer deletion', a rule which can apply to delete unspecified and therefore irrelevant causers. This view can be made consistent with the classification in Table 3 by assuming that *sam po sebe* modification applies to a representation which has already undergone this deletion.

Nevertheless, the opposite view – that the intransitive alternant is conceptually basic – is also consistent with the pattern in Table 3 on a different set of theoretical assumptions. For instance, this could be the case under the assumption that *sam po sebe* modifies the transitive variant only after it has been derived via causativization.

Therefore, I think it's important to note that the pattern in Table 3 is potentially consistent with competing, and in this case mutually inconsistent, proposals regarding the causative alternation, depending on which additional theoretical assumptions one chooses to adopt. Thus while *by itself* phrases are clearly relevant to the study of lexicalized causal properties, their ability to serve as a simple diagnostic for one causative alternant being more 'basic' than the other is not a given.

## 5 Conclusion

In this paper I have proposed a descriptive, empirically-driven analysis of the meaning of *sam po sebe* phrases in Russian, wherein *sam po sebe* is used both to assert the presence of a cause for an event, and to identify a particular referent as the causal locus. We have seen that the interpretation of *sam po sebe* with verbs from different verb classes is consistent with there being a distinction between externally-caused eventualities, internally-caused eventualities, and acausal eventualities in this language, and that it is sometimes possible to override these default semantics in context. Having investigated the meaning of *sam po sebe*, we are now in a better position to assess what this phrase can tell us about causal semantics in Russian, as well as what it cannot.

## References

- Chierchia, Gennaro. 2004. A semantics for unaccusatives and its syntactic consequences. In *The Unaccusativity Puzzle: Explorations of the Syntax-Lexicon Interface*, eds. Artemis Alexiadou, Elena

- Anagnostopoulou, and Martin Everaert. 22–59. Oxford: Oxford University Press.
- Comrie, Bernard. 1974. The second dative: a transformational approach. In *Slavic Transformational Syntax*, eds. Richard D. Brecht and Catherine V. Chvany. 123–150. Ann Arbor, MI: Michigan Slavic Publications.
- Koontz-Garboden, Andrew. 2009. Anticausativization. *Natural Language and Linguistic Theory* 27:77–138.
- König, Ekkehard, Peter Siemund, and Stephan Töpper. 2013. Intensifiers and reflexive pronouns. In *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Accessed online 2014-09-12.
- Levin, Beth & Malka Rappaport-Hovav. 1995. *Unaccusativity: at the Syntax-Lexical Semantics Interface*. Cambridge: MIT Press.
- Madariaga, Nerea. 2006. Why Russian semi-predicative items always agree. *Journal of Slavic Linguistics* 14:45–78.
- Paducheva, Elena. 2003. Is there an “anticausative” component in the semantics of decausatives? *Journal of Slavic Linguistics* 11:173–198.
- Schäfer, Florian. 2007. ‘By itself’. Ms., Universität Stuttgart, Germany.

katie.sardinha@berkeley.edu

## **On Clitics, their Place in the Prosodic Structure, and Accent\***

*Aida Talić*  
University of Connecticut

This paper investigates syntax-prosody interaction in the context of Bosnian/Croatian/Serbian (BCS) clitics. I show that clitic mapping from the syntactic to the prosodic structure depends on the syntactic complexity of the host at the output of the syntax. This effect is visible in accent shifts from hosts to proclitics in some dialects of BCS in (1).

- (1) **ú**\_kući  
in\_house

Not all BCS dialects allow the shift in (1). Selkirk (1996) in fact suggests BCS dialects differ in the way their clitics are mapped in the prosody: in dialects that allow the shift, clitics map as either *internal* clitics (incorporated into the prosodic word of the host) (2a) or *affixal* clitics (adjoined to the prosodic word of the host) (2b). In dialects that disallow it, clitics are *free* clitics outside the prosodic word of the host (2c).

---

\* I am grateful to Harry van der Hulst, Željko Bošković, Jonathan Bobaljik, Andrea Calabrese, Nadira Aljović, two anonymous reviewers, and the audience of FASL 23 for valuable discussion and insightful comments.

- (2) a. (clitic lexical word)<sub>ω</sub> → internal clitic  
 b. (clitic (lexical word)<sub>ω</sub>)<sub>ω</sub> → affixal clitic  
 c. (clitic (lexical word)<sub>ω</sub>)<sub>φ</sub> → free clitic<sup>1</sup>

I show that in dialects with the accent shift,<sup>2</sup> we need to allow for all options in (2) even within the same dialect, and that the choice of clitic mapping in these dialects depends on the level of morphosyntactic complexity of the host. I investigate the relevant accent shift with nouns, adjectives, and verbs hosting proclitics in different constructions, as well as certain contexts where enclitics interact with the accent of their host.

### 1 Accent Shift to Proclitics and Prefixes

BCS is a pitch-accent language where accented syllables carry a falling or a rising tone.<sup>3</sup> A falling accent on the first syllable is a result of either the default initial High tone insertion (in words lacking lexical High tones) or a word-initial lexical High tone, while a rising tone results from High tone spreading (see e.g. Inkelas & Zec 1988). As described by Riđanović & Aljović (2009), proclitics in certain BCS dialects can take over a falling tone from the first syllable of the host and the resulting tone on the proclitic is either falling or rising, as illustrated with two types of hosts in (3)-(4).

- (3) **Typel: Toneless roots** → *Falling tone on the PCL*<sup>4</sup>  
 a. zà\_tobɔ:m ‘after you’ (PRN) (cf. tòbom)  
 b. zà\_ra:d ‘for the work/article’ (N) (cf. rà:d)  
 c. nè\_igra:m ‘I am not playing’ (V) (cf. ìgra:m)

<sup>1</sup> Notation and abbreviations used: ω=prosodic word, φ=phonological phrase, σ=syllable, √=root, [ ` ]=falling tone, [ ´ ]=rising tone, [ : ]=long vowel; PCL/ECL=proclitic/enclitic; PFX/SFX=prefix/suffix.

<sup>2</sup> Previous literature reports dialects in south Bosnia and Herzegovina (=Herzegovina) and Montenegro as allowing the shift. Consultants of accent shifting dialects are from central, northeast, and south Bosnia and Herzegovina.

<sup>3</sup> The notion “accent” in this paper means “prominence” rather than “accent mark on the metrical grid”, and “accent shift” refers to the shift of prominence.

<sup>4</sup> The shift to PRNs and Ns (3a-b) & (4a-b) is limited to what I refer to as “shifting dialects” (see fn. 2), while shifting to Vs (4c) occurs in a larger area. I will show below why the negation has a closer bond to the verb than proclitics preceding PRNs, Ns, and As. The shift in (3c) seems to be limited to a few verbs (*nè\_idem* – ‘neg\_go’; *nè\_odem* –

- (4) **Type2: Toneful roots (=with initial H) → Rising tone on the PCL**
- a. zá\_tebe ‘for you’ (PRN) (cf. *tèbe*)
  - b. zá\_čovjeka ‘for the man’ (N) (cf. *čòvjeka*)
  - c. né\_pi:še:m ‘I am not writing’ (V) (cf. *pi:še:m*)

Selkirk (1996) proposes that in BCS dialects that allow such accent shift, clitics are either internal or affixal clitics, within the prosodic word of the host.<sup>5</sup> As such, they are in the domain of accent assignment rules, and can interact with the accent of the host.

Additional support for treating proclitics in (3)-(4) as internal clitics incorporated into the prosodic word of their host comes from the fact that they behave identical to prefixes regarding the accent shift: A prefix preceding a toneless host has a falling tone (5a) and a prefix preceding a host with a High tone has a rising tone (5b). Compare (5a) and (5b) to (3b) and (4b) respectively.

- (5) a. nèra:d (neg+work) ‘idleness’  
 b. néčovjek (neg+man) ‘brute’<sup>6</sup>

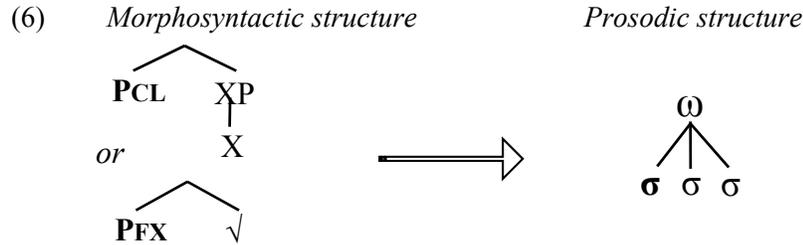
The parallelism between (3b)/(5a) and (4b)/(5a) indicates that, just like prefixes, prepositions (proclitics) enter, into the domain of accent assignment in BCS. We can capture this by assuming that both proclitics and prefixes *incorporate* into the prosodic word ( $\omega$ ) of the host in (3)-(5), as in (6).

---

‘neg\_leave’); i.e. most BCS verbs seem to have (or receive) a High tone before they combine with the clitic as in (4c).

<sup>5</sup> Selkirk discusses simple hosts without lexical High tones. I will show below that while her account can be extended to simple hosts with an initial High tone, it is necessary to modify it to capture contexts with more complex hosts.

<sup>6</sup> Prefixes other than *ne-* behave the same: e.g. *práčovjek* ‘Early Man’.



Given the mapping in (6), in (3a-c) the proclitic is the first syllable in the prosodic word with no lexical High tone and it surfaces with a default falling tone, but in (4a-c) the proclitic precedes a host with a lexical initial High tone, which spreads to the proclitic and is realized as rising. Roots with a lexical High tone are more frequent and the accent shift in most cases results in a rising tone on the proclitic.

In the following two sections, I show different levels of morphosyntactic complexity of the host have different effects on the shift. Ridanović & Aljović (2009) describe some examples of phrase-internal complexity blocking the shift, but the effect of word-internal complexity has not been reported in the literature. Also, to the best of my knowledge, there is no account of these two effects in the existing literature.

### 1.1 *Word-internally Branching Hosts*

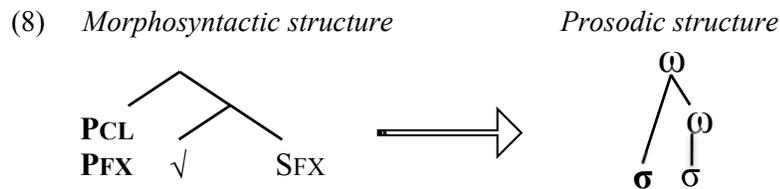
We have seen that a proclitic/prefix preceding a toneless host incorporates into its prosodic word and gets a falling tone by the default H-insertion in (3). In contrast, when a proclitic/prefix precedes a toneless root followed by a derivational suffix, the tone on the proclitic/prefix must be rising.<sup>7</sup>

- (7) a. **ra:dni:ke**      b. **prédra:dni:ke**      c. **zá\_ra:dni:ke**  
      workers.acc      foremen.acc      for\_workers

Given that the root *ra:d* ‘work’ is toneless, the High tone that spreads to the prefix/proclitic can only be a result of initial H-insertion. If the prefix/proclitic were in the domain of this rule here (as in (6)), the High

<sup>7</sup> Given that *-ni:k* is a suffix that creates agent nominals, *ra:dni:k* ‘worker’ could have even more internal structure. I am focusing on overt pieces for ease of exposition, but the point remains the same – word internal complexity of the host has an effect on what kind of tone the clitic gets.

tone would be inserted to the prefix/proclitic and realized as falling. The rising tone on the prefix/proclitic in (7) indicates that they are not in the same domain as the host for the purposes of H-insertion, but they *are* in the domain of the host for H-spreading. That is, prefixes/proclitics behave as if they are both inside and outside the prosodic word of the host, which is exactly what Selkirk (1996) proposed for affixal clitics. Thus, I assume that prefixes/proclitics preceding a morphologically complex host as in (7) map as in (8).



The effect in (7) shows that there are two prosodic-word-internal domains of accentual rule application:<sup>8</sup>

- (9) a. The inner (minimal) prosodic word = root + suffixes  
 b. The outer (maximal) prosodic word that contains the clitic

Therefore, in (7) initial H-insertion applies within the minimal prosodic word, which contains only the host. H-spreading then applies within the maximal prosodic word, which contains the host and the clitic, so the proclitic gets a rising tone.

In the following section I show what effect phrase-internal branching of the host has on the shift.

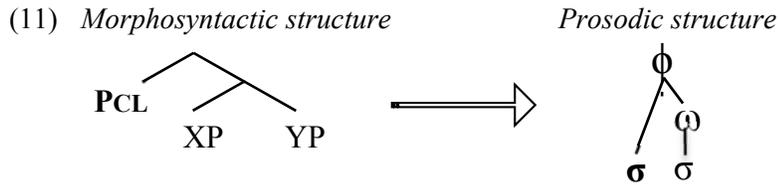
### 1.2 *Phrase-internally Branching Hosts*

A proclitic preceding an NP with a PP-adjunct or a complement cannot interact with the accent of the noun. (10a) shows that the proclitic is not in the domain of initial H-insertion, and (10b) shows that it is also not in the domain of H-spreading.

<sup>8</sup> See also Blumenfeld (2014) for domains of phonological processes in Russian.

- (10) a. \* $\dot{u}$  [ra:d [o klitikama]]  
 b. \* $\acute{u}$  [ra:d [o klitikama]]  
 c. u [rà:d [o klitikama]]  
 in article about clitics

Recall that for shifting dialects, Selkirk (1996) suggests that clitics map as either internal or affixal clitics. Under such an analysis, a proclitic would interact with the accent of the host in a dialect that in principle allows the shift, contrary to what we see in (10). The proclitic here behaves as if it were outside of the prosodic word of the host completely. Crucially, Selkirk (1996) argues that a clitic can attach as *a sister to the prosodic word*, creating a phonological phrase ( $\phi$ ) with it (2c). With this, she captures dialects that disallow the accent shift to proclitics. I take (10) to suggest that even in the dialects that allow the shift, the clitic preceding a syntactically branching host is completely outside of the prosodic word as shown in (11). As the clitic is both outside  $\omega_{min}$  and  $\omega_{max}$ , it is out of reach of either H-insertion and H-spreading.



In sum, BCS proclitics map to prosody in each of the three ways proposed by Selkirk (1996) even within the same dialect that allows accent shift to proclitics. Their mapping depends on the morphosyntactic complexity of the host, and has consequences on the interaction of proclitics with the accent of the host.

- |                                                                                                                                                                                                                           |                                                                                                                                                                                                                                                                         |                                                                                                                                                                                                                                                                    |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>(12) a. <b>internal clitic</b><br/>         (clitic host)<math>_{\omega}</math><br/>         ✓ falling<br/>         ✓ rising<br/>         ✓ H-insertion<br/>         ✓ H-spreading<br/>         non-branching host</p> | <p>b. <b>affixal clitic</b><br/>         (clitic (host)<math>_{\omega}</math>)<math>_{\omega}</math><br/>         *falling<br/>         ✓ rising<br/>         *H-insertion<br/>         ✓ H-spreading<br/>         word-internal<br/>         branching of the host</p> | <p>c. <b>free clitic</b><br/>         (clitic (host)<math>_{\omega}</math>)<math>_{\phi}</math><br/>         *falling<br/>         *rising<br/>         *H-insertion<br/>         *H-spreading<br/>         phrase-internal<br/>         branching of the host</p> |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

We have reached (12) based on nominal hosts; below I turn to proclitics preceding adjectival and verbal hosts.

## 2 Adjectival Hosts

### 2.1 Simple and Branching APs

Proclitics preceding adjectival hosts reveal two types of adjectives. A proclitic preceding the first type of adjectives illustrated in (13a-b) can optionally get either a falling or a rising tone without any change in meaning, which indicates that there is a free variation between adjectival roots that are lexically marked or unmarked for a High tone.

- (13) **Type1A: Toneful/Toneless** → *Falling/rising tone on the PCL*
- a. **ú**\_ova:j/koji/svaki zi:d ‘in(to) this/which/every wall’- lexical H
  - b. **ù**\_ova:j/koji/svaki zi:d ‘in(to) this/which/every wall’  
- no lexical H

A proclitic preceding adjectives of the second type can get only a rising tone, which indicates that such adjectives are lexically marked for a High tone, without a toneless variant.

- (14) **Type2A: Toneful roots** → *Rising tone on the PCL*
- a. **ú**\_našu/Selminu kuću ‘in(to) our/Selma’s house’<sup>9</sup> - lexical H  
\***ù**\_našu/Selma’s kuću
  - b. **ú**\_ta:j zi:d ‘in(to) that wall’ - lexical H  
\***ù**\_ta:j zi:d
  - c. **ú**\_veliku kuću ‘in(to) the big house’ - lexical H  
\***ù**\_veliku kuću

The availability of accent shift here suggests that proclitics enter the minimal prosodic word of adjectives in (13)-(14), as in (6) above.

Similar to syntactic branching with NP hosts, if there is additional branching within the AP, i.e. if an intensifier precedes the adjective, accent shift from the first syllable of the intensifier is degraded.

---

<sup>9</sup> Note that demonstratives, possessives and agreeing quantifiers are morphologically adjectives in BCS, and they occupy the same syntactic position as adjectives (Zlatic 1997; Bošković 2005; Despić 2011).

- (15) a. \*ù\_ja:ko veliku sobu  
 b. \*ú\_ja:ko veliku sobu  
 c. u jà:ko veliku sobu  
    in very big room

This indicates that the clitic can neither incorporate nor adjoin to the adverb within the AP. Thus, the proclitic in (15) maps as in (11) above.

We have seen so far that a proclitic preceding a syntactically branching NP or AP maps as a free clitic. This raises the question of whether a proclitic preceding an adjective precedes a non-branching element in the output of the syntax. Furthermore, given that the accent shift is allowed from simple adjectives to proclitics, we know they constitute a prosodic word. The question is whether the incorporation of the proclitic into its adjectival host takes place in the syntax or only in the prosody.

In its base position P precedes a branching NP (P+[NP AP [NP ]]): if the mapping to the prosodic structure summarized in (12) is right, then that P has to cliticize to the adjective in the syntax to enter into its prosodic word in PF. Independent evidence that this is indeed the case comes from extractions that seem to be moving non-constituents, referred to “extraordinary left-branch extraction (LBE)” (see Bošković 2005). BCS allows P+AP to move out of a PP:

- (16) [Ú\_sta:roj]<sub>i</sub> je živjela t<sub>i</sub> kući.  
       in\_old is lived house

Borsley & Jaworska (1988) and Bošković (2013) argue that such cases involve ordinary LBE where P adjoins to the moving adjective.<sup>10</sup> Given that P+AP can undergo syntactic movement, it cannot be the case that P incorporates into the adjective only in the prosody. Bošković (2013) offers two alternative analyses for this approach to extraordinary LBE: downward vs. upward P-cliticization. In Talić (2013, 2014), I argue for the latter analysis based on a correlation between accent shift and adjective extraction. I turn to this effect in the following section.

<sup>10</sup> See Bošković (2005) for evidence that extraordinary LBE involves extraction of the AP rather than either remnant PP-fronting (Franks & Progovac 1994; Abels 2003) or scattered-deletion (Čavar & Fanselow 2000). See Stjepanović (2010; 2011) for evidence that LBE in general involves extraction of the AP out of the NP.

### 2.2 *The Effect of Syntactic Movement on Accent Shift to Prepositions*

We have seen that accent shifts from BCS adjectives to prepositions in contexts where one adjective modifies a noun as shown in (13) and (14). Interestingly, when two adjectives modify a noun, the shift is degraded if both adjectives are descriptive (17a), but it is allowed if adjectives belong to different classes (17b-d). Note that BCS possessives, demonstratives and some quantifiers are morphologically and syntactically adjectives (Zlatić 1997; Bošković 2005; Despić 2011).

- (17) a. \*?ú\_sta:roj velikoj kući                    *descriptive + descriptive*  
           in\_old big house  
       b. %ú\_sta:roj bratovoj kući                    *descriptive + possessive*  
           in\_old brother.poss house  
       c. ú\_ovoj bratovoj kući                    *demonstrative + possessive*  
           in\_this brother.poss house  
       d. ú\_svakoj sta:roj kući                    *quantifier + descriptive*  
           in\_every old house

Interestingly, these contexts where accent shift is allowed and where it is blocked exactly replicate contexts of another phenomenon in BCS. This language allows left-branch extraction of adjectives (LBE) (18a). With two descriptive adjectives modifying a noun, the shift is degraded (18b), but it improves if the adjectives belong to different classes (18c-d), as noticed by Bošković (2005).

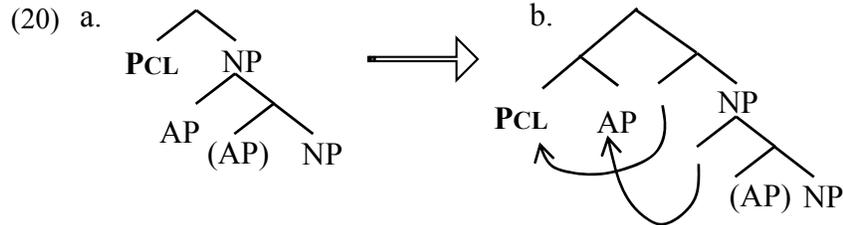
- (18) a. Sta:ru<sub>i</sub> je voljela t<sub>i</sub> kuću.                    cf. (14c)  
           old is loved house  
       b. \*Sta:ru<sub>i</sub> je voljela t<sub>i</sub> veliku kuću.                    cf. (17a)  
           old is loved big house  
       c. %Sta:ru<sub>i</sub> je voljela t<sub>i</sub> bratovu kuću.<sup>11</sup>                    cf. (17b)  
           old is loved brother.poss house  
       d. Ovu<sub>i</sub>/Svaku<sub>i</sub> je voljela t<sub>i</sub> bratovu kuću.                    cf. (17c,d)  
           this/every is loved brother.poss house

<sup>11</sup> Interestingly, one speaker who rejects (17b), also does not allow LBE in (18c), which makes these examples parallel to (17a) and (18b) for this speaker.

Based on this striking parallelism between accent shift in (17) and LBE in (18), which shows that accent shift is possible only in contexts where it is possible to move the adjective, I reach the following generalization:

- (19) A proclitic (preposition) can take over the accent from its host only if the host is allowed to move independently.

What this implies is that the preposition does not cliticize to the adjective in a downward fashion. Rather, the adjective first moves to a position c-commanding the proclitic, and then the clitic adjoins to it (see Talić (2013, 2014) for more details).



Therefore, the proclitic is a sister to a branching NP in situ, but after it cliticizes to the AP, it reaches PF as adjoined to a non-branching AP (unless the AP itself branches). Thus, P+AP can map as in (6).<sup>12</sup>

Additional evidence that the proclitic attaches to the AP in the syntax after the AP moves out of the NP comes from the fact that in NPs with adjectives, additional complexity introduced by adding PP adjuncts does not change the facts about the shift, i.e. the shift is still possible.

- (21) a.  $\acute{u}$ \_naše:m/to:m stanu na Jaliji  
 in\_our/that apartment at Jalija  
 b.  $\acute{u}$ \_Selmino:j novo:j kući pored rijeke  
 in Selma's new house by river  
 c.  $\acute{u}$ \_sta:ro:j bratovo:j kući pored rijeke  
 in\_old brother's house by river

<sup>12</sup> Note that under a lowering analysis, it would be impossible to capture why accent shift is degraded in cases where the adjective cannot move, since for a lowering P, all the examples in (13), (14), and (17) look the same.

Further support for upward cliticization in cases where the first branch can move out of a branching NP comes from other cases where both the movement of the first branch and the accent shift are blocked, which was the case with NPs containing PP-adjuncts. The example in (10) can be understood in the same way as (17). Only full phrases and heads can move. Moving only a segment of a phrase is not possible (22b).

- (22) a. Ovaj primjer mi treba [za [rà:d [o klitikama]]].  
 this example me.dat needs for article about clitics  
 ‘I need this example for an article about clitics.’  
 b. \* [Za rad] mi ovaj primjer treba [ t o klitikama].

The branching structure of the NP containing a PP in (22) cannot be split by moving the first branch above the proclitic, hence the proclitic in such constructions always reaches PF preceding a branching NP and can only map to prosody as in (11).

### 3 Verbal Hosts and the Negation

BCS verbs have roots with and without lexical High tones just like nouns and adjectives, and prefixes/proclitics can take over the accent of verbal hosts as well. As before, a prefix/proclitic preceding a toneless verbal host gets a falling tone, but a prefix/proclitic preceding a verbal host with a lexical tone gets a rising tone.

- (23) a. šè:ta:m ‘walk.1sg’                      d. igra:m ‘play.1sg’  
 b. próše:ta:m ‘take.walk.1sg’            e. zàigra:m ‘start.play.1sg’  
 c. né\_še:ta:m ‘neg\_walk.1sg’            f. nè\_igra:m ‘neg\_play.1sg’  
     -lexical H                                      -no lexical H  
     -rising on PCL                                -falling on PCL

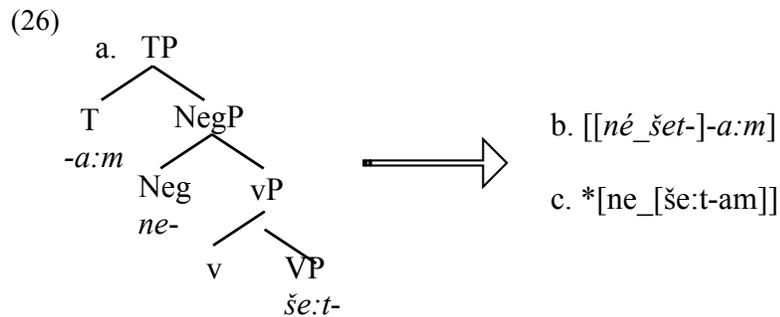
However, there are several differences that separate verbal hosts from nominal and adjectival hosts. First, the accent shift to the sentential negation, which is traditionally assumed to be a clitic, is not optional, and is widely present cross-dialectally. Unlike with branching NPs and APs (see (10) and (15) above), a (PP) adjunct or a complement following a verb does not block accent shift to negation.

- (24) a. Nikad **né\_štam** pored rijeke  
 never neg\_walk by river  
 ‘I never walk by the river.’  
 b. Nikad **nè\_igraju** fudbal pored rijeke.  
 never neg\_play soccer by river  
 ‘They never play soccer by the river.’

Furthermore, even if a VP-internal adverb separates the negation from its host at the beginning of the derivation, the negative particle is still able to incorporate into it, and the shift is possible.

- (25) a. [ on **né\_gleda** često ~~gleda~~ filmove]  
 he neg\_watches often movies  
 ‘He doesn’t often watch movies.’  
 b. \*on ne često gleda filmove.

I assume that the verb in BCS undergoes head movement to a syntactic position higher than V (Bošković 1995, 2001; Stjepanović 1999). If the negation is present, the verb carries it in further V-head movement steps.



Moreover, notice that the verbs in (23)-(25) have tense morphology. At first glance, branching at the word level does not seem to block the shift to the prefix or to the negation, in contrast to the effect that word internal branching has on the shift from nominal hosts (7). However, what is different about the two cases of word internal complexity is that in (7), the root is combined with the suffix before the clitic enters the derivation, which is not necessarily the case with the verbal host and the negation. In particular, if we take the negation to project its phrase between the VP and

the TP (26a) (Pollock 1989; Bobaljik 1995), then the verb combines with the negation before tense morphology is added (26b), not after (26c).

The negative element *ne* is traditionally assumed to be a clitic, so the structure in (26) raises the question of how it can attach to the verb before the affix. However, the traditional treatment of the negation as a clitic may be merely based on the fact that it is orthographically separated from the verb; it is not well motivated by its syntactic and phonological behavior. First, unlike other items taken to be clitics in BCS (prepositions, particles *ni* 'neither' and *i* 'also'), the negation is very picky with respect to what category its host can belong to, which is the defining characteristic of affixes, not clitics (cf. Zwicky & Pullum's 1983 criteria separating clitics from affixes). It always has to be followed by the finite verb wherever it is in the sentence, and it cannot be separated from it by another element (25b). Second, the fact that the accent shift from the verb to the negation obligatorily takes place in a wide array of BCS dialects, even where accent shift to other clitics is disallowed, also indicates that the negation is much more closely bound to its verbal host than other clitics are to their hosts. I take this to mean that BCS negation is morphologically a prefix rather than a proclitic. It follows from the structure in (26a) that the negation always precedes a simple host in the output of the syntax. Thus, based on the mapping mechanism developed above (6), the negation is internal to the prosodic word of the host, hence can interact both with H-insertion and H-spreading, as other prefixes do.

Finally, a clitic added to the newly created morphologically branching verbal host is only in the domain of H-spreading and gets a rising tone, as it was also the case in (7) above.

- (27)     $\text{ní\_ne\_igram}$                       cf.  $*\text{nì\_ne\_igram}$   
           neither\_neg\_play  
           'I don't even play.'

In sum, BCS proclitics in shifting dialects attach at three different levels in the prosodic structure depending on the amount of (morpho)-syntactic branching of their host as shown in (12). As a result, proclitics preceding nominal, adjectival, and verbal hosts can take over the accent from the host if they reach PF in a syntactic configuration that leads to their mapping as internal or affixal clitics, but not if they are mapped to prosody as free clitics.



Although such contexts are limited (and not noted before), there are hosts that exhibit this behavior.<sup>15</sup>

- (29) a. *Dá\_li\_mu vjeruješ?*  
 that\_Q\_him trust  
 ‘Do you trust him?’
- b. *Dá\_li vjeruješ političarima?*  
 that\_Q trust politicians  
 ‘Do you trust politicians?’
- c. *Gdjé\_su\_te zvali?*  
 where\_are\_you invited  
 ‘Where did they invite you?’
- d. *Gdjé\_ste parkirali auto?*  
 where\_are parked car  
 ‘Where did you park the car?’
- e. *Kó\_te\_je zvao?*  
 who\_you\_is called  
 ‘Who called you?’
- f. *Štá\_ti\_je govorila?*  
 what\_you\_is told  
 ‘What was she telling you?’

However, this is not allowed in all constructions. Below I give similar host+clitic sequences, where the spreading does not take place and the host surfaces with a default falling tone.

- (30) a. *Gdjè\_te zovu?*  
 where\_you invite  
 ‘Where are they inviting you?’
- b. *Kò\_te zove?*  
 who\_you call  
 ‘Who is calling you?’
- c. *Kò\_ti\_ga donosi?*  
 who\_you\_it bring  
 ‘Who is bringing it for you?’
- d. *Štâ\_ti kaže?*  
 what\_you tell  
 ‘What is she telling you?’

Based on the clitic mapping proposed above (12), it seems that enclitics can incorporate into the host in (29), but not in (30). (29a-b) can get a simple account assuming that the question particle *-li* shares the head C position with the complementizer. This enables incorporation into the prosodic word, and H-spreading.

The hosts in (29c-d) are wh-words immediately followed by an auxiliary clitic. I assume the auxiliary moves to a head position in the left

<sup>15</sup> This seems to be characteristic of the shifting dialect in the central Bosnia and Herzegovina, while in the south (where other shifts are allowed), these hosts get a falling tone in all environments. In the dialect where this interaction is allowed, the number of enclitics following the host does not play a role.

periphery and the *wh*-word moves to its specifier.<sup>16</sup> Thus, there is nothing separating the host and the first clitic in such cases, and prosodic incorporation can take place. The rest of the clitics are in their lower positions in the syntax (see Bošković 2001). However, in (29c-d), there is no potential host for clitics lower than the *wh*-word, so the rest of the clitics also incorporate into it, yielding the following prosodic word:<sup>17</sup>

(31) [gdjé\_su\_te]<sub>ω</sub>

The lexical High tone from the auxiliary *su* ‘are’ spreads to the toneless host, giving it a rising tone.

Regarding (29e-f), Bošković (2001) gives a number of arguments that the clitic *je* ‘is’ is higher than pronominal clitics in the syntax and occupies the same position as other auxiliary clitics, its placement following other clitics being a result of a low level PF effect. Following this approach, I suggest that the order of clitics following the host in (29e-f) starts as (32a). However, the clitic *je* ‘is’ has an idiosyncratic requirement to occur after all other clitics (see Bošković 2001 on its nature), so it undergoes PF reordering yielding the order in (32b).

(32) a. [ko\_je\_te]<sub>ω</sub> → b. [kó\_te\_je]<sub>ω</sub>

Thus, in all the cases in (29), the clitics are in the same prosodic word with the host. The High tone can spread from the first clitic following the host, giving the host a rising tone.

Concerning (30), notice that there is no auxiliary in any of these cases and that the lexical verb is finite. Assuming the finite lexical verb moves in a position higher than *V* in questions, the word order that the syntax generates in (30) is: [*host V ECL*].

(33) a. [CP *gdje zovu te ...*] cf. (30a) → b. [CP *gdje ~~zovu~~ te zovu...*]

<sup>16</sup> This can be any projection in the left periphery, as long as there are no null copies between the auxiliary clitic and its host.

<sup>17</sup> For ease of exposition, I take all enclitics to be internal clitics here, but it could be the case that only the first clitic is an internal clitic and the rest of them are affixal, or free clitics. However, this cannot be tested since only the first clitic interacts with the host.

In PF, it is clear that (33a) violates the second position requirement of the enclitic because two words that do not make up a constituent precede it. Crucially, following Franks' (1998) proposal that in the cases where pronouncing the highest copy of an element causes a PF violation, it is possible to pronounce its lower copy to satisfy the PF requirement, Bošković (2001) argues that exactly in such cases the lower copy of the element immediately preceding the clitic is pronounced as exemplified in (33b). Thus, the enclitic cannot incorporate into the null copy of verb, nor can it incorporate to *gdje* across the null copy of the verb. The lack of interaction with the accent of the host suggests that the clitic in these cases must attach as a free clitic to *gdje*.

Interestingly, the clitic *je* 'is' is omitted in the presence of the clitic *se* (Bošković 2001), regardless of what function *se* has (see Ridanović (2003; 2012), I give four of them in (34), using Ridanović's terminology). The falling tone on the host in such cases indicates that H-spreading from the clitic to the host does not take place.

- |         |                                                |                      |
|---------|------------------------------------------------|----------------------|
| (34) a. | Kò_ti_se obradovao?                            | <i>Lexemic</i>       |
|         | who_you_SE be.happy                            |                      |
|         | 'Who was happy to see you?'                    |                      |
| b.      | Štà_tj_se još potrošilo u kuhinji?             | <i>Intransitive</i>  |
|         | what_you_SE more run.out in kitchen            |                      |
|         | 'What else did you run out of in the kitchen?' |                      |
| c.      | Štà_tj_se pilo?                                | <i>Desiderative</i>  |
|         | what_you_SE drink                              |                      |
|         | 'What did you feel like drinking?'             |                      |
| d.      | Gdjè_tj_se zavrtjelo u glavi?                  | <i>Physiological</i> |
|         | where_you_SE spin in head                      |                      |
|         | 'Where did you start feeling dizzy?'           |                      |

The absence of H-spreading from clitics to the host in such cases can be captured in the following way. Given that all the sentences in (34a-d) are questions, as noted above, I suggest that the auxiliary (*je*) moves to a head position in the left periphery in the syntax parallel to (29c-f), and that clitics incorporate into the prosodic word of the host, as illustrated in (35) for (34a).

- (35) [ko\_je\_tj\_se]<sub>ω</sub>

As discussed above, the clitic *je* has to undergo PF-reordering to occur after all other clitics. However, the clitic *se* is another clitic that has to follow all other clitics causing a conflict in (35), and blocking PF-reordering which would place *je* to the final position. The only way to resolve the conflict here is to not pronounce *je*, which explains why this clitic is missing in clitic clusters where *se* is present (Bošković 2001 notes that *je* is preferably dropped in the presence of *se*, but does not give an analysis of this kind of *je*-drop).

(36) [ko\_je\_ti\_se]<sub>ω</sub>

The host is thus separated from the overt clitics by a silent *je* and the High tone cannot spread to it. Thus, the host is realized with a default falling tone.

As expected, contexts where the presence or the absence of *se* does not matter are when the host is followed by the clitic *li* or an auxiliary other than *je* (37).

- |                                                                      |                                                                             |
|----------------------------------------------------------------------|-----------------------------------------------------------------------------|
| (37) a. Dá_li_se vratio?<br>that_Q_SE returned<br>'Has he returned?' | c. Gdjé_su_se vratili?<br>where_are_SE returned<br>'Where did they return?' |
| b. Kó_li_se vratio?<br>who_Q_SE return<br>'I wonder who returned?'   |                                                                             |

In such cases, High tone spreading always takes place because the first enclitic that carries a High tone is immediately adjacent to the host.

## 5 Conclusion

I have examined the nature of interaction between PF and the syntax by investigating the phenomenon of accent mobility in the context of clitics in some dialects of BCS. I have shown that prosodization of proclitics and enclitics is sensitive to (morpho)syntactic structure. Contrary to previous proposals, I provided evidence that in a dialect that allows accent shift from nouns, adjectives, and verbs to proclitics, proclitics map to prosody as internal, affixal, or free clitics. I have also observed some novel cases of interaction of BCS enclitics with their hosts.

### References

- Abels, Klaus. 2003. \*[P clitic] – Why? In P. Kosta (ed.), *Investigations into Formal Slavic linguistics*, Part II, 443-460.
- Blumenfeld, Lev. 2012. Vowel-zero alternations in Russian prepositions: prosodic constituency and productivity. In V. Makarova (ed.), *Russian Language Studies in North America*, 43-70. London: Anthem Press.
- Bobaljik, Jonathan. 1995. *Morphosyntax: The Syntax of Verbal Inflection*. Cambridge, MA: MIT dissertation.
- Borsley, Robert & Ewa Jaworska. 1988. A note on prepositions and case marking in Polish. *Linguistic Inquiry* 19, 685-691.
- Bošković, Željko. 1995. Participle movement and second-position cliticization. *Lingua* 95. 245-266.
- Bošković, Željko. 2001. *On the Nature of Syntax-Phonology Interface: Cliticization and Related Phenomena*. Amsterdam: Elsevier.
- Bošković, Željko. 2005. On the locality of Left Branch Extraction and the structure of NP. *Studia Linguistica* 59(1). 1-45.
- Bošković, Željko. 2013. Traces do not head islands: what can PF deletion rescue? In Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki, Asako Uchibori (eds.), *Deep Insights, Broad Perspectives: Essays in Honor of Mamoru Saito*, 56-93. Japan: Kaitakusha.
- Browne, Wayles. 1974. On the problem of clitic placement in Serbo-Croatian. In Richard Brecht and Catherine V. Chivany, *Slavic Transformational Syntax*, 36-52. University of Michigan.
- Ćavar, Damir & Gisbert Fanselow. 2000. Discontinuous constituents in Slavic and Germanic languages. University of Hamburg and University of Potsdam, Ms.
- Franks, Steven. 1998. Clitics in Slavic. Paper presented at the Comparative Slavic Morphosyntax Workshop, Indiana University, Bloomington, June 1998.
- Franks, Steven & Ljiljana Progovac. 1994. On the placement of Serbo-Croatian clitics. *Indiana Linguistic Studies* 7. 69-78.
- Inkelas, Sharon & Draga Zec. 1988. Serbo-Croatian pitch accent: the interaction of tone, stress, and intonation. *Language* 64 (2). 227-248.
- Kiparsky, Paul & Morris Halle. 1977. Towards a reconstruction of Indo-European accent. In Larry Hyman (ed.), *Studies in Stress and Accent*, 209-238. Los Angeles: University of Southern California.

- Riđanović, Midhat & Nadira Aljović. 2009. On the shift of Bosnian accent from host to proclitic: new insights. In Steven Franks, Vrinda Chidambaram, & Brian Joseph (eds.), *A Linguist's Linguist: Studies in South Slavic Linguistics in Honor of E. Wayles Browne*, 387-402. Bloomington, IN: Slavica.
- Riđanović, Midhat. 2003. *Totalni promašaj: Prikaz bosanske gramatike bosanskoga jezika Dž. Jahića, S. Halilovića i I. Palića* (Drugo izdanje). Sarajevo, Bosnia and Herzegovina: Šahinpašić.
- Riđanović, Midhat. 2012. *Bosnian for Foreigners with a Comprehensive Grammar*. Sarajevo, Bosnia and Herzegovina: Rabic Publishing Company.
- Selkirk, Elisabeth. 1996. The prosodic structure of function words. In K. Demuth & J. Morgan (eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, 187-213. Lawrence Erlbaum.
- Stepanović, Sandra. 2010. Left Branch Extraction in multiple wh-questions: a surprise for question interpretation. In W. Browne et al. (eds.), *Proceedings of the 18<sup>th</sup> Formal Approaches to Slavic Linguistics (FASL)*, 502-517. Ann Arbor: Michigan Slavic Publications.
- Stepanović, Sandra. 2011. Differential object marking in Serbo-Croatian: evidence from Left Branch Extraction in negative concord constructions. In J. Bailyn et al. (eds.), *Proceedings of the 19<sup>th</sup> Formal Approaches to Slavic Linguistics (FASL)*, 99-115. Ann Arbor: Michigan Slavic Publications.
- Stjepanović, Sandra. 1999. What Do Second Position Cliticization, Scrambling, and Multiple Wh-Fronting Have in Common? Storrs, CT: UConn dissertation.
- Talić, Aida. 2013. Extraordinary complement extraction: PP-complements and inherently case-marked nominal complements. *Studies in Polish Linguistics* 8. 127-150.
- Talić, Aida. 2014. Off with their heads! – Commonalities between “extraordinary LBE” and apparent N-complement extraction. In C. Chapman et al. (eds.), *Formal Approaches to Slavic Linguistics (FASL)* 22. 380-401.
- Zlatic, Larisa. 1997. *The Structure of the Serbian Noun Phrase*. Austin, TX: University of Texas at Austin dissertation.
- Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59(3). 502-513.

## **Deriving Null Pronouns: A Unified Analysis of Subject Drop in Russian\***

*Egor Tsedryk*  
Saint Mary's University

Russian is well-known for its ability to drop arguments that can be recovered from a linguistic antecedent or a situational context. Nevertheless, the status of null pronouns (*pro*) in this language remains unclear. From one point of view, any null argument (subject or object) in Russian could be attributed to an optional topic drop or a contextually licensed ellipsis (Franks 1995, Lindseth 1998, Gordishevsky and Avrutin 2003, Fehrman and Junghanns 2008, McShane 2009). On the other hand, Müller (2006, 2008) argues that verbal inflectional paradigms in Russian do not involve morphological impoverishment rules (found, for example, in German). For Müller, the lack of such rules implies the ability to license the subject *pro*, meaning that – at least to some extent – Russian is a pro-drop language (see also Růžička 1986, Perlmutter and Moore 2002, Müller 1988, Gribanova 2013). In fact, Russian could be an instance of partial pro-drop (Madariaga 2014).<sup>1</sup> However, unlike other partial pro-drop languages (e.g., Hebrew, Brazilian Portuguese and Finnish), Russian does not impose clear person or tense restrictions on its null subjects: the subject can be dropped with any tense and person (except inverted constructions; see Avrutin and Rohrbacher 1997).

The goal of this paper is to shed light on subject drop in Russian, focusing on interaction between *pro* and fronted XPs. The paper is

---

\* I would like to thank two anonymous reviewers for their insightful comments.

<sup>1</sup> According to Holmberg (2010:94), a partial pro-drop language should (i) be able to drop a definite pronoun if the latter is bound by a local c-commanding antecedent (aka “finite control”) and (ii) have a null indefinite pronoun. Russian has these two properties.

structured as follows: Section 1 outlines the differences between subject and object drop. Section 2 focuses on root clauses, analyzing referential and non-referential null subjects as a by-product of D-linking, probing and  $\phi$ -feature inheritance from C. Section 3 extends this analysis to subordinate clauses, discussing “finite control”, and section 4 concludes.

## 1 Background

It is often assumed that subject and object gaps in finite clauses in Russian are similar in kind (Gordishevsky and Avrutin 2003, Fehrmann and Junghanns 2008). Gribanova (2013) has recently shown that an object gap can in fact be an instance of Verb-Stranding Verb Phrase Ellipsis (VVPE), and that null subjects should be differentiated from the genuine object drop. In what follows, I highlight the relevant points of Gribanova’s analysis (sections 1.1 and 1.2) and discuss its implications for the clausal architecture in Russian (section 1.3).

### 1.1 VVPE and Object Drop

The gap ( $\emptyset$ ) in (1b) and the one in (2) look identical on the surface. However, according to Gribanova, they are different syntactically: (1b) is an example of VVPE, while (2) instantiates an empty category in the internal argument position (the gap and its antecedent are in bold). Crucially, the gap in (1b) has a linguistic antecedent in (1a), while the gap in (1b) can only be recovered from a situational context.

- (1) a. Kto-nibud’ podnimet **vazu kotoraja ležit na polu**?  
 somebody will.pick.up vase that is.lying on floor  
 ‘Will anybody pick up the vase that is lying on the floor?’  
 b. Ne volnujtes’, ja sejčas podnimu  $\emptyset$ .  
 NEG worry I now will.pick.up  
 ‘Don’t worry, I’ll pick (it) up now.’
- (2) [**Something** is lying on the floor.]  
 Ne vstavaj, ja podnimu  $\emptyset$ .  
 NEG get.up I will.pick.up  
 ‘Don’t get up, I’ll pick (it) up.’  
 (Gribanova 2013:107; based on Gordishevsky and Avrutin 2003:7)

Gribanova argues that VVPE in Russian is made possible by a short verb movement to an aspectual head, whose projection (AspP) is sandwiched between TP and  $v$ P. The remnant  $v$ P (including the internal argument) is

then elided under identity. The difference between VVPE and object drop (i.e. object gap with a situational antecedent) is sketched in (3).

- (3) a. [TP T [AspP [Asp V-complex] [<sub>NP</sub> t<sub>v</sub> [VP t<sub>v</sub> ~~DP~~]]]]  
 b. [TP T [AspP [Asp V-complex] [<sub>VP</sub> t<sub>v</sub> [VP t<sub>v</sub> [DP *pro*]]]]]

Empirical evidence for this bifurcation comes from islands. As shown in (4), object drop cannot occur within a complex NP island, as opposed to VVPE in (5).

- (4) [**Something** falls, no one wants to get it.]  
 #Tot fakt, čto nikto ne podnjajl **∅**, menja očen' ogorčæet.  
 the fact that no-one NEG picked.up me very upsets  
*intended*: 'The fact that no one picked (it) up very much upsets me.' (Gribanova 2013:109)<sup>2</sup>
- (5) Kažetsja, čto nikto ne podnjajl **tu vazu**.  
 seems that no-one NEG picked.up that vase  
 'It seems that no one picked up that vase.'  
 Tot fakt, čto nikto ne podnjajl **∅** menja očen' ogorčæet.  
 the fact that no-one NEG picked.up me very upsets  
 'The fact that no one picked (it) up very much upsets me.'  
 (Gribanova 2013:110)

Once we distinguish VVPE and object drop, it is important to exclude the former from the discussion pertaining to pro-drop phenomena in Russian. The subsequent question is whether or not null subjects should be put under the same umbrella as null objects.

### 1.2 Subject-Object Asymmetry

Gribanova argues that null subjects in Russian finite clauses should have different licensing conditions in syntax than null objects, even though both types of pro-drop are ultimately related to discourse in some way or another. Let me summarize Gribanova's argument, which is based on the island-like behaviour of *čto*-clauses in Russian.

<sup>2</sup> Gribanova reports that (4) has an acceptability rate of 2.9/7 (based on judgements of 130 speakers).

In (6a), it is shown that a subject pronoun can be omitted in *čto*-clauses, provided that there is an antecedent in the matrix clause. In (6b), we see that an object pronoun, on the other hand, must be overt.

- (6) a. **Volodja** skazal, čto **(on)** kupit zelěnuju lampu.  
 Volodja said that he will.buy green lamp  
 ‘Volodja said that (he) will buy a green lamp.’
- b. **Volodja** skazal, čto Nadja ljubit **\*(ego)**.  
 Volodja said that Nadja loves \*(him)  
 ‘Volodja said that Nadja loves him.’ (Avrutin and Babyonyshev 1997:248; cited in Gribanova 2013:114-115)

Interestingly, an object with a situational antecedent cannot be dropped either:

- (7) [**Something** is lying on the floor.]  
 #Ja byl uveren, čto kto-to uže podnjajl **∅**.  
 I was sure that someone already picked.up  
*intended*: ‘I was sure that someone already picked (it) up.’  
 (Gribanova 2013:115)

Finally, it is a well-established fact that *čto*-clauses do not allow extraction, as we can observe in (8).

- (8) a. \*Kakuju knigu<sub>i</sub> ty ne uveren, čto Petja pročital *t<sub>i</sub>*?  
 which book you NEG sure that Petja read  
*intended*: ‘Which book are you not sure that Peter read?’
- b. \*Kto<sub>i</sub> ty ne uveren čto *t<sub>i</sub>* pročital etu knigu?  
 who you NEG sure that read this book  
*intended*: ‘Who are you not sure (that) read this book?’  
 (Gribanova 2013:116)

Based on a parallel between examples like (7) and (8a), Gribanova concludes that object drop involves an operator-variable relation. As for null subjects, since they can occur within *čto*-clauses, their syntactic distribution should be different. Gribanova suggests a Rizzi-style (1986) account of subject drop in Russian, assuming that *pro* has unvalued  $\phi$ -features valued by an Agree relation with T. The difference between subject and object drop is schematized in (9) (based on Gribanova 2013:113).

- (9) a. [CP [TP T [Aspp *pro* ... V ... Obj]]]                      *Subject drop*  
 b. [CP Op<sub>i</sub> [C' C [TP Subj ... V ... t<sub>i</sub>]]]                      *Object drop*

To implement Gribanova's proposal, I would like to suggest that object drop in Russian is an instance of a null aboutness-shift topic (A-Topic), which is the uppermost topic at the left periphery (see Frascarelli and Hinterhölzl 2007, and Bianchi and Frascarelli 2010). Note that if A-Topic is overt, the object pronoun has to be overt as well, as we can observe in (10).<sup>3</sup>

- (10) a. Ax, da, **moja novaja šuba**... ja \*(eë) kupila na rynke.  
 oh yes my new fur.coat I it boght at market  
 'Oh, yes, my new fur coat... I bought it at the market.'  
 b. Kstati o **šube**, ja \*(eë) kupila na rynke.  
 by.the.way about fur.coat I it boght at market  
 'By the way, about the fur coat, I bought it at the market.'

Furthermore, compare (10) with (11), where the null object ( $\emptyset_2$ ) has a situational referent, and it co-occurs with a null subject ( $\emptyset_1$ ). In (10), the A-Topic is base-generated, and it is linked to a resumptive pronoun in situ. In (11), the A-Topic is a *pro* moved to Spec,CP, and  $\emptyset_2$  is its copy interpreted as a bound variable.

- (11) [Sveta<sub>1</sub> says, showing a new fur coat<sub>2</sub> to her friend:]  
 Vot posmotri,  $\emptyset_1$  včera kupila  $\emptyset_2$  za polceny.  
 here have.a.look yesterday bought for half.price  
 'Have a look, (I) bought (it) yesterday for half price.'

<sup>3</sup> An anonymous reviewer points out an asymmetry between subject and object drop in (i), where the topic is introduced by *vot* 'here is' (everything else is the same as in (10a)):

- (i) a. Vot moja novaja šuba,  $\emptyset$  kupila  $\emptyset$  na rynke.  
 b. \*Vot moja novaja šuba, ja kupila  $\emptyset$  na rynke.  
 c. ?Vot moja novaja šuba, ja kupila eë na rynke.  
 d. Vot moja novaja šuba,  $\emptyset$  kupila eë na rynke.

I would suggest analyzing (i) as two independent clauses: *vot* + DP (presentational clause) and CP, as in (ii).

- (ii) [Vot moja novaja šuba<sub>i</sub>] [CP Op<sub>i</sub> [C' C [TP  $\emptyset_{\text{Subj}}$  kupila t<sub>i</sub>]]] (cf. ia)

If an overt (but not covert) subject is another topic intervening in the operator-variable relation, (ib) is expected. In (ic-d), there is no null operator in Spec,CP. Note that the subject cannot be dropped in (10), showing that an overtly topicalized object involves an overt subject.

As for  $\theta_1$ , I will show in section 2 that it is a *pro* moved to Spec,TP and probed by  $\phi$ -features of C. Both  $\theta_1$  and  $\theta_2$  are linked to C-domain (and ultimately to context), but this linking is mediated by different syntactic processes, involving left-edge movement in the case of  $\theta_2$  and morpho-syntactic features in the case of  $\theta_1$ . Before I expand upon the syntactic processes standing behind subject drop in finite clauses in Russian, it is important to make a few clarifications about the clause structure.

### 1.3 Clause Structure

The key assumption in Gribanova's (2013) analysis of VVPE in Russian is that the verb moves to Asp without reaching T. Preverbal placement of adverbs in Russian (S Adv VO) further implies that adverbs are merged outside  $\nu$ P, but inside AspP (see Dyakonova 2009:33-35 for a discussion). I assume that adverbs can be stacked in multiple specifiers of AspP. For example in (12a), the manner adverb (*podrobno*) marks the lower boundary of AspP, the time adverb (*sejčas*) marks its higher boundary, and there are scrambled pronouns in-between (see 12b)).

- (12) a. Ja sejčas tebe vsë podrobno ob"jasnju.  
 I now you everything in.detail will.explain  
 'I'll now explain you everything in detail.'
- b. [<sub>TP</sub> Ja [<sub>T</sub> T [<sub>AspP</sub> sejčas tebe vsë podrobno [<sub>Asp</sub> ob"jasnju] [<sub>VP</sub> ... t<sub>V</sub> ... ]]]]]

If an object moves in front of the subject, it tends to be contrastive, as in (13), but it can also be a given (familiar) topic that has a salient referent. More generally, dislocated XPs are D(iscourse)-linked.<sup>4</sup>

- (13) *Tebe* ja sejčas vsë podrobno ob"jasnju.  
 you I now everything in.detail will.explain  
 'To you (in particular) I'll now explain everything in detail.'

I undertake a non-cartographic approach, adopting Bailyn's (2013) proposal that left dislocation results in adjunction to TP. Spec,TP is a so-called "EPP position" associated with a grammatical subject (Subj). This

<sup>4</sup> "A constituent is D-linked if it has been explicitly mentioned in the previous discourse, is situationally given by being physically present at the moment of communication, or can be easily inferred from the context by being in the set relation with some other entity or event in the preceding discourse." (Dyakonova 2009:73)

position is not restricted to nominative DPs, and any XP can end up in Spec,TP, provided that it is structurally prominent enough to satisfy EPP (Extended Projection Principle). Finally, adding an A-Topic in Spec,CP (section 1.2), we have the following picture.<sup>5</sup>

- (14) [CP A-Topic [C' C [TP *D-linked XP(s)* [TP Subj [T' T [AspP XP(s) V ...

In sum, we should retain the following points: (i) TP and vP are separated by AspP; (ii) V moves to Asp, but not to T; (iii) AspP can have multiple specifiers hosting discourse-neutral XPs; (iv) D-linked XPs are outside AspP; (iv) EPP can be satisfied by a non-agreeing XP. In the subsequent discussion, I will focus on the cases where the subject is dropped and a preverbal XP can be either inside or outside AspP. I show that only D-linked XPs block referential subject drop in Russian.

## 2 Subject Drop in Root Clauses

Section 2.1 introduces data that, to my knowledge, have not been previously discussed in the literature on subject drop in Russian. Section 2.2 presents assumptions about feature specification of C, T and *pro*. Section 2.3 shows how referential and non-referential null subjects are derived from a single feature specification without postulating a special kind of an empty pronoun for each particular case.

### 2.1 Subject Drop and Fronted XPs

In (15), we have a 1st person plural null subject, as we can infer from the second sentence (cf. *nas* 'us'). The subjectless verb (in bold) is in the past tense and does not show person agreement. (I do not indicate the null subject to abstract away from its structural position for a moment.)

- (15) Nedavno **videli** Svetu. Ona byla očen' rada nas vstretit'.  
 recently saw<sub>PL</sub> Sveta<sub>ACC</sub> she was very happy us to.meet  
 '(We)'ve recently seen Sveta. She was very happy to meet us.'

Consider now (16), where the accusative DP is a given topic, preceding the adverb, as opposed to (15). The verb in (16) has exactly the same form as in (15), but the null subject cannot have a referential reading (only an arbitrary one).

<sup>5</sup> A-Topic in (14) corresponds to Bailyn's (2013:268-273) "left-edge topicalization".

- (16) [Sveta disappeared. Nobody knows where she is. In a conversation about Sveta, somebody says:]  
*Svetu*          *nedavno*    ***videli***    *na*    *rynke*.  
*Sveta*<sub>ACC</sub>    recently    saw<sub>PL</sub>    at    market  
 ‘Sveta was recently seen at the market.’ (# We’ve recently seen...)

A minimal pair in (17) presents another interesting case. It is possible to drop the subject of the second clause in (17a), but not in (17b) where *mne* ‘to me’ is inserted in front of the verb.<sup>6</sup>

- (17) a. *Ja*    *tol’ko*    *čto*    *videl*    *Svetu*.    (*Ona*)    ***skazala***    *čto*    *naš*  
 I    just          saw    Sveta    she    said<sub>F</sub>    that    our  
*dom*    *uže*    *prodan*.  
 house    already    sold  
 ‘I’ve just seen Sveta. (She) said that our house had already been sold.’
- b. *Ja*    *tol’ko*    *čto*    *videl*    *Svetu*.    \*(*Ona*)    *mne*    ***skazala***    *čto*  
 I    just          saw    Sveta    she    me    said<sub>F</sub>    that  
*naš*    *dom*    *uže*    *prodan*.  
 our house    already    sold  
 ‘I’ve just seen Sveta. She told me that our house had already been sold.’

Intuitively, what seems to be at stake in (17b) is that *mne* intervenes between the verbal agreement and the antecedent. Note that the verb agrees in gender (and number), but not in person. According to Avrutin and Rohrbacher (1997:45), the person feature should be “supplied from the discourse presupposition via [...] discourse binding.” This amounts to saying that *mne* in (17b) blocks discourse binding, preventing the verbal inflection from getting a referential index. The intuition behind discourse binding of the past tense inflection in Russian might be on the right track, but we need a formal account of the blocking effect observed in (17b).

In (18), we have another case where the null subject is blocked by a D-linked constituent in front of the subject. The demonstrative *eto* ‘this’ is a given topic referring to a background situation (i.e. a strange noise in the garage); it is not a cleft construction (‘It is me that...’).<sup>7</sup>

<sup>6</sup> According to an anonymous reviewer, (17b) can be fixed as *Ja tol’ko čto videl Svetu. (Ona) skazala mne čto...* In this case, *mne* does not move to TP, staying inside VP.

<sup>7</sup> Imagine the same situation, in which Peter answers with irritation:

- (18) [Peter repairs his car in the garage, but his wife does not know it. Suddenly she hears a strange noise and asks: “What are you doing there?” Peter says:]  
*Eto* \*(ja) **čínju** mašinu.  
 this I repair<sub>1SG</sub> car<sub>ACC</sub>  
*intended*: ‘[It is that] I repair the car.’

The verbal inflection in (18) does not need to be discourse bound in the sense of Avrutin and Rohrbacher (1997), since the person feature is already present in the structure. We still have a blocking effect, indicating that *eto* may intervene in a more local relation, which could be a prerequisite for a more general relation with discourse.<sup>8</sup>

All in all, examples (16), (17b) and (18) show that (i) the null subjects in Russian are linked to discourse, (ii) verbal inflection in Russian is not pronominal and (iii) referential null subjects are blocked by left peripheral topics.

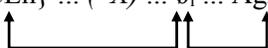
From the point of view of Sigurðsson’s (2011) theory of argument drop, Russian subject drop seems to pattern like a Germanic type of topic drop. Sigurðsson (2011:282) proposes a C/Edge linking generalization, stating that “any definite argument, overt or silent, *positively* matches *at least one* [C/edge-linking feature] in its local C-domain.” In Sigurðsson’s terms, a C/edge linking feature can be a logophoric agent (1st person), a logophoric patient (2nd person), or it can be any kind of topic. According to Sigurðsson, there are three main types of null subjects: (i) Chinese topic drop, when the subject does not move to the C/edge, (ii) Romance pro-

- 
- (i) Čto, čto... MašInu **čínju**.  
 what what car<sub>ACC</sub> repair<sub>1SG</sub>  
 ‘What, what... I repair a CAR.’

The accusative object precedes the verb, but the null subject is still possible. Here the object receives a focal stress (cf. capitalized syllable), but it is not D-linked. This means that the accusative object in (i) is part of the middle-field focus (it is inside AspP). I thank Ora Matushansky for bringing this example to my attention.

<sup>8</sup> An anonymous reviewer suggests that *\*Eto čínju mašinu* could be ungrammatical because *eto* prevents the focused VP from moving to the C-domain, since the question “What are you doing?” would most naturally be answered as *Čínju mašinu*. I agree. However, it should be made clear that (18) is only felicitous as an explanation to a background event (e.g., strange noise in the garage); it is not a strictly speaking all-new-information utterance. My argument is: if *Eto ja čínju mašinu* is felicitous in context A, and if Russian verbal inflection alone were able to license a null subject, *\*Eto čínju mašinu* would also be grammatical in context A, contrary to fact. Ungrammaticality in context B (all-new-information answer) is immaterial for this argument.

drop, when an agreement on the verb is basically an incorporated pronoun, and (iii) Germanic topic drop that requires movement to the C/edge. In Icelandic, for example, the subject can be dropped if there is no adverb or any other category in the sentence-initial position (see (19a, b))—that is, if there is no intervener (X) between the null subject and a C/edge linking feature (CLn), as shown in (19c) (Sigurðsson 2011:287; Agr stands for subject-verb agreement).

- (19) a. **Tala** stundum íslensku. *Icelandic*  
 speak<sub>1SG</sub> sometimes Icelandic  
 ‘I sometimes speak Icelandic.’  
 b. \**Stundum tala* íslensku.  
 sometimes speak<sub>1SG</sub> Icelandic  
 c. [CP ... {CLn} ... (\*X) ...  $\emptyset_i$  ... Agr<sub>i</sub> ...  


Russian is not a V2 language: the verb stays low and adverbial modifiers do not necessarily intervene. Consider for example (20a). Note that if the PP *včera na rynke* is inside AspP, and the null subject moves to Spec,TP to check EPP, there is nothing that could intervene between the subject and a C/edge linking feature, as shown in (20b).

- (20) a. *Včera na rynke videli* Svetu, no ona prošla mimo  
 yesterday at market saw<sub>PL</sub> Sveta but she passed by  
 i s nami daže ne pozdarovalas’  
 and with us even NEG greeted  
*intended:* ‘(We) saw Sveta yesterday at the market, but she  
 passed by and didn’t even greet us.’  
 b. [CP ... {CLn} ... [TP  $\emptyset_i$  [T’ T- $\phi_i$  [AspP *včera na rynke videli* ...  


If the PP moved to Spec,TP, and the null subject stayed inside AspP, we would expect a blocking effect similar to the one in (19b).

There are two more comments with regard to Sigurðsson’s C/edge linking model. First, the null subject does not have to move to Spec,CP in order to match a C/edge linking feature. Sigurðsson (2011:282) assumes that “movement tucks in to the right of its probe instead of adding structure to its left.” So every time a null subject moves to the C/edge, it moves to a position c-commanded by a corresponding C/edge linking feature. Such

a movement does not create a Spec-head relation, but a probe-goal relation. I do not need to make this assumption in a non-cartographic approach adopted here. The null subject can move to Spec,TP (in a standard way), where it would be probed by C.

Second, if every argument has its own feature to match in the C-domain, it is not clear why we should have intervention effects at all. For example in (18), *eto* would match a topic feature, while the subject matches the 1st person. Under relativized minimality, *eto* should not be an intervener. In (17b), we can stipulate that the 1st person object blocks the 3rd person subject, but we run into a problem in (16), where the 3rd person object blocks the 1st person subject.

My analysis is inspired by Sigurðsson's idea that there is a C/edge feature that scans the context and ensures a linking between discourse and syntax. However, my implementation of this idea is somewhat different. Section 2.2 presents my assumptions about  $\phi$ -feature specifications, and section 2.3 shows how null subjects are derived.

## 2.2 Assumptions

In line with Holmberg (2005), I take *pro* to be a bundle of unvalued  $\phi$ -features, labeled as  $\phi P$ . Following Pesetsky and Torrego (2007), I assume that a feature F can be valued or unvalued and either interpretable (*i*F) or uninterpretable (*u*F). In this way, there are four logical possibilities for  $\phi$ -features, distributed as follows:

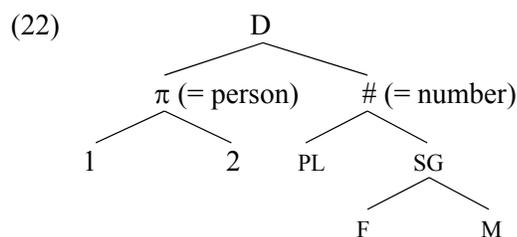
- (21) a. C valued *u* $\phi$   
 b. T unvalued *u* $\phi$   
 c.  $\phi P$  unvalued *i* $\phi$   
 d. DP valued *i* $\phi$

Unvalued  $\phi$ -features are valued via Agree, understood here as a process of feature sharing (Pesetsky and Torrego 2007). Uninterpretable  $\phi$ -features probe their local c-commanding domain (to find the closest interpretable  $\phi$ -set) and trigger feature sharing between two syntactic positions. For convenience, it suffices to say that a head (C or T) probes the corresponding category (DP or  $\phi P$ ). The ultimate goal of the syntactic derivation is to value uninterpretable features so they can be deleted before the syntactic structure is sent to interfaces; otherwise, the derivation would crash.

Following Chomsky (2008), I assume that T can inherit  $\phi$ -features from C. However, I do not assume that feature inheritance is an automatic

process – it applies only if needed (this will be clarified in section 2.3). In fact, it is not just any feature that can be inherited from C. To make this point more explicit, I need first to spell out my assumptions about  $\phi$ -features in C.

Generally, I assume that C does not have a predetermined set of  $\phi$ -feature values (1st person plural, 2nd person singular, 3rd person feminine, etc.). That is, C's  $\phi$ -feature composition depends on discourse or context. This assumption rejoins Sigurðsson's (2011) idea of C/edge linking implemented in a feature-geometric perspective. More precisely, I assume that C enters the derivation with a maximally specified set of  $\phi$ -features organized in a hierarchical manner, as depicted in (22).<sup>9</sup>



Following Camacho (2013:120), I use [D] as the highest node in a  $\phi$ -feature geometry, assuming that its role is similar to Sigurðsson's C/edge linking. To be more specific, the role of [D] is twofold. On the one hand, it scans the context and checks those features that are contextually relevant. On the other hand, it probes the C-domain (D-probing). Discursively incompatible features are deleted from the bottom up, going from a more specific structure, as in (22), to a less specific one, depending on the context. The remaining  $\phi$ -features are D-linked in a sense that they have a discursive or a contextual referent. D-linking and D-probing are two

<sup>9</sup> The feature geometry in (22) is based solely on Russian verbal inflection (cross-linguistic perspective is beyond the scope of this paper). Every node dominated by  $\pi$  or  $\#$  is exclusive with regard to its sister and implicational with regard to its mother. For example, there is no gender marking in plural; masculine or feminine implies singular, etc. Bare singular [SG] corresponds to neuter (neither masculine nor feminine) and bare  $\pi$  corresponds to 3rd person (neither speaker nor addressee). There is nothing that would correspond to bare  $\#$  (number is either singular or plural in Russian).  $\pi$  and  $\#$  are unified under D (referential node) without excluding each other; these nodes represent two independent means of discursive identification, based on speech act participant roles and individuation, respectively.

simultaneous processes: [D] probes the first available category in the syntactic domain and ensures that it matches a salient contextual antecedent. As a result, there is a maximal correspondence between the C/edge and the context.

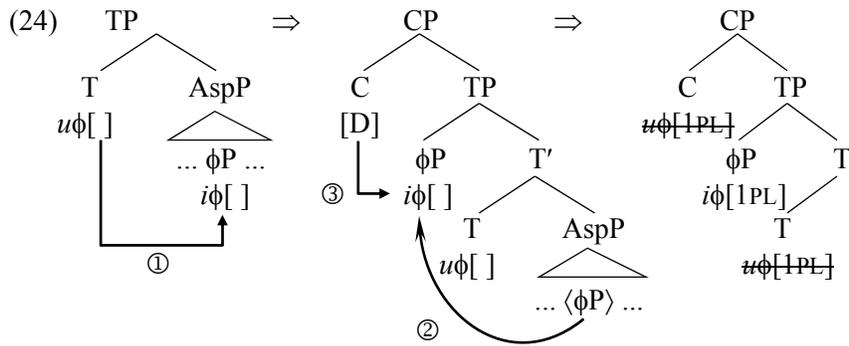
Now in connection with feature inheritance mentioned earlier, I assume the following condition:

- (23) *Condition on D-linking of  $\phi$ -features*  
 Only the edge head can have D-linked  $\phi$ -features.

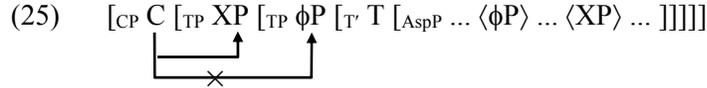
T is not an edge head. Thus, if T inherits  $\phi$ -features from C, it inherits only those features that are not D-linked. In section 2.3, I suggest that non-referential null subject constructions are the only instance of feature inheritance in Russian.

2.3 *Deriving Null Subjects*

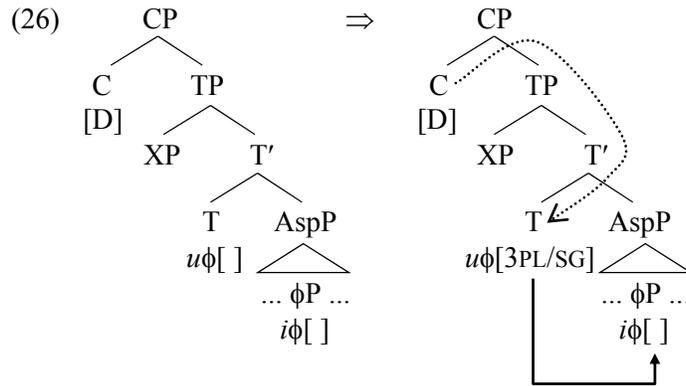
Derivation of referential null subjects proceeds in three major steps, as numbered in (24) (unvalued features are presented with empty brackets). As a first step, T probes  $\phi$ P but T cannot value its  $u\phi$  because the targeted  $i\phi$  does not have values. The second step is movement of  $\phi$ P to Spec,TP (EPP). This movement allows the third step: [D] probes  $\phi$ P resulting in an Agree relation between C and  $\phi$ P. That is, D-linked  $\phi$ -features (e.g., 1PL) are shared with  $\phi$ P and, subsequently, with T. Finally, all  $u\phi$ -features are deleted, and what is interpreted is the bundle of  $\phi$ -features in Spec,TP.



If there is a D-linked XP adjoined to TP, as in (25), it would intervene between C and  $\phi$ P, and  $\phi$ -features of T would remain unvalued.



Now suppose that  $\phi P$  stays inside  $AspP$  and an  $XP$  moves to  $Spec,TP$  (EPP), as in (26). In this case,  $XP$  prevents  $[D]$  from probing  $\phi P$ , and  $T$  cannot value its  $\phi$ -features. The derivation would crash unless there is a last resort strategy, which – I assume – is feature inheritance:  $T$  inherits  $\phi$ -features from  $C$ . According to (23),  $T$  cannot have D-linked  $\phi$ -features, and it inherits a default value of  $[\pi]$  (3rd person) with either  $[SG]$  or  $[PL]$  value of  $[\#]$ . Subsequently,  $T$  shares its inherited  $[3PL/SG]$  value with  $\phi P$ , which becomes a non-referential null subject.



If  $T$  inherits  $[3PL]$ , we obtain an arbitrary *pro* construction, as in (27a). If it inherits  $[3SG]$ , we have an impersonal construction, as in (27b).

- (27) a. Menja sejčas toržestvenno **vstretjat** u trapa samolëta.  
 me<sub>ACC</sub> now solemnly will.meet<sub>3PL</sub> at exit of.aircraft  
 ‘I will now be solemnly greeted at the aircraft’s exit.’  
 b. Menja sejčas **vybrosit** za bort.  
 me<sub>ACC</sub> now throw<sub>3SG</sub> over board  
 ‘I will now be thrown overboard.’

The hallmark of the constructions in (27) is that *pro* is not in  $Spec,TP$ . Staying inside  $AspP$ , it may incorporate into the verb, which raises to  $Asp$ . In other words, impersonal and arbitrary *pro* are the only cases in Russian where the verbal agreement could be considered as an incorporated

pronoun, remotely resembling the Romance type of pro-drop; otherwise, the referential *pro* stays in Spec,TP, and it has to be linked to C for full interpretation.<sup>10</sup>

Note that I do not postulate a special kind of null pronoun for constructions like those in (27). Both referential and non-referential null subjects are derived from feature specifications in (21) (with a condition in (23)) and general syntactic processes, such as feature probing and feature inheritance. The next section discusses embedded CPs.

### 3 Subject Drop in Embedded Clauses

Null subjects in the embedded finite clauses are derived in the same way. Embedded null subjects in Russian are referentially dependant on a matrix antecedent. This is known as “finite control” (FC), illustrated in (28), where the null subject cannot be anybody else but *Petja*.<sup>11</sup>

- (28) **Petja**<sub>1</sub> uveren, čto **Ø**<sub>1/\*2</sub> skoro **stanet** čempionom.  
 Petja sure that soon will.become<sub>3SG</sub> champion  
 ‘Petja is sure that he (= Petja) will become a champion soon.’

Livitz (2014) has recently proposed an Agree-based analysis of FC in Russian. She assumes that the embedded C has unvalued *uφ*-features. Agree relations are established by probing heads, as shown in (29) (based on Livitz 2014:93). Abstracting away from technical details, C probes  $\phi$ P, the matrix little *v* probes C, but the latter has unvalued  $\phi$ -features, and the little *v* has to probe again, targeting the matrix subject. Finally, the  $\phi$ -features of the matrix subject are shared with the little *v*, C and  $\phi$ P.

- (29) [<sub>VP</sub> *Petja* [<sub>v'</sub> *v* [<sub>VP</sub> V [<sub>CP</sub> C [<sub>TP</sub>  $\phi$ P ...  
           *iφ*[3SG]        *uφ*[ ]                    *uφ*[ ]                    *iφ*[ ]  
           ↑                    |                    ↑                    |                    ↑

<sup>10</sup> For constructions like *Temneet* ‘It becomes dark’, there are two possibilities: (i) there is a null locative in Spec,TP, or (ii)  $\phi$ -features in C are not D-linked, since there is no salient antecedent (the condition in (23) states that  $\phi$ -features in C can be D-linked, but they do not have to).

<sup>11</sup> In Tsedryk (2012, 2013), I propose a movement analysis of FC in Russian to account for: (i) subject orientation, (ii) nominative antecedent, (iii) obligatory control diagnostics, and (iv) partial sensitivity to islands. See also Livitz 2014 for a discussion of some of these properties.

Let us now consider how XP-fronting interacts with the null subjects in the embedded clause.

Interestingly, the first person accusative *nas* does not block the embedded null subject in (30a). However, if PP is moved in front of *nas*, the null subject becomes degraded and the overt pronoun is a preferred option, as shown in (30b).

- (30) a. Sveta skazala, čto *nas* **videla** [<sub>PP</sub> včera na rynke].  
 Sveta said<sub>F</sub> that <sub>USACC</sub> saw<sub>F</sub> yesterday at market  
 ‘Sveta said that she (= Sveta) saw us yesterday at the market.’  
 b. Sveta skazala, čto [<sub>PP</sub> včera na rynke] ?\*(ona)  
 Sveta said<sub>F</sub> that yesterday at market she  
*nas* **videla**.  
<sub>USACC</sub> saw<sub>F</sub>  
 ‘Sveta said that yesterday at the market she saw us.’

In (30a),  $\phi$ P checks EPP, moving to Spec,TP, while *nas* presumably stays lower (inside AspP). In (30b), EPP is checked by PP;  $\phi$ P stays lower, and C cannot probe it. Thus, only an overt pronoun (DP) is possible in (30b).

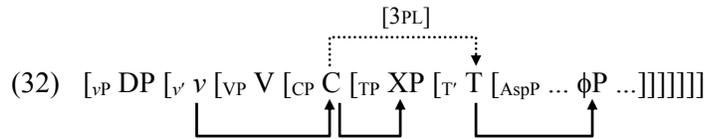
It is possible that the root and the embedded C have a different  $\phi$ -feature properties—that is, a valued  $\phi$ -set for the former, but an unvalued one for the latter (in line with Livitz 2014). However, a question arises with respect to non-referential null subjects in the embedded contexts. Compare (30) with an analogous pair of sentences in (31), where the embedded verb has an arbitrary 3rd person plural inflection that does not match the subject agreement in the matrix clause. Note that the fronted PP in (31b) does not cause any problems.

- (31) a. Sveta skazala, čto *nas* **videli** [<sub>PP</sub> včera na rynke].  
 Sveta said<sub>F</sub> that <sub>USACC</sub> saw<sub>PL</sub> yesterday at market  
 ‘Sveta said that we were seen yesterday at the market.’  
 b. Sveta skazala, čto [<sub>PP</sub> včera na rynke] *nas* **videli**.<sup>12</sup>  
 Sveta said<sub>F</sub> that yesterday at market <sub>USACC</sub> saw<sub>PL</sub>  
 ‘Sveta said that yesterday at the market we were seen.’

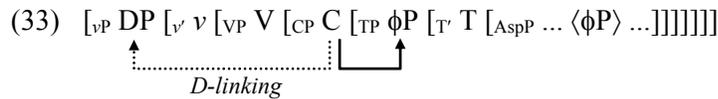
If  $\phi$ P, T and C all have unvalued  $\phi$ -features, where would the plural number come from in (31)? One possible solution is to stipulate that there is special kind of  $\phi$ P that has a valued  $\phi$ -set. Another possibility is to admit

<sup>12</sup> (31b) sounds better with a perfective form *uvideli*.

that the embedded finite C (just like the root one) has valued  $\phi$ -features, along the lines of my assumptions presented in section 2.2. Probing relations would be similar to those in (30) except: (i) that  $\nu$  does not probe its specifier, and (ii) that C probes XP, not  $\phi$ P, as illustrated in (32). As a result of (ii), T inherits  $\phi$ -features from C, and T probes  $\phi$ P staying inside AspP. XP in (32) is either *nas* in (31a) or PP in (31b).<sup>13</sup>



As for FC, it is derived from D-linking and D-probing (triggered by [D] in C). C is linked to the most salient contextual antecedent (i.e. the subject of the matrix clause), and it probes  $\phi$ P in Spec,TP, as shown in (33) (the relation between  $\nu$  and C is irrelevant).



Recall from section 2.2 that D-linking and D-probing are two simultaneous processes that ensure a maximal match between the most prominent constituent in the syntactic domain (EPP position) and the most prominent entity in the context or discourse. In (32), D-linking is not shown. In this case, C cannot be D-linked to the matrix subject because the latter would

<sup>13</sup> According to my analysis, (30a) and (31a) are not structurally identical:

(i) ...[<sub>CP</sub> *čto* [<sub>TP</sub>  $\phi$ P [<sub>T'</sub> T [<sub>AspP</sub> *nas*  $\langle \phi$ P *videla včera na rynke*]]]]] (cf. 30a)

(ii) ...[<sub>CP</sub> *čto* [<sub>TP</sub> *nas* [<sub>T'</sub> T [<sub>AspP</sub>  $\langle nas \rangle \phi$ P *videli včera na rynke*]]]]] (cf. 31a)

There is another interesting option that is worth exploring with regard to (30a): nothing moves to Spec,TP and EPP is satisfied by merging an overt C (*čto*):

(iii) ... [<sub>CP</sub> *čto* [<sub>TP</sub> T [<sub>AspP</sub> *nas*  $\phi$ P *videla včera na rynke*]]]]] (instead of (i))

Assuming that *nas* and  $\phi$ P are in multiple Specs of Asp, they are both equally accessible for (equidistant from) C. That is, *nas* is not an intervener between C and  $\phi$ P unless it moves to Spec,TP. Technically, we could replace EPP with a more general *T-extension requirement* (TER) (Kučerová 2014:137): “If Merge(T,  $\alpha$ ) applies, Merge(T',  $\beta$ ) must be the next step of the derivation, where T' is a projection of T and  $\beta$  belongs to the same phase domain as T”. According to Kučerová (2014:139), *pro* should not be in Spec,TP if TER can be satisfied by other means. This hypothesis is in line with my analysis of non-referential *pro*, which should never be in Spec,TP. As for referential *pro* in root clauses, we would have to assume that TER is satisfied by a null C (not *pro*). But why should we replace one empty category by another if both share their features via Agree in any case?

not match XP (*nas* or PP) in Spec,TP. C is D-linked to a prominent antecedent outside the matrix clause.

#### 4 Conclusion

To conclude, Russian has fairly heterogeneous properties of its argument drop phenomena, which led to contradictory claims with regard to its pro-drop status. According to Gribanova (2013), object drop (with a discursive antecedent) is the only instance of a genuine topic-drop in Russian (a null operator binding a variable). As for the subject drop, she suggested a Rizzi-style analysis, assuming that subject *pro* is licensed by Agree. Nevertheless, Russian is not a pro-drop language whose subject-verb agreement is an incorporated pronoun. Focusing on the interaction between *pro* and dislocated XPs, I have shown that Russian has a pattern similar to Germanic languages: a fronted XP blocks a referential *pro*. However, this situation is obscured in Russian by the fact that the verb stays low, and preverbal XPs do not necessarily interfere with *pro* that raises to Spec,TP. According to my analysis, there is no special null pronoun for referential and non-referential uses. Interpretation of *pro* largely depends on its structural position and accessibility to C:

- (34) a.  $[\text{CP } C [\text{TP } \textit{pro} [\text{T}' \text{T} [\text{AspP} \dots \langle \textit{pro} \rangle \dots]]]]$  *referential*
- 
- b.  $[\text{CP } C [\text{TP } \textit{XP} [\text{T}' \text{T} [\text{AspP} \dots \textit{pro} \dots]]]]$  *non-referential*
- 

More generally, I consider C as a locus of D-linked  $\phi$ -features, suggesting that only non-D-linked values, [3PL] or [3SG], can be inherited by T (dotted arrow in (34b)). The same analysis was applied to the embedded finite clauses and, overall, it offers a unified view of null subjects in Russian, without additional stipulations.

#### References

- Avrutin, Sergey, and Maria Babyonyshev. 1997. Obviation in subjunctive clauses and AGR: evidence from Russian. *Natural Language and Linguistic Theory* 15: 229-262.
- Avrutin, Sergey, and Bernhard Rohrbacher. 1997. Null subjects in Russian

- inverted constructions. In *Proceedings of FASL 4*, 32-53. Ann Arbor: Michigan Slavic Publications.
- Bailyn, John. 2013. *The Syntax of Russian*. Cambridge: Cambridge University Press.
- Bianchi, Valentina, and Mara Frascarelli. 2010. Is topic a root phenomenon? *Iberia: An International Journal of Theoretical Linguistics* 2.1: 43-88.
- Camacho, José A. 2013. *Null Subjects*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 2008. On phases. In *Foundational Issues in Linguistic Theory*, ed. by Robert Freidin, Carlos Otero, and Maria Luisa Zubizarreta, 133-166. Cambridge: MIT Press.
- Dyakonova, Marina. 2009. *A Phase-based Approach to Russian Free Word Order*. Ph.D. dissertation, University of Amsterdam. LOT Dissertation Series 230.
- Frascarelli, Mara, and Ronald Hinterhölzl. 2007. Types of topics in German and Italian. In *On Information Structure, Meaning and Form*, ed. by Susanne Winkler and Kerstin Schwabe, 87-116. Amsterdam: John Benjamins.
- Fehrmann, Dorothee, and Uwe Junghanns. 2008. Subjects and scales. In *Scales*, ed. by Marc Richards and Andrej Malchukov, 189-220. Linguistische Arbeitsberichte 86. Leipzig: Universität Leipzig.
- Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. New York: Oxford University Press.
- Gordishevsky, Galina, and Sergey Avrutin. 2003. Subject and object omissions in child Russian. In *Israeli Association for Theoretical Linguistics Conference*, Vol. 19, ed. by Yehuda Falk. Israel: Ben Guiron University of the Negev.
- Gribanova, Vera. 2013. Verb-stranding verb phrase ellipsis and the structure of the Russian verbal complex. *Natural Language and Linguistic Theory* 31: 91-136.
- Holmberg, Anders. 2005. Is there a little *Pro*? Evidence from Finnish. *Linguistic Inquiry* 36: 533-564.
- Holmberg, Anders. 2010. Null subject parameters. In *Parametric Variation: Null Subjects in Minimalist Theory*, eds. Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 88-124. Cambridge: Cambridge University Press.
- Kučerová, Ivona. 2014. The syntax of null subjects. *Syntax* 17.2: 132-167.
- Lindseth, Martina. 1998. *Null-Subject Properties of Slavic Languages: With Special Reference to Russian, Czech and Sorbian*. München:

- Verlag Otto Sagner.
- Livitz, Inna. 2014. *Deriving Silence through Internal Reference: Focus on Pronouns*. Ph.D. dissertation, New York University.
- Madariaga, Nerea. 2014. Diachronic change and the nature of pronominal null subjects: the case of Russian. Ms., University of the Basque Country UPV/EHU.
- McShane, Marjorie. 2009. Subject ellipsis in Russian and Polish. *Studia Linguistica* 63(1):98-132.
- Müller, Gereon. 1988. Zur Analyse subjektloser Konstruktionen in der Rektions-Bindungs-Theorie. Thattrace-Effekte und pro-Lizensierung im Russischen. Ms., Universität Frankfurt am Main.
- Müller, Gereon. 2006. Pro-drop and impoverishment. In *Form, Structure, and Grammar. A Festschrift Presented to Günther Grewendorf on Occasion of his 60th Birthday*, eds. Patrick Brandt and Eric Fuß, 93-115. Berlin: Akademie Verlag.
- Müller, Gereon. 2008. Some consequences of an impoverishment-based approach to morphological richness and pro-drop. In *Elements of Slavic and Germanic Grammars: A Comparative View*, eds. Jacek Witkoś and Gisbert Fanselow, 125-145. Polish Studies in English Language and Linguistics 23. Frankfurt am Main: Peter Lang.
- Perlmutter, David, and John Moore. 2002. Language-internal explanation: The Distribution of Russian Impersonals. *Language* 78: 619-650.
- Pesetsky, David, and Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In *Phrasal and Clausal Architecture*, ed. by Simin Karimi, Vida Samiian, and Wendy K. Wilkins, 262-294. Amsterdam: John Benjamins.
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17: 501-557.
- Růžička, Rudolf. 1986. Funkcionirovanie i klassifikacija pustyx kategorij v russkom literaturnom jazyke. *Zeitschrift für Slavistik* 31: 388-392.
- Sigurðsson, Ármann Halldór. 2011. Condition on argument drop. *Linguistic Inquiry* 42: 267-304.
- Tsedryk, Egor. 2012. Finite control in Russian as movement: subject extraction, feature deficiency and parametric variation. Ms., Saint Mary's University.
- Tsedryk, Egor. 2013. Internal merge of nominative subjects and pro-drop in Russian. In *Proceedings of the 2013 Annual Conference of the Canadian Linguistic Association*, ed. by Shan Luo, University of Victoria.

egor.tsedryk@smu.ca

## Russian Anaphoric Possessive in Context\*

*Ksenia Zanon*  
Indiana University

The paper discusses the distribution of the reflexive possessive *svoj* in Russian. An aggregate of previously unreported facts concerning the behavior of *svoj* in different contexts—with superlatives and various types of quantifiers—is argued to instantiate new evidence for QR in Russian (a controversial issue, given the traditional analysis of Russian as a language lacking covert QR operation, as, e.g., in Ionin 2002, but see Antonyuk-Yudina 2009, Zanon 2014), as well as additional arguments for the contextual approach to phases as in Bošković (2014a, in press).

### 1 Introduction

I take Bošković's recent proposals as a point of departure in the investigation of properties associated with the anaphoric possessive. I am adopting two of his theoretical contributions. First, I assume a contextual approach to phases, whereby the highest projection in the extended domain of a lexical head counts as a phase (Bošković 2014a). Second, I exploit the idea that only the highest edge is extricable in configurations with multiple phasal edges (Specs/adjuncts) (Bošković, in press). What follows is a brief elaboration on each point.

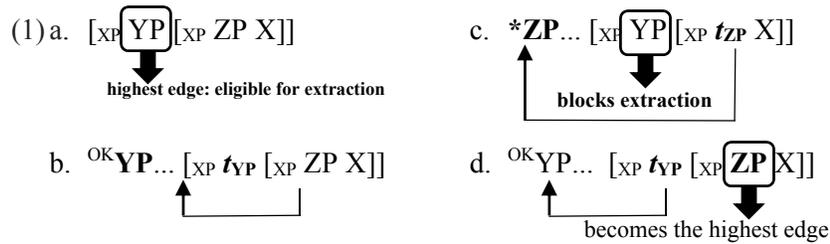
Bošković (2008, 2012, 2014b) argues that languages are subject to parametric variation in the nominal domain, whereby the English-type incarnations project DP above NP, but Russian-type lack this D-layer.

---

\* Many thanks to my indefatigable informants, especially to Natasha Fitzgibbons and Misha Yadroff, as well as the audience, the reviewers, and the editors of *FASL* 23.

Further, Bošković (2014a) shows that the highest projection in the extended domain of every lexical head functions as a phase: it follows then that DP constitutes a phase in English, and NP – in Russian.

Bošković (2014b, in press) also contends that in cases of multiple edges (Specs/adjuncts) of a phase, only the highest one counts as the phasal edge for the Phase Impenetrability Condition (PIC). Abstractly, YP in (1a) is the highest edge. This configuration renders it accessible for extraction (1b). (1c) with movement of ZP over an overt higher edge is illicit. Traces, however, void “edgehood” for the purposes of PIC, much like they rescind the effects of Relativized Minimality (RM) (i.e. traces do not count as interveners for RM, Chomsky 1995): in (1d) the extraction of YP renders ZP the highest edge suitable for subsequent operations.



This technology explains Bošković’s (in press) BCS paradigm in (2): (2a) is equivalent to (1c): the extraction out of AP is blocked by the higher edge *Jovanovog*. In (2b) *ponosnog na tebe* is on the edge, so movement out of it is possible. Once *Jovanovog* moves in (2c), the AP becomes the edge licit for subsequent operations as predicted by (1d).

- (2) a. \*Na tebe<sub>1</sub> sam vidio [<sub>NP</sub> Jovanovog [<sub>NP</sub> [ponosnog t<sub>1</sub>] [<sub>NP</sub> oca]]]  
 of you am seen Jovan’s proud father  
 ‘I saw Jovan’s father (who is) proud of you.’
- b. Na tebe<sub>1</sub> sam vidio [<sub>NP</sub> [ponosnog t<sub>1</sub>] [<sub>NP</sub> oca]]
- c. ?Jovanovog<sub>1</sub> na tebe<sub>2</sub> sam vidio [<sub>NP</sub> t<sub>1</sub> [<sub>NP</sub> [ponosnog t<sub>2</sub>] [<sub>NP</sub> oca]]]

Anaphor binding is conceived in similar terms: the anaphoric reflexive must occupy the outermost edge to be bound outside its minimal phasal XP under the phasal approach to Condition A. This captures Russian (3):

the order Adj>*svoj* in (3b) is disallowed, since the adjective, occupying the highest edge, blocks the binding of *svoj*.<sup>1</sup>

- (3) a. Ona ščitaet pobedu v ètom zabege **svoim**  
 she considers win in this race self's  
personal'nym rekordom.  
 personal record  
 'She considers the victory in this race her personal record.'  
 b. ?\*Ona ščitaet pobedu v ètom zabege personal'nym **svoim**  
 rekordom.

The prediction hence is that *svoj* in Russian ought to occupy the outermost edge in every context. In the ensuing sections I discuss two environments that are superficially problematic for Bošković's approach. Certain quantifiers *must* precede the possessive whereas others *can*. Similarly, the superlatives appear to be freely ordered with respect to *svoj*. The analysis that I develop here does not contradict the claims above; in fact, it provides additional arguments in their favor. I suggest that quantifiers appear at the edge only in the contexts of obligatory QR, using the contexts in question as a test for QR in Russian.

## 2 Quantifiers

In this section I consider the interaction of quantifiers with *svoj*. The novel observation concerns the split between strong and weak quantifiers in this context. The former obligatorily precede the possessive, whereas the latter can appear either before or after *svoj*. Ultimately, I will show that whenever the quantifier surfaces in front of *svoj*, it is subject to QR.

### 2.1 Agreeing Quantifiers: Facts

The most natural order of quantifiers and *svoj* is in (4), where quantifiers of every flavor precede the anaphoric possessive. Observe that this pattern seemingly contradicts the "edgehood" requirement imposed on *svoj* discussed above. Conversely, a permutation of this order yields an unexpected result in (5). The latter demonstrates that strong quantifiers

---

<sup>1</sup> The judgments reported here involve neutral interpretation. If the adjective is focused, (3b) becomes acceptable for the majority of speakers. See Section 3 for details.

(like *vse* ‘all’ and *každyj* ‘each’) are illicit if *svoj* occupies the edge. However, the indefinites (i.e., *-to* and *-nibud*’ series) can freely precede or follow the possessive. The summary of these facts is provided in (6).

- (4) a. Dostoevskij postavil na polku každuju **svoju** knigu.  
 Dostoevskij put on shelf each self's book  
 ‘Dostoevsky put each book of his on the shelf.’  
 b. Dostoevskij postavil na polku vse **svoi** knigi.  
 D put on shelf all self's books  
 c. Dostoevskij postavil na polku kakie-to **svoi** knigi.  
 D put on shelf some self's books  
 d. Opublikoval li D kakuju-nibud' **svoju** knigu?  
 published QD some self's book  
 ‘Did Dostoevsky publish some book of his?’
- (5) a. ?\*Dostoevskij postavil na polku **svoju** každuju knigu.  
 b. ?\*Dostoevskij postavil na polku **svoi** vse knigi.  
 c. Dostoevskij postavil na polku **svoi** kakie-to knigi.  
 d. Ja ne znaju, objavio li Dostoevskij **svoju** kakuju-nibud' knigu.
- (6) a. ?\**svoj* > Q<sub>STRONG</sub>                      c. ✓ *svoj* > Q<sub>INDEF</sub>  
 b. Q<sub>STRONG</sub> > *svoj*                      d. ✓ Q<sub>INDEF</sub> > *svoj*<sup>2</sup>

## 2.2 Agreeing Quantifiers: Analysis

Following Bošković (2012) and Despić (2011), I treat agreeing QPs and the possessive *svoj* as NP-adjuncts. I propose that quantifiers in (4) undergo QR.<sup>3</sup> Since traces do not count in the calculation of edge, the operation of QR ensures that *svoj* is on the outermost edge, which renders it accessible to its binder in LF as desired, as shown in (7a).

Now the question is how to capture the observed split between (5a,b) and (5c,d)? In (5a,b) the possessive is on the edge of the phase, so it blocks the extraction of the QP, as in (7b). It follows that quantifiers of this type are subject to obligatory QR. Indefinites, on the other hand, can be interpreted in-situ as choice-functional elements (Yanovich 2005). The

<sup>2</sup> Some of my informants find (5c,d) as degraded as (5a, b). For these speakers, the analysis I give for (5a, b) holds of instances involving indefinites. Note, these are the same speakers for whom the wide scope interpretation of the indefinites discussed below is impossible.

<sup>3</sup> I am arguing here that only the quantified element is subject to QR. This is exactly the analysis entertained in Chomsky (1993, 1995); see also Ruys (1997) for a semantic implementation of this proposal.



Yanovich (2005) argues that *-to* items in Russian should be likewise analyzed as CFal elements. He reports that (10) is ambiguous between two readings: (10-i) instantiates the meaning where the indefinite gets the widest scope; in (10-ii) it is interpreted inside the conditional.

- (10) Petja budet sčastliv, esli kakaja-to devuška pridet.  
 Petya will.be happy if some girl comes
- (i) <sup>OK</sup>  $\exists >$ if: ‘There is a property  $p_{\langle e,t \rangle}$  such that Petja will be happy if a girl  $y$  such that  $p(y) = 1$  comes.’
- (ii) <sup>OK</sup> if  $> \exists$ : ‘Petja will be happy if there is a girl who comes.’

Yeremina (2012) claims that wide scope is generally more prominent for *-to* indefinites in Russian.

Now consider (11) in light of these claims. (11a) is judged by the majority of my informants as realizing the wide scope interpretation, which is in line with ‘indefinites as choice functions’ account (i.e. we expect the widest scope here in compliance with Yeremina’s observation about indefinites).<sup>4</sup> But in (11b), the most prominent reading is the one where the indefinite takes narrow scope, indicating that it is best treated as a quantificational element (again, in consonance with my analysis).

- (11) a. ? Každyj professor думаet, čto Ivan pročitaet **svoj**  
 each professor thinks that Ivan will.read self’s  
kakoj-to doklad na konferencii.  
 some-TO paper on conference  
 ‘Every professor thinks that Ivan<sub>1</sub> will deliver some paper of his<sub>1</sub> at the conference.’

*Preference for  $\exists > \forall$  (there is a particular paper of Ivan’s such that every professor thinks that Ivan will deliver it at the conference)*

- b. Každyj professor думаet, čto Ivan pročitaet **kakoj-to svoj**  
 doklad na konferencii.

*Preference for  $\forall > \exists$  (each professor thinks that Ivan will deliver a [possibly different] paper of his)*

---

<sup>4</sup> One informant rejects *svoj > -to* order altogether. For this informant the only interpretation of (10b) is the one where the indefinite realizes narrow scope.

The upshot of the discussion is the following: strong quantifiers are subject to QR. Hence, they must be the highest adjuncts in the NP, which ensures that the possessive is on the edge of the phase when binding takes place. Weak Qs (like indefinites) must undergo QR, if they are on the edge of NP preceding the possessor. If they are merged lower than the possessive, they are obligatorily interpreted in-situ as choice functional items.

### 2.3 Genitive of Q Assigning Quantifiers: Facts

The genitive of Q assigning quantifiers evince the same split as agreeing quantifiers: some (*mnogo*, *bol'sinstvo*) must be obligatorily merged after *svoj* is introduced, as in (12), while others (*neskol'ko* and cardinal numerals) have an option of either preceding or following the possessive, as in (13). The descriptive summary is provided in (14).

- (12) a. Sberbank priostanovit bol'sinstvo **svoix** operacij.  
Sberbank will.suspend most self's operations  
'Sberbank will suspend most of its operations.'
- b. \*Sberbank priostanovit **svoix** bol'sinstvo operacij.
- c. On opublikoval mnogo **svoix** fotografij.  
he published many self's photos  
'He<sub>1</sub> published many photos of his<sub>1</sub>.'
- d. \*On opublikoval **svoix** mnogo fotografij.
- (13) a. On postavil na polku neskol'ko **svoix** knjig.  
he put on shelf several self's<sub>GEN</sub> books
- b. On postavil na polku **svoix** neskol'ko knjig. [has more than several]  
self's<sub>GEN</sub>
- c. On postavil na polku **svoi** neskol'ko knjig. [only has several]  
self's<sub>ACC</sub>
- d. Dostoevskij postavil na polku 12 **svoix** knjig.  
D put on shelf 12 self's<sub>GEN</sub> books<sub>GEN</sub>
- e. Dostoevskij postavil na polku **svoix** 12 knjig. [has more than 12]
- f. Dostoevskij postavil na polku **svoi** 12 knjig. [only has 12]
- (14) a. [*t*Q [*svoj*...]] QR, *svoj* is on the edge
- b. \*[*svoj* [Q...]] Qs must undergo QR, *svoj* blocks QR
- c. <sup>OK</sup>[*svoj* [Q...]] Qs are interpreted in-situ

It is easy to capture (14) for agreeing Qs (assuming agreeing Qs are adjuncts). But the constructions involving Gen of Q presumably boast a

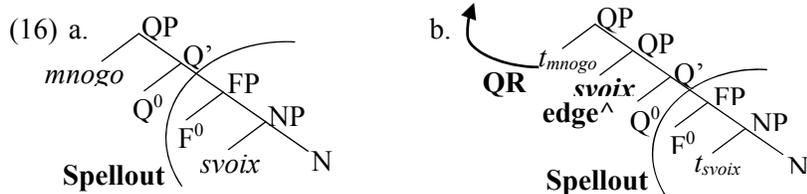
richer internal structure. For the latter I am essentially adopting Bošković's (2014a,b) analysis for the distribution of numerals and *svoj* in BCS and extend it to all Gen of Q-assigning quantifiers.

2.4 Genitive of Q Assigning Quantifiers: Analysis

Bošković (2014a,b) argues that numerals implicate a QP (see also Despić 2011), which constitutes the highest projection within the nominal domain and hence functions as a phase. Contra standard analyses of numerals as Q<sup>0</sup>-elements, he endorses the view that numerals are phrasal elements (adjuncts to QP) on the grounds that they undergo Left Branch Extraction (LBE), an instance of phrasal movement. As (15) demonstrates, *mnogo*, *neskol'ko*, and *bol'sinstvo* also undergo LBE. Genitive case is assigned by F<sup>0</sup>, which intervenes between QP and NP.

- (15) My mnogo/neskol'ko/12/ bol'sinstvo (na-/s-)kupili knjig.  
 we many/several/12/most bought books  
 'We bought many/several/12/most books.'

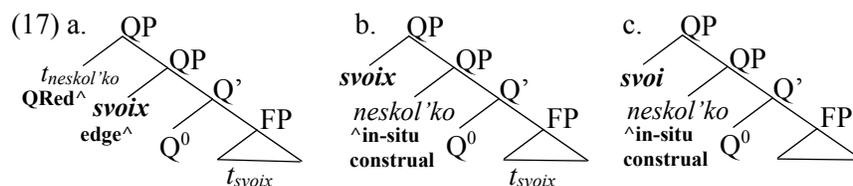
These assumptions produce the structure in (16a). It is problematic, however: since QP is a phase (it is the highest projection within the NP domain), its complement is subject to Spellout, rendering only its edge (*mnogo*) and its head accessible for further computations. But *svoj*, being inside the complement, belongs in the Spellout domain, where it cannot be bound. To avoid this problem Bošković suggests that the anaphor tucks in under the quantifier as in (16b), which precludes it from being spelled out. In this position it can be bound (after QR) by the higher binder, as below.



Now the question is: what motivates this movement of the anaphor to the edge? I suggest that the mechanism akin to Chomsky's (1993) Greed is responsible for this operation: items move to satisfy their own requirements (see also Bošković 2007). Intuitively, the possessive *wants* to be bound, so it moves to the position where it *can* be bound. Suppose

*svoj* enters the derivation with a feature [+anaphoric], which requires that the possessive occupies the edge. Note this is the opposite of the edge feature on the head (e.g., edge feature on  $T^0$ , which drives the movement to Spec TP). Hence, whenever the anaphor can move, it must move (in compliance with restrictions on movement such as (anti-)locality).<sup>5</sup>

The explicit derivations of (13) are provided in (17) below. In (13a=17a) the possessive tucks in under the quantifier. The latter undergoes QR, thus cinching the “edgehood” of *svoj*. In (13b=17b), the extracted anaphor occupies the outermost edge blocking QR and forcing the in-situ construal of *neskol'ko*. Finally in (13c=17c), *svoi* surfaces in the accusative. One can plausibly infer that in this case the possessive is base-generated in the QP-adjoined position (rather than moved there from an FP-internal, genitive-assigning slot). The quantifier in (17c) likewise forgoes QR (due to the blocking effect of the higher edge), requiring the in-situ interpretation instead. Note, incidentally, that (13b)/(13e) and (13c)/(13f) are distinct in terms of meaning: the latter instantiates the reading, whereby the agent is in possession of only several/12 particular books and all of these books are manipulated in some way (dubbed here “definite” reading), while the former obligatorily engenders a subset construal (under which the agent handles only a subset of all the books that he owns). Though a formal semantic account of this dichotomy is orthogonal to this project, one can reasonably conjecture that the two available interpretations correspond to the two distinct derivations. To wit: the definite reading is linked to instances like (17c) with high generation of *svoj*, whereas the partitive meaning corresponds to the NP-internal generation of the possessive in (17b).



An additional argument for the analysis in (17b) involves ellipsis. A number of researchers (e.g., Boeckx 2009, Bošković 2014a, van Craenenbroeck 2010 and references therein) have argued that only phasal

<sup>5</sup> Note that in (3b) the possessive *is* on the edge (just not the highest edge).

complements (and phases) can be elided. So, if *svoj* remains in the NP-internal position (as in the hypothetical (16a)), it cannot survive ellipsis (since it is buried in the complement of the phasal head position). On the other hand, if *svoj* moves to a higher position, as I argue, then it should survive ellipsis. The latter is borne out, as shown in (18b) (cf. (18a)).

- (18) a. \*Maša prodala neskol'ko svoix knig, a Petja  
 Masha sold several self's books and Petja  
 obmenjal mnogo [~~svoix knig~~].  
 exchanged many self's books  
 'Masha sold several of her books and Petya – lots of his.'
- b. Maša prodala neskol'ko svoix knig, a Petja obmenjal mnogo  
 svoix [knig].

The main takeaway point from this section pertains to the similar patterning of quantifiers with respect to *svoj* placement. Despite the need for more technology and a richer structure in the Gen of Q assigning contexts, I propose a blanket analysis that stands for all types of quantifiers: whenever the quantified element precedes the possessive, it is subject to QR. The operation is obligatory for strong quantifiers (irrespective of their agreement properties), so they must always be merged last in the noun phrase to be accessible for extraction. On the other hand, quantifiers that are interpretable in-situ can be introduced into the structure earlier than the possessive.

### 3 Interactions with Focus

At this juncture, I detour briefly into the interaction with focus, returning to the issue mentioned in ft. 1. Consider (3) again: (3b) is degraded in neutral contexts. However, it improves significantly if the adjective is focused. In fact, whenever a focus element (*-to*, *-že*, *daže* 'even', *tol'ko* 'even') is explicitly present, the judgments become opposite to those in (3). To capture (19) and the grammatical incarnation of (3b), I defend essentially the same analysis as above for quantifiers. In fact, following much recent work on the nature of focus, I assimilate the behavior of focused items to that of quantified elements.

- (19) a. On umudrilsja opublikovat' [daže/tol'ko erotičeskie]<sub>F</sub>  
 he managed to.publish even/only erotic  
svoi novelly.  
 self's novellas  
 'He<sub>1</sub> managed to publish even/only his<sub>1</sub> erotic novellas.'
- b. \*... svoi [daže/tol'ko erotičeskie]<sub>F</sub> novelly
- c. (Ved') ona poterjala [novye-TO]<sub>F</sub> svoi krossovki.  
 (Foc) she lost new-FOC self's tennis.shoes  
 '(Imagine,) she lost her NEW tennis shoes.'
- d. \*... svoi [novye-TO]<sub>F</sub> krossovki
- e. [Starye ŽE]<sub>F</sub> svoi proizvedenija on ne predelyvaet.  
 old- Foc self's oeuvres he not redo  
 'As for the old oeuvres of his, he does not redo them.'
- f. \*...svoi [starye ŽE]<sub>F</sub> proizvedenija

There exist a few proposals (Krifka 2004, Wagner 2005, a.o.) that invoke movement analysis for English 'even' and 'only' items. The following is a paradigm from Wagner (2005) showing that 'only' licenses NPIs, but exclusively in the unfocused part. He argues that (20a,c) are derived by moving the focused associate to the complement position of 'only' in LF, as in (20b,d).

- (20) a. John only gave **kale** to any of his friends.  
 b. [only kale] [λx. John gave x to any of his friends]  
 c. John only gave any kale to **his friends**.  
 d. [only his friends] [λx. John gave any kale to x]

This analysis is directly applicable to the Russian cases above. I suggest that focus movement is akin to QR. Whenever the adjective is focused, it has to be ex-situ to be interpretable as a focus-bearing element, whether it is associated with *daže* 'even', *tol'ko* 'even' or a null operator (for arguments that *-to* and *-že* items instantiate "operator-like interpretation of focus" see McCoy (2003)). The only licit derivation for such instances is in (21a), where the adjective is composed after the introduction of *svoj*, since the alternative in (21b) results in the configuration prohibiting the adjective extraction. The treatment explains the pattern in (19): a focalized element must occupy the highest edge in (19a,c,e), as it is subject to subsequent extraction; the order *svoj*>Adj in (19b,d,f) examples precludes

the obligatory focus movement over *svoj*. (21a) likewise underlies the mandatory focus construal of adjectives in the grammatical instances of strings in (3b), which involve a silent focus operator.



If correct, this analysis is indistinguishable from my proposals in Section 2, which yields a conceptual benefit of reducing quantifier movement and adjective focus movement of adjectives to the same operation.<sup>6</sup>

#### 4 Binding: Agree or Movement?

Hornstein (2001) argues that anaphors are a residue of overt movement of the antecedent. Marelj (2011) proposes that the availability of the anaphoric possessive is reducible to the availability of LBE. In the languages that allow for LBE, the antecedent is extricable and its trace is spelled out as *svoj*. I will offer some evidence against this treatment.

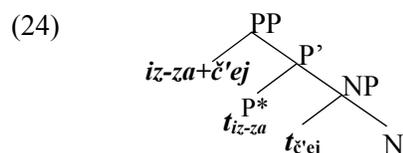
One problematic paradigm for Hornstein is reported in Despić (2011). In (22) I replicate his BCS paradigm for Russian. The anaphoric possessive can be bound inside the PP, as in (22a). However, the overt extraction out of this PP in (22b) is prohibited. The problem is evident: if binding were treated on a par with movement, why is (22a) good, but (22b) is not? In what follows I present a solution to (22), compliant with my analysis of binding.

- (22) a. *Politik<sub>1</sub> ušelv otstavku iz-za svoje<sub>1</sub> suprug<sub>i</sub>.*  
 politician went toresignation because self's spouse  
 'The politician<sub>1</sub> resigned because of his<sub>1</sub> spouse.'
- b. \**Č'ej politik ušelv otstavku iz-za t suprug<sub>i</sub>?*  
 whose politician went toresignation because spouse

<sup>6</sup> Alternatively, one can argue for overt string vacuous focus movement, as in Bošković (2012). For arguments that Slavic languages have a specialized lower focus position see Stjepanović (1995).

(22b) becomes grammatical if the extracted complex includes both the preposition and the wh-word, as in (23). Bošković (2013) handles cases like (23) as in (24). The adjective (*č'ej* in this case) moves to Spec PP. This movement is followed by adjunction of the preposition to the extracted adjective. The resulting complex (*iz-za č'ej*) can then move out of this PP. In (24) the PP constitutes a phase.

(23) *Iz-za č'ej politik supruzi ušel v otstavku?*



As conceived, the derivation in (24) violates a number of constraints. Movements of *č'ej* and *iz-za* are both too short under Bošković's definition of anti-locality.<sup>7</sup> To fix these violations he exploits a modified version of the \*-notation. Traditionally, the \* is assigned to an island if an element crosses it. To salvage such a derivation some rescue operation is required, such as, e.g., PF ellipsis. The paradigm in (25) constitutes the empirical underpinning of this claim. Once the island is deleted, the \* is removed along with it, resulting in the grammatical (25b).

- (25) a. \*Ben will be mad if Abby talks to one of the teachers, but she couldn't remember which (of the teachers) Ben will be mad if she talks to.  
 b. Ben will be mad if Abby talks to one of the teachers, but she couldn't remember which ~~\*[(of the teachers) Ben will be mad if she talks to]~~. (Merchant 2001: 88)

Bošković's innovation is to place the \* on the head of the island. This allows him to explain the instances, in which the islandhood is voided when the head moves, leaving the copy that is deleted in PF. He extends this analysis to the PIC/antilocality violations, arguing that the \* is placed on the head of the phase. So, in (24), it is the preposition that bears the \*.

<sup>7</sup> An element has to cross one full phrase boundary (not just a segment): this requirement rules out, e.g., the instances where the complements move to the Spec of the same phrase.



- (29) ?\*Dostoevskij postavil na polku **svoju** každyju knigu.  
 D put on shelf self's each book
- (30) ?\*Dostoevskij<sub>i</sub> postavil na polku [<sub>t<sub>1</sub></sub> **každyju** knigu].  
 ^Spellout: svoju ^highest edge

It stands to reason, hence, that binding is not established via movement. Rather, it is analogous to an agree-type operation. *Svoj* then is a lexical item rather than the mere spellout of the trace of the antecedent. As such, it hinders QR in cases like (29) and focus movement in (19).

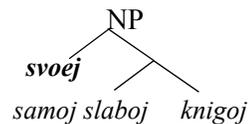
## 5 Superlatives

The facts surrounding the distribution of superlatives with *svoj* in (31) instantiate yet another case that appears to run counter to the predictions in (1). *Svoj* can either precede or follow the superlative (cf. (31a) and (31b)). In principle, (31b) is expected to be bad, since the highest edge is occupied by the superlative complex.

- (31) a. Ja ščitaju *Pridorožnuju travu* **svoej** samojslaboj knjigoj.  
 I consider *Roadside grass* self's most weak book  
 'I consider *Roadside grass* my weakest book.'
- b. Ja ščitaju *Pridorožnuju travu* samoj slaboj **svoej** knjigoj.  
 'I consider *R.g.* the weakest one of my books.'

(31a) is straightforward and fits the analysis developed above. On the assumption that (31a) has the representation in (32), the possessive, appearing on the edge, is bound.

(32)



(31b), on the other hand, seems problematic: the binding of the possessive here is predicted to be impossible, since *svoj* does not occupy the highest edge. The ensuing discussion does not pretend to solve the problem posed

by (31b) in its entirety, rather it catalogs some previously unreported facts and sketches a possible analysis.<sup>8</sup>

One pertinent observation in conjunction with (31) concerns the parallelism with English: (31a) and (31b) are distinct in terms of meaning in the way English (33) is. (33a) is ambiguous, as indicated in (i) and (ii); (33b), on the other hand, can only have the ‘possession’ interpretation. That the meaning (33a-i) is missing for (33b) becomes obvious in contexts with *ever*, as in (34). It is not a new observation that superlatives license NPIs (see, e.g., von Stechow 1999, Gajewski 2014). What is new is the interaction of superlatives, possessives and NPIs reported in (34). The prepositional construction in (34b), in which the possessive is inside the PP, is incompatible with *ever*.

- (33) a. John finished his longest book yesterday.  
       (i) = of all the books he has ever read, the one he  
           finished yesterday was the longest  
       (ii) = of all the books in his possession, the one he  
           finished yesterday was the longest

---

<sup>8</sup> Semantic literature boasts robust research on the nature of superlatives. Below is the paradigmatic example, which became the focus of much recent attention:

- (i) John climbed the highest mountain.  
       (a) **Absolute reading:** John climbed the mountain that is higher than any other mountain (i.e., he climbed Everest).  
       (b) **Comparative reading:** John climbed the mountain higher than anybody else. (i.e., John has conquered Hoosier Hill, provided nobody else in some contextually salient set managed such an ascent).

There are two general approaches that account for the available readings in (i): the in-situ approaches (Sharvit and Stateva 2002, Stateva 2002), and the “movement” approaches (Heim 1999, 2000, Szabolsci 1986). In essence, the “movement” analyses derive the two readings in (i) by using two distinct mechanisms: the comparative reading in (i-b) is captured via movement of the comparative –EST morpheme out of DP, while the absolute reading in (i-a) is argued to arise via an in-situ mechanism.

It is easy to see why I find the spirit of this analysis tempting, given my earlier proposal for the quantifiers. One can speculate that something like the “movement” analysis is applicable to (31): the superlative in (31a) is interpreted in-situ, while the superlative in (31b) undergoes QR, thus “granting” the edge status to the possessive anaphor. To make this work two non-trivial caveats need to be addressed. First, the entire superlative complex, not just the –EST morpheme, must be extracted. Second, (31) is not entirely clear with regard to comparative/absolute readings. My informants seem to converge on the notion that this is not the relevant difference in (31). I leave this alternative at that pending further inquiry.

- b. John finished the longest one of his books yesterday.  
 (i) ≠ of all the books he has ever read, the one he finished yesterday was the longest  
 (ii) = of all the books in his possession, the one he finished yesterday was the longest
- (34) a. John finished his longest book ever yesterday.  
 b. \*John finished the longest one of his books ever yesterday.

The facts in (34) are replicable for Russian. Consider (35) and (36). In both examples the NPIs are only possible if the possessive appears in the initial position. The permutation of this order, whereby the possessive appears closest to the noun, is impossible.

- (35) a. Včera Ivan dočital svoju samuju dlinnuju iz  
 yesterday Ivan finished self's most long of  
 kogda-libo pročitaných (im knig) knigu.  
 ever read by.him books book  
 'Yesterday Ivan finished his longest book ever.'  
 b. \*Včera Ivan dočital samuju dlinnuju iz kogda-libo pročitaných  
 (im knig) svoju knigu.
- (36) a. Ja uže soveršila svoju samuju krupnuju v žizni  
 I already realized self's most big in life  
 pokupku.  
 purchase  
 'I already made my biggest purchase in life.'  
 b. \*Ja uže soveršila samuju krupnuju v žizni svoju pokupku.

Based on this similarity of patterning between the prepositional construction and the superlative+*svoj* order, I propose the structure in (37) for (31b). I submit that in (37) the null  $N_2$  is responsible for supplying the partitive meaning (much like overt English counterpart *one* in (33b)).

- (37)
- 
- ```

  graph TD
    NP2 --> S["samoj slaboj+svoej"]
    NP2 --> N_prime["N'"]
    N_prime --> N2_star["N2*"]
    N_prime --> NP1["NP1"]
    N2_star --> Empty["∅"]
    N2_star --> t_svoej["t_svoej"]
    NP1 --> N1["N1"]
  
```

In (37) *svoj* moves to form a constituent with the superlative complex in the higher NP (driven by [+anaphoric] feature). The noun adjoins to the resulting complex. The movements violate anti-locality, so the \* is placed on  $N_2$ , which is subsequently removed in PF. In this configuration the anaphor is on the edge.

There are two pieces of evidence in support of this analysis. First, superlative + possessive form a prosodic unit, hence the contrast in (38).

- (38) a. ???Ivan opublikoval #samuju skandal'nuju #svoju# knigu.  
           Ivan published       most       scandalous       self's book  
       b. Ivan opublikoval #samuju skandal'nuju svoju#knigu.

The second argument comes from overt extraction in (39). The best example here is in (39d), predictably so given (37): the entire superlative + possessive complex can move, which rules out all the other cases in (39).

- (39) a. \*Samuju Ivan opublikoval skandal'nuju svoju knigu.  
           most       Ivan published       scandalous       self's book  
       b. ?\*Samuju skandal'nuju Ivan opublikoval svoju knigu.  
       c. ?\*Svoju knigu Ivan opublikoval samuju skandal'nuju.  
       d. Samuju skandal'nuju svoju Ivan opublikoval knigu.

## 6 Conclusion

The paper defends the position that phases are established contextually with only the highest edge counting as the edge for the purposes of PIC. It also provides novel evidence for QR in Russian. In the position where they should block binding, this operation ensures that the possessive anaphor is on the edge, hence accessible to the binder. Some quantifiers are subject to obligatory QR (these Qs always precede the possessive). Those quantifiers that *can* be interpreted in-situ (e.g., indefinites) may follow the possessive anaphor. In this case, they receive a particular interpretation, consistent with the indefinites-as-choice functions analysis. If the indefinites precede the possessives, however, they are construed as quantificational elements. Some focus operations are QR-like in nature (and hence are subject to the same treatments as Qs). In certain contexts (with Gen of Q-assigning quantifiers, inside PPs and with superlatives), the feature [+anaphoric] drives the movement of the anaphor to the edge.

I also provided some evidence against treating anaphoric possessives as residue of overt movement of the antecedent.

## References

- Antonyuk-Yudina, Svitlana. 2009. Long-distance scrambling, VP ellipsis, and scope economy in Russian. In *Proceedings of PLC 32*: 1-9.
- Boeckx, Cedric. 2009. On the locus of asymmetry in UG. *Catalan Journal of Linguistics* 8: 41-53.
- Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An Even More Minimal Theory. *Linguistic Inquiry* 38(4): 589-644.
- Bošković, Željko. 2008. What will you have, DP or NP? In *Proceedings of NELS 37*: 101-114.
- Bošković, Željko. 2012. On NPs and clauses. In *Discourse and Grammar: From Sentence Types to Lexical Categories*, 179-242.
- Bošković, Željko. 2013. Traces do not head islands: what can PF deletion rescue? Ms., U. of Connecticut.
- Bošković, Željko. 2014a. Now I'm a phase, now I'm not a phase: on the variability of phases with Extraction and Ellipsis. *Linguistic Inquiry* 45(1): 27-89.
- Bošković, Željko. 2014b. More on the edge of the edge. In *Proceedings of FASL 22*. Ann Arbor: Michigan Slavic Publications: 44-66.
- Bošković, Željko. In press. Getting Really Edgy: On the edge of the edge. *Linguistic Inquiry*.
- Chomsky, Noam. 1993. A minimalist program for linguistic theory. In *The view from Building 20*, 1-52. Cambridge: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Despić, Miloje. 2011. *Syntax in the Absence of Determiner Phrase*. Ph.D. diss., U of Connecticut.
- Gajewski, Jon. 2014. *Superlatives, NPIs and most*. Ms., U of Connecticut.
- Heim, Irene. 1999. *Notes on Superlatives*. Ms., MIT.
- Heim, Irene. 2000. Degree operators and scope. In *Proceedings of SALT 10*, 40-64.
- Hornstein, Norbert. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.

- Ionin, Tanya. 2002. The one girl who kissed every boy: scope, scrambling and discourse function in Russian. In *ConSole X*: 65-80.
- Kratzer, Angelika. 1998. Scope or pseudo-scope? Are there wide-choice indefinites? In *Events and Grammar*. Dordrecht: Kluwer. 163-196.
- Krifka, Manfred. 2004. *Association with Focus Phrases*. Ms., Humboldt U.
- Marelj, Marijana. 2011. Bound variable and left-branch extraction. *Syntax* 14(3):205-229.
- McCoy, Svetlana. 2003. Connecting information structure and discourse structure through "kontrast". *Journal of Logic, Language and Information* 12:319-335.
- Merchant, Jason. 2001. *The Syntax of Silence*. Oxford: Oxford U Press.
- Reinhart, Tanya. 1997. Quantifier scope: how labor is divided between QR and choice functions. *Linguistics and Philosophy* 20:335-397.
- Ruys, E. G. 1997. On quantifier raising and "Move-F." In *Utrecht Institute of Linguistics OTS yearbook 1997*, 83-104. Utrecht University.
- Sharvit, Yael, and Penka Stateva. 2002. Superlative expressions, context, and focus. *Linguistics and Philosophy* 25:453-504.
- Stateva, Penka. 2002. *How Different are Different Degree Constructions?* Ph.D. diss., U of Connecticut.
- Stjepanović, Sandra. 1995. *Movement of Wh-Phrases in Serbo-Croatian Matrix Clauses*. Ms., U of Connecticut.
- Szabolcsi, Anna. 1986. Comparative superlatives. In *MIT Working Papers in Linguistics*, vol. 8. Cambridge: MIT Press. 245-266.
- van Craenenbroeck, Jeroen. 2010. *The Syntax of Ellipsis: Evidence from Dutch Dialects*. New York: Oxford U Press.
- von Stechow, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2):97-148.
- Wagner, Michael. 2005. NPI-licensing and focus movement. In *SALT 15*: 275-293.
- Yanovich, Igor. 2005. Choice-functional series of indefinite pronouns and Hamblin series. In *SALT 15*: 309-326.
- Yeremina, Olga. 2012. *The Semantics of Russian Indefinite Pronouns*. Ph.D. diss., Michigan State U.
- Zanon, Ksenia. 2014. Two Russian hybrids. *Journal of Slavic Linguistics* 22(2):215-258.

## **Scalar Implicatures of Russian Verbs**\*

*Yulia Zinova*

Heinrich-Heine University, Düsseldorf

*Hana Filip*

Heinrich-Heine University, Düsseldorf

In this paper we address the common claim that perfective verbs presuppose the initial phase (or a process part) of events denoted by them, and assert their final phase (or a culmination part), while the meaning of imperfective verbs lacks both components. Different formulations of this claim have been proposed by Padučeva (1996, 2011) and Romanova (2006) for Russian, and by Dočekal and Kučerová (2009) for Czech, among others. We argue that what is regarded as a matter of presupposition in the semantic structure of Russian perfective verbs is best analyzed in terms of scalar implicature in the negated contexts and entailment in the affirmative sentences. The main evidence for our analysis is based on some recent work in the presupposition projection theories; of particular interest is Chemla's (2009) experimental study.

### **1 The Main Idea**

According to the proposals by Padučeva (1996, 2011), Romanova (2006) and Dočekal and Kučerová (2009), the semantic structure of (1) consists of two components: (i) a process part of an event of reading, which is

---

\* We would like to thank the audience, organizers, and the anonymous reviewers of FASL 23. The research reported here was supported by the German Science Foundation (DFG) as part of the SFB 991.

taken to be presupposed, followed by (ii) a culmination at a point when the whole book has been read, which is taken to be asserted.

- (1) Ivan pročital ètu knigu.  
 Ivan <sub>PREF.read</sub><sub>PAST.SG.M</sub> this book  
 ‘Ivan read this book completely through.’

The presuppositional nature of the process component of perfective verbs is viewed as being confirmed by the observation that it is preserved under negation and in questions, as shown in (2a) and (2b), respectively:

- (2) a. Ivan ne pročital<sup>PF</sup> ètu knigu.  
 Ivan <sub>NEG</sub> <sub>PREF.read</sub><sub>PAST.SG.M</sub> this book  
 ‘Ivan did not read this book completely through.’  
*Inference:* Ivan started reading/read a part of this book.  
*Assertion:* Ivan did not finish reading this book.
- b. Ivan pročital<sup>PF</sup> ètu knigu?  
 Ivan <sub>PREF.read</sub><sub>PAST.SG.M</sub> this book  
 ‘Has/Did Ivan read this book completely through?’  
*Inference:* Ivan started reading/ read a part of this book.  
*Question:* The speaker asks the addressee to confirm or deny whether Ivan finished reading this book.

In (2a), what is negated is the culmination, but not the process (initial) part of described events, i.e., (2a) can be felicitously uttered in a situation in which it is known that Ivan started reading the book. In (2b), the speaker takes it for granted that Ivan started reading the book, and what is questioned is whether he finished it. To the extent that the previous studies rely on the negation and question tests, it is fair to assume that what they have in mind is a semantic presupposition.

In this paper, we argue that the existential inference in question is not a matter of semantic presupposition, contrary to most analyses, but instead, a matter of scalar implicature in negative contexts (2a) and in questions (2b), and an entailment in affirmative sentences (1). We will provide empirical tests allowing us to tease apart presuppositions, entailments and (scalar) implicatures associated with Russian perfective verbs. The tests are based on the recent research in the domain of

projective content (Chemla 2009; Romoli 2011; Schlenker 2008, among others).

The paper is organized as follows. In Section 2 we present several variants of a presuppositional analysis of the inferences associated with perfective verbs and point out their weaknesses. In Section 3 we apply standard tests for semantic and pragmatic presuppositions to Russian verbs, and introduce Grønn's (2004, 2006) idea that the inference in question is a pragmatic implicature. In Section 4 we discuss the results of the recent experiment by Chemla (2009) and the questionnaire study we have done on the basis of those results. The empirical data obtained from the questionnaire is then used to advocate the scalar implicature analysis of the inferences associated with perfective verbs.

## 2 Presuppositional Analyses of Slavic Perfective Verbs

### 2.1 Russian Linguistic Tradition

In the Russian linguistic tradition, the idea that perfective verbs have a bipartite structure can be traced back to Maslov (1984). On his view, Russian perfective verbs consist of an 'eventive' part (*sobytijnyj komponent*) and a 'stative / resultative' part (*statal'nyj komponent*).

Building on Maslov (1984), Padučeva (1996, 2011) proposes that these two components of perfective verbs differ in their communicative status. What roughly corresponds to Maslov's 'eventive' component is presupposed and concerns backgrounded information. On her view, it comprises not only the process part of events described by perfective verbs, but also their preparatory conditions and various associated pragmatic factors like intentions, expectations and obligations associated with the utterance of sentences headed by perfective verbs. The second, asserted, component regards focused information, including the 'reaching of a/the boundary', i.e., the final phase of events involving goals, results, and limits of various sorts. Padučeva (1996) illustrates these points with the following contrast, among others:

- (3) a. Taksi vyzyvali<sup>IPF?</sup> [=(1)] Padučeva 1996  
 Taxi call<sub>PAST.PL</sub>  
 'Did you call a cab?'

- b. Vy vyzvali<sup>PF</sup> taksi?  
 you<sub>PL</sub> call<sub>PAST.PL</sub> taxi  
 ‘Did you call a cab?’

*Presupposition:* The hearer was expected/required to call a cab.

(3a), which is headed by an imperfective verb, is a neutral question about whether a cab was called. (3b), which is headed by a perfective verb, in addition, strongly suggests that from the point of view of the speaker the addressee was required or obliged to call a cab.

What is important for the purposes of this paper is that Padučeva (1996, p. 54) also claims that “the first [i.e., presupposed, backgrounded, YZ&HF] component does not fall within the scope of negation.” In evoking a standard test for a semantic presupposition, she implicitly suggests that ‘the first [meaning] component’ of perfective verbs is, on her view, akin to a semantic presupposition, even if she does not use this term.

Although Padučeva (1996) adduces a number of valid and subtle intuitions in support of her approach to the uses of perfective verbs, as opposed to imperfective ones, its major weakness is that it fails to separate between the semantic meaning components of perfective verbs, on the one hand, and various speech act related pragmatic inferences (such as speaker’s deontic and normative expectations on the addressee) associated with utterances of sentences with perfective verbs, on the other hand.

The second problem, and one that is also mentioned in Grønn (2004), is that the observed speaker-oriented modality inferences are not consistently attached to all the uses of sentences with perfective verbs. For instance, as Grønn (2004) observes, they are not associated with the utterances of affirmative perfective sentences. Take, for example, (4), which is an affirmative correspondent of (3b), but unlike (3b) it does not suggest (under the most neutral circumstances) that the referent of *you* was required or obliged to call a cab:

- (4) Vy vyzvali<sup>PF</sup> taksi.  
 you call<sub>PAST.PL</sub> taxi  
 ‘You called a cab.’

Padučeva (1996, p. 56) also observes that there is no reason to assume that the utterance of (4) triggers the inference of an “expectation component” (“komponent ožidanija”) on the part of the speaker, but she does not motivate this observation any further. That is, Padučeva (1996) is aware of the fact that not all (utterances of) sentences with perfective verbs carry the relevant inference (or “presupposition” in her wide sense), but she does not provide any account when it may, must or must not be present in sentences with perfective verbs.

## 2.2 *Contemporary Syntactic Approaches to the Decomposition of Perfective Verbs*

Following Padučeva (1996), Romanova (2006) proposes that “perfective verbs must have a complex semantic structure, where one part is asserted, the other is presupposed” (p.29). She adopts Padučeva’s (1996) characterization of the presupposed part, but has a different understanding of the asserted component.

Most importantly, according to Romanova (2006), “it is not true that only resultative verbs or the verbs with ‘reaching-the-boundary’ component, can bear the presupposition of perfectives” (p. 29); rather, all perfectives are “words that encode decomposable structures (informational, semantic and therefore syntactic)” (ibid., p. 53). For example, even the class of inceptive verbs like those with the prefix *za-* like *zapet* ‘to begin to sing’ which fail to entail culmination or result, limit of some sort (under the most usual understanding) are taken to have a complex semantic structure, whereby the first part is presupposed. (5) (example (64a) in Romanova 2006, p.29), for instance, asserts that Tonja did not start to sing and presupposes that Tonja was expected to sing her song, according to Romanova (2006).

- (5) Tonja ne zapela<sup>PF</sup> svoju pesnju.  
 T. not INCEP.Sing<sub>PAST.SG.F</sub> self<sup>SF.ACC</sup> song<sub>ACC</sub>  
 ‘Tonja didn’t start to sing her song (contrary to the expectation).’

To give another example, (6) ((65) in Romanova 2006, p. 30) is claimed to be associated with a “presupposition” that the addressee was supposed to buy bread:

- (6) Ty            kupila<sup>PF</sup>            xleb?  
 You<sub>SG.NOM</sub> bought<sub>PAST.SG.F</sub> bread<sub>ACC</sub>  
 ‘Did you buy bread?’

*Presupposition:* You were supposed to buy bread.

This move then allows her to assimilate the semantics of perfective verbs as a whole class to accomplishments which are commonly assumed to have a bipartite structure. Romanova (2006) follows a syntactic approach of Ramchand (2006), according to which accomplishments are syntactic structures that consist of two separate projections, namely *process* (ProcP) and *result* (resP) corresponding to the presuppositional and assertive components of the meaning of perfective verbs, respectively.

There are three main problems with Romanova’s (2006) account. First, the meaning of perfective verbs as a whole class cannot be assimilated to that of accomplishments (for counterarguments see Filip 2000, Filip and Rothstein 2005). Obviously, there are perfective verbs that cannot be meaningfully decomposed into two subevents, a process and a result subevent. One good example is the class of semelfactive verbs with the suffix *-nu-* in Russian: e.g., *prygnut* ‘to jump’.

Second, what remains entirely unclear is the representation of speaker and/or addressee oriented attitudes in terms of syntactic structures. For instance, the syntactic representation of the alleged ‘contrary to the expectation’ (5) and obligation (6) inference that is supposed to be associated with the *process* (ProcP) part of the syntactic structure of perfective verbs remains on a pre-theoretic level.

Third, it is easy to show that the alleged presuppositional meaning components (here, the expectation of the speaker on the addressee or on some participant of the situation described by perfective sentences) are not tied to the uses of perfective verbs only, which is a point of criticism that also applies to Padučeva’s (1996) proposal. Compare (5) with (7). The main difference between them is in their main verbs: (5) is headed by a perfective verb, while (7) by its corresponding imperfective simplex. Also (7), and not only (5), triggers the inference that Tonja was expected to sing her song.

- (7) Tonja ne pela<sup>IPF</sup> svoju pesnju.  
 T. not sing<sub>PAST.F.SG</sub> self<sup>SF.ACC</sup> song<sub>ACC</sub>  
 ‘Tonja wasn’t singing/didn’t sing her song.’

Romanova’s (2006) account also inherits the problems that we observed with Padučeva’s (1996) proposal: namely, first, the failure to distinguish between semantic components of perfective verbs and pragmatic factors having to do with obligations, expectations and the like on the part of the interlocutors, and second, the fact that the alleged presuppositions of perfective verbs fail to be present in all their uses, most notably in utterances of affirmative sentences.

### 2.3 *Event Semantics*

One illustrative example of an event semantics approach is Dočekal and Kučerová (2009). As is widely assumed, they take it for granted that all perfective verbs have a uniform meaning of telic predicates, drawing on Czech and Russian data. Telic predicates are equated with accomplishment predicates, which means that they are decomposed into two subevents, where  $e_1$  is a process and  $e_2$  is the result state (mainly following Giorgi and Pianesi 2001). Their main innovation is the claim that perfective verbs carry the ‘activity presupposition’ (‘process’ in our terms) tied to  $e_1$  or ‘the first homogeneous part of telic events’. The evidence for this claim comes from the observation that it exhibits the usual projective properties of a semantic presupposition: namely, it ‘projects under negation and under a question operator.’

One immediate problem with this account is that the meaning of perfective verbs as a whole class cannot be equated with that of accomplishments (see also above the criticism of Romanova’s (2006) account).

Another problem is the one that Dočekal and Kučerová (2009) noticed themselves: namely, imperfective verbs can also carry the ‘activity presupposition’. A case in point is the class of secondary imperfective verbs (explicitly marked with the imperfective suffix *-yva-*) that are formed with the ‘completive’ (or ‘terminative’) prefix *do-*, as in (8a). The sentence (8a) denies that Vasya was about to finish reading the book yesterday, and implies that he read a part of it, but was nowhere near being close to finishing reading it. But notice that the same

inference – namely that Vasya read a part of the book – is also triggered by the sentence with the corresponding perfective verb (8b):

- (8) a. Včera Vasya ne dočityval<sup>IPF</sup> tu knigu.  
 Yesterday Vasya not COMP.read<sub>IMP.PAST.SG.M</sub> that book  
 ‘Yesterday Vasya was not finishing reading that book.’  
*Inference:* He started reading that book.
- b. Včera Vasya ne dočital<sup>PF</sup> tu knigu  
 Yesterday Vasya not COMP.read<sub>PAST.SG.M</sub> that book  
 ‘Yesterday Vasya did not finish reading that book.’  
*Inference:* He started reading that book.

Dočekal and Kučerová (2009) acknowledge that terminative (uses of prefixes like *do-*, when used to form secondary imperfective verbs, are problematic for their account, because secondary imperfectives with such prefixes can also trigger the ‘activity presupposition’ just like perfective verbs. They set this problem aside for future research.

#### 2.4 Summary and Questions

First, all the works summarized here share the claim that all and only perfective verbs can be decomposed into two parts, effectively have the bipartite structure of accomplishments. In this bipartite structure, the first part, ‘process’ or ‘activity’, is presupposed while the second, ‘result’, part is asserted. However, there is a number of perfective verbs that do not have the structure of accomplishments, i.e., that cannot be plausibly decomposed into a process and a result component (see Filip 2000, Filip and Rothstein 2005, and references therein).

Second, the studies of perfective verbs, especially those conducted in the Russian tradition (here represented by Padučeva 2006 and Romanova 2006), often contain claims about the association of perfective verbs with certain speaker-oriented modalities; particularly prominent are speaker’s normative and deontic expectations on the addressee. Such speech act related factors clearly lie outside of the lexical semantic structure of perfective verbs (which is not to deny that they may arise from the interaction of the lexical meaning of perfective verbs with pragmatic factors). This raises the question about the distribution and robustness of such pragmatic inferences that are allegedly associated with the uses/meaning of perfective verbs.

Third, despite frequent claims about the ‘presupposition’ of perfective verbs, there seems to be little reflection on the status of such claims, and if any concrete empirical evidence is adduced at all, it is their preservation under negation and in questions. However, not all that projects is a presupposition (see e.g., Chierchia and McConnell-Ginet 1990, Beaver 2001, Potts 2005), so further tests must be applied in order to establish the nature of the inferences associated with perfective verbs. This is the main question of the current paper.

### 3 Probing Perfectives: Presupposition or Implicature?

#### 3.1 *Presupposition?*

3.1.1 Evidence against Semantic Presupposition. Projection from embeddings, negation and antecedents of conditionals, is standardly used as a diagnostic test for a semantic presupposition. Let us consider the examples (9) and (10). In both cases, the inference of the affirmative sentences (9a) and (10a) survives under negation in (9b) and (10b), and hence would qualify as a presupposition:

- (9) a. John won the marathon.  
 b. John didn’t win the marathon.  
*Inference:* John participated in the marathon.
- (10) a. John read all the books.  
 b. John didn’t read all the books.  
*Inference:* John read some of the books.

However, the inferences in question do not always project a conditional out of the antecedent:

- (11) a. If John won the marathon, he will celebrate tonight.  
*Inference:* John participated in the marathon.  
 b. If John read all the books, he will pass the exam.  
 ⇒ John read some of the books.

This difference is used to distinguish the inferences of (9) and (10): the projected component of (9a) is a semantic presupposition and the projected component of (10a) is a scalar implicature.

Now let us turn to Russian sentences with perfective verbs that denote accomplishments. As (12) shows, the alleged ‘process presupposition’, which is claimed to be triggered by perfective verbs, does not project out of the antecedent of a conditional, and hence it fails to exhibit one of the properties of semantic presupposition.

- (12) Esli Vasja pročital<sup>PF</sup> učebnik, on sdast<sup>PF</sup> èkzamen.  
 if Vasja PREF.readPAST.SG.M textbook, he PASSPRES.3SG exam  
 ‘If Vasja completely read the textbook, he will pass the exam.’  
 → Vasja read/began reading at least a part of the textbook.

It may also be observed that the alleged ‘process presupposition’ of sentences with perfective verbs (denoting accomplishments) is also easily defeasible. This speaks against its presuppositional nature too, on the assumption that a semantic presupposition is generally non-cancellable.<sup>1</sup> For instance, the discourse in (13) is felicitous even though the first sentence (equivalent to (2a) given at the outset) is followed by a second sentence that denies its alleged presupposition, namely, ‘Ivan started reading the book.’

- (13) Ivan ne pročital<sup>PF</sup> ètu knigu. On daže ne otkryl eë.  
 Ivan NEG PREF.read this book he even NEG open it<sub>ACC.F</sub>  
 ‘Ivan didn’t read this book. He did not even open it.’

3.1.2 Evidence against Pragmatic Presupposition. Theories of pragmatic presuppositions regard those as requirements on the common ground (see e.g., Heim 1983; Karttunen 1973; Stalnaker 1973; Shanon 1976). One good test for pragmatic presupposition is known as “Hey, wait a minute!” test, which builds on Shanon’s (1976, p. 248) observation: “[u]pon uttering S, a speaker P pragmatically presupposes Q if it is suitable for the hearer to utter ‘One moment, I did not know that Q’ in response to S.”

Using this test can show easily that the alleged ‘process presupposition’ of Russian sentences with perfective verbs that denote

---

<sup>1</sup> The non-cancellability of semantic presuppositions is less reliable than projection tests.

accomplishments does not qualify as pragmatic presupposition. Consider the examples in (14):

- (14) a. Katya pročitala<sup>PF</sup> skazki Puškina.  
 Katya<sub>PREF</sub>.read<sub>PAST.SG.F</sub> fairy tales Pushkin<sub>GEN</sub>  
 ‘Katya read the fairy tales by Pushkin completely through.’
- b. #Pogodi-ka! Ja ne znal, što ona ix čitala<sup>IPF</sup>!  
 wait! I<sub>NEG</sub> knew that she them read  
 ‘Wait a minute! I didn’t know that she was reading them!’
- c. Pogodi-ka! Ja ne znal, što ona umeet čitat’!  
 wait I<sub>NEG</sub> knew that she can read  
 ‘Wait a minute! I didn’t know that she can read!’

(14a) is headed by the perfective verb *pročitala* ‘she read completely (through)’. If pronounced with neutral intonation, it would be odd to follow it with (14b) that indicates the hearer’s surprise about the alleged ‘process presupposition’. In contrast, (14a) can be followed by (14c) which indicates that the ability of Katya to read is a pragmatic presupposition of (14a).

3.1.3 Summary. In this section, we used standard presuppositional tests to show that the ‘process presupposition’ that is claimed to be triggered by sentences with perfective verbs denoting accomplishments is not a matter of semantic or pragmatic presupposition.

### 3.2 Pragmatic Implicature

As Grønn (2004, p. 61) points out, “[t]he negation test in itself is not a sufficient argument for associating perfective accomplishments with a presupposition [of the existence of their process part, YZ&HF].” Instead, he proposes to treat it as a matter of pragmatic strengthening by the Gricean maxim of quantity (Grice 1975). He relies on speaker’s and hearer’s economy effort in communication that he takes to be related to “the Gricean idea that the best form-meaning pairs are the ones which minimize both the speaker’s and hearer’s effort (whose interests are, in a sense, conflicting)” (Grønn 2006, p. 71). He also assumes the markedness theory of Slavic aspect, according to which the perfective member of the aspectual opposition is marked, while the imperfective

member is semantically unmarked, i.e., unspecified with respect to the distinguishing semantic feature of Perfective.

Under negation, what we observe is aspectual competition: namely, when the existence of a whole (culminated) event is to be denied, the use of an unmarked imperfective, as in (15), is the default choice of the speaker:

- (15) Ivan ne čital<sup>IPF</sup> ètu knigu.  
 Ivan NEG read<sub>PAST.SG.M</sub> this book  
 ‘Ivan did not read this book.’  
*Interpretation:* denial of the existence of a whole event.

If the speaker uses an utterance with the marked perfective verb, as in (16) (which is equivalent to (2a) given at the outset), the hearer infers that there was some attempt or activity on the part of the Agent which did not culminate because it would have been more economic for the speaker to use a sentence with an unmarked imperfective, if it were possible/relevant:

- (16) Ivan ne pročital<sup>PF</sup> ètu knigu.  
 Ivan NEG PREF.read<sub>PAST.SG.M</sub> this book  
 ‘Ivan did not read this book completely through.’

Based on such data and observations, Grønn (2004, 2006) suggests that the alleged presupposition of perfective verbs is best seen in terms of an implicature, rather than in terms of a presupposition. Grønn’s (2004, 2006) suggestion points in the right direction. In what follows, we propose that the existential inference associated with the process part of perfective verbs that denote accomplishments is a matter of scalar implicature.

#### 4 Proposal: Scalar Implicature

##### 4.1 Background: Projection Theories

In developing our approach to the analysis of the semantics of Russian perfective verb, recent findings in the research on presupposition projection are of particular importance. Building on the presupposition projection theories (e.g., Heim 1983; Schlenker 2008, and references

therein), Chemla (2009) provides experimental evidence that distinguishes the projection properties of presuppositions from those of scalar implicatures.

Among his most relevant insights is the following one: If a sentence  $S$  (e.g., (17a)) with the presupposition  $P(x)$  (17b) is embedded under universal quantifiers *every/each* or *no* (as in (17c) and (17d)) the presupposition of the whole sentence is universal:  $\forall x:P(x)$ , (17e). Hence, the presupposition is the same in sentences with a universal affirmation (*every/each*, (17c)) or a universal negation (*no*, (17d)).

- (17) a. The student knows that he is lucky.  
 b. The student is lucky.  
 c. Each student knows that he is lucky. [= (4) in Chemla (2009)]  
 d. No student knows that he is lucky. [= (8) in Chemla (2009)]  
 e. Each student is lucky.

This property does not hold for scalar implicatures: if a sentence  $S$  (18a) entails that  $I(x)$  (20b), then  $S$  embedded under *every/each* (18c) entails that  $\forall x:I(x)$  (universal inference, (18d)) and  $S$  embedded under *no* (18e) implicates that  $\exists x:I(x)$  (existential inference, (18f)).

- (18) a. John read all books. [= (13) in Chemla (2009)]  
 b. John read some of the books.  
 c. Each student read all the books. [= (14) in Chemla (2009)]  
 d. Each student read some of the books.  
 e. No student read all the books. [= (18) in Chemla (2009)]  
 f. Some student read some books.

The universal inference like the one in (18d) in the universal assertion context such as (18c) is a trivial property of entailments. The existential inference (18f) in the universal negation context such as (18e) follows from the Gricean maxims and the construction of alternatives. Let us illustrate this point with a simple example. First, recall how scalar implicatures that involve a scalar item (e.g., *all*) in a downward entailing context (here negation) are derived (following suggestions in Grice 1975; Ducrot 1969; Horn 1972, among others).

- (19) a. John didn't read all the books. [= (12) in Chemla (2009)]  
 b. *Alternative*: John didn't read any of the books.  
 c. *Scalar implicature*: John read some of the books.

Sentences with *all* (19a) and *any* (19b) belong to an implicational scale that consists of a set of linguistic alternatives of the same grammatical category which can be arranged in a linear order by degree of informativeness or logical (semantic) strength. Sentence in (19b) is an alternative to (19a), whereby (19b) is logically stronger than (19a). If the speaker does not use (19b), the most natural assumption on the part of the hearer is to conclude that the alternative sentence (19b) is false. The negation of (19b), "it is not the case that John didn't read any of the books" or "John read some of the books," is then an indirect scalar implicature (19c) of (19a) (the two negations cancel each other out).

Similar reasoning works for deriving an implicature (20c) from (20a). The sentence (20b) is an alternative to (20a). As this alternative is informationally stronger, but was not uttered, it gets negated, giving rise to the scalar implicature in (20c).

- (20) a. No student read all the books. [= (18) in Chemla (2009)]  
 b. *Alternative*: No student read any book.  
 c. *Scalar implicature*: (At least) one student read some of the books.

#### 4.2 *Empirical Evidence: Questionnaire*

If the results reported in Chemla (2009) are correct, then embedding sentences that contain inferences of unknown nature under negative universal quantifiers can be seen as a test for distinguishing between presuppositions and scalar implicatures. The reasoning is then as follows, put in the simplest terms: if the inference is universal, the embedded sentence contains a presupposition trigger; if the inference is existential, the embedded sentence involves a scalar implicature. To illustrate how this test can be applied to Russian data consider (21):

- (21) a. Nikto iz nas ne pročital učebnik.  
 nobody of us NEG PREF.read<sub>PAST,SG,M</sub> textbook  
 ‘None of us read the textbook.’
- b. *Alternative*: None of us read any part of the textbook.
- c. *Scalar implicature*: Some of us read/started reading at least a part of the textbook.
- d. *Presupposition*: All of us read/started reading at least a part of the textbook.

The inference in (21c) is existential and arises as the negation of the stronger alternative (21b) to (21a). If only this inference is attested, the sentence (21a) contains a scalar item that triggers an implicature. If, on the other hand, the inference (21d) is attested,<sup>2</sup> (21a) must contain a presupposition trigger.

To test which inferences native speakers of Russian get, we ran a simple questionnaire. Similarly to the experimental design by Chemla (2009), we provided participants with two sentences and asked them to judge if the first one suggested (*predpolagaet* in Russian instructions) the second one. We also asked to assume that the first sentence was uttered by a reliable, honest and well-informed speaker (*nadežnyj, iskrennij i informirovannyj sobesednik* in Russian) in order to establish a natural context in which the Gricean maxims could be applied, which was a necessary condition for the derivation of scalar implicatures.

For the test material, we had sentences of three different types. The first group of sentences were sentences like (21a) that were designed to test the type of inference associated with perfective accomplishments. They were constructed by means of embedding Russian sentences that contained perfective accomplishments under negative universal quantifiers (analogous to examples like (12) and (18) from Chemla (2009)). Apart from (21), another example of such sentence is (22).

---

<sup>2</sup> Note that in this case, in fact, both (21c) and (21d) must hold, as (21c) is a weaker statement than (21d).

- (22) a. Nikto iz moix studentov ne dočital statju.  
 nobody of my students NEG COMP.read<sub>PAST.SG.M</sub> article  
 ‘None of my students finished reading the article.’  
*Tested inferences:*
- b. Vse studenty načali čitat’ statju.  
 All student.PL start<sub>PAST.PL</sub> read article  
 ‘All students started reading the article.’
- c. Kto-to načal čitat’ statju.  
 Somebody start<sub>PAST.SG.M</sub> read article  
 ‘Somebody started reading the article.’

The second group of sentences included perfective sentences denoting accomplishments that contained negation but no quantifier. They were intended to explore if/when native speakers of Russian would report inferences concerning the process component and/or speech-act related speaker-oriented modalities like his/her normative and deontic expectations on the addressee. Some representative examples are given below:

- (23) a. Vasja ne sdelal domašnee zadanie.  
 Vasja NEG PREF.do<sub>PST.SG.M</sub> homework  
 ‘Vasja didn’t do his homework.’  
*Tested inferences:*
- b. Vasja načinal delat’ domašnee zadanie.<sup>3</sup>  
 Vasja start<sub>PAST.SG.M</sub> do homework  
 ‘Vasja started doing the homework.’
- c. Vasja dolžen byl sdelat’ domašnee zadanie.  
 Vasja obliged be<sub>PAST.SG.M</sub> do homework  
 ‘Vasja had to do the homework.’
- (24) a. Vasja ne dodelal domašnee zadanie.  
 Vasja NEG COMP.do<sub>PST.SG.M</sub> homework  
 ‘Vasja didn’t do his homework.’

---

<sup>3</sup> An imperfective verb *načinal* ‘started’ is used here as the more neutral one in comparison with the perfective variant *načal* ‘has started’ that tends (in the absence of a temporal adverbial) to denote an event in the recent past.

*Tested inferences:*

- b. Vasja načinal delat' domašnee zadanie.  
 Vasja start<sub>PAST.SG.M</sub> do homework  
 'Vasja started doing the homework.'
- c. Vasja dolžen byl sdelat' domašnee zadanie.  
 Vasja obliged be<sub>PAST.SG.M</sub> do homework  
 'Vasja had to do the homework.'

The last group included control sentences with presupposition triggers like 'know' and possessive pronouns. One illustrative example is the following one:

- (25) Petja ne znaet, čto Katja včera xodila v kino.  
 Petja NEG know<sub>PRES.3SG</sub> that Katja yesterday went in cinema  
 'Petja does not know that Katja went to cinema yesterday.'

*Tested inference:*

Katja včera xodila v kino.  
 Katja yesterday go<sub>PAST.SG.F</sub> in cinema  
 'Katja went to cinema yesterday.'

We collected answers from 100 native speakers of Russian, using the free version of Survey Monkey (surveymonkey.com) questionnaire platform. The questionnaire design differed from that of Chemla (2009) with respect to possible answers. Anticipating the difficulty of some sentences and inferences, we allowed not only two variants "yes" and "no", but also the weaker versions "probably yes" and "probably no". The answers then were assigned numerical values (1 for "no", 2 for "probably no", 3 for "probably yes" and 4 for "yes") and the mean values were calculated. Control sentences received the rating of 3.61.

Our results strongly suggest that the inferences in question do not have the properties of presupposition. We observed a significant difference in the acceptance rates of existential and universal inferences when the target sentence involved the universal negation. In this case, the universal inferences (e.g., 'all of us at least started reading the textbook', as in (21d) and (22b)) were strongly dispreferred (rating 1.65), while the existential inferences (i.e., 'some of us started reading the textbook', as in (21c) and (22c)) were accepted (rating 3.11). Such behavior, according

to the results of Chemla (2009), corresponds to that of scalar implicatures and not presuppositions.

As far as the question about the presence of speech-act related speaker oriented modalities is concerned, which are emphasized by Padučeva (1996, 2011) and Romanova (2006), participants highly rated (3.16 overall rating) the relevant proposed inference, of the type given in (23b) above. This indicates that their observations are empirically valid. It is an open question how exactly they should be motivated based on independently motivated generalizations concerning the functioning of the Russian aspectual system and its interactions with speech-act factors.

In contrast, we did not find sufficient empirical evidence for the alleged semantic process presupposition, which plays a role in the analysis of perfectivity in Padučeva (1996, 2011), Romanova (2006) as well as in Dočekal and Kučerová (2009). Inferences of the type given in (23a) seem to be dispreferred (rating 1.39), with one notable exception: namely, sentences headed by perfective verbs that contain the completive prefix *do-*. For such sentences, an inference concerning the process component of denoted accomplishments (see (24a)) was rated high (3.39). However, this result is clearly tied to completive prefix *do-*, rather than to perfective aspect of verbs in general.

## 5 Conclusion

In this paper we have shown that the projection properties of Russian perfective verbs in downward entailing contexts (under the universal quantifier *no*) indicate that the projected inference concerning the ‘process’ part of perfective accomplishments is a scalar implicature, rather than a presupposition, contrary to common analyses of Russian perfective verbs. Although our main data come from Russian, the methodology developed here is extendable to other Slavic languages.

## References

- Beaver, David I. 2001. *Presupposition and Assertion in Dynamic Semantics*, Vol. 29. CSLI Publications Stanford.
- Chemla, Emmanuel. 2009. Presuppositions of Quantified Sentences: Experimental Data. *Natural Language Semantics*, 17(4), 299–340.

- Chierchia, Gennaro and McConnell-Ginet, Sally. 1990. *Meaning and Grammar. An Introduction to Semantics*. MIT Press.
- Dočekal, Mojmír and Kučerová, Ivona. 2009. Bound Ability Readings of Imperfective Verbs: A Case for Presupposition. Submitted.
- Ducrot, Oswald. 1969. Présupposés et sous-entendus. *Langue française*, (4), 30–43.
- Filip, Hana. 2000. The Quantization Puzzle. In *Events as Grammatical Objects*, ed. Carol Tenny and James Pustejovsky, 39–96. CSLI Publications, Stanford.
- Filip, Hana, and Susan Rothstein. 2005. In *Formal Approaches to Slavic Linguistics 14 (The Princeton Meeting)*, ed. J. Lavine, S. Franks, M. Tasseva-Kurktchieva and H. Filip. 139–156. Michigan Slavic Publications, Ann Arbor.
- Giorgi, Alessandra and Pianesi, Fabio. 2001. Ways of Terminating. In *Semantic Interfaces: Reference, Anaphora and Aspect*, ed. Maria Teresa Guasti, 211–277. CSLI publications, Stanford.
- Grice, H. Paul. 1975. Logic and Conversation. In *Syntax and Semantics 3: Speech acts*, ed. P. Cole and J. L. Morgan, 41–58. Academic Press, New York.
- Grønn, Atle. 2004. *The Semantics and Pragmatics of the Russian Factual Imperfective*. Ph.D. thesis, Acta Humaniora 199, Oslo.
- Grønn, Atle. 2006. Information Structure and Aspectual Competition. In *Proceedings of the Ninth Symposium on Logic and Language*, 70–77.
- Heim, Irene. 1983. On the Projection Problem for Presuppositions. In *Formal Semantics – the Essential Readings*, 249–260.
- Horn, Laurence R. 1972. *On the Semantic Properties of Logical Operators in English*. Ph.D. thesis, UCLA.
- Karttunen, Lauri. 1973. Presuppositions of Compound Sentences. *Linguistic Inquiry*, (4), 167–193.
- Maslov, Jurij S. 1984. *Očerki po aspektologii [Aspectology studies]*. Izd-vo Leningradskogo universiteta.
- Padučeva, Elena V. 1996. *Semantičeskie issledovanija: Semantika vremeni i vida v ruskom jazyke; Semantika narrative [Semantic studies: The semantics of tense and aspect in Russian; The semantics of narrative]*. Škola “Jazyki Russoj Kultury”, Moscow.
- Padučeva, Elena. V. 2011. Presupposition [Prezumpcija]. In *A project of Corpora-Based Description of Russian Grammar [Proekt korpusnogo opisanija russoj grammatiki]*, rusgram.ru.

- Potts, Christopher. 2005. *The Logic of Conventional Implicatures*. Oxford University Press.
- Ramchand, Gillian. 2006. Verb Meaning and the Lexicon: A First Phase Syntax. Manuscript. <http://ling.auf.net/lingBuzz/000307>. Published as Ramchand (2010).
- Ramchand, Gillian. 2010. *Verb Meaning and the Lexicon*, Vol. 116. Cambridge University Press.
- Romanova, Eugenia. 2006. *Constructing Perfectivity in Russian*. Ph.D. thesis, University of Tromsø.
- Romoli, Jacopo. 2011. The Presuppositions of Soft Triggers Aren't Presuppositions. In *Proceedings of SALT*, Vol. 21, 236–256.
- Schlenker, Philippe. 2008. Be Articulate: A Pragmatic Theory of Presupposition Projection. *Theoretical Linguistics*, 34(3), 157–212.
- Shanon, Benny. (1976). On the Two Kinds of Presuppositions in Natural Language. *Foundations of Language*, 247–249.
- Stalnaker, Robert. 1973. Presuppositions. *Journal of Philosophical Logic*, 2(4), 447–457.

zinova@phil.uni-duesseldorf.de  
hana.filip@gmail.com