Prepositional Repercussions in Russian: Pronouns, Comparatives and Ellipsis

Thesis submitted in partial fulfillment
of the requirements for the degree of
"DOCTOR OF PHILOSOPHY"

by

Tatiana Philippova

Submitted to the Senate of Ben-Gurion University
of the Negev

April 24, 2018

Beer-Sheva

This work was carried out under the supervision of

Prof. Idan Landau

Prof. Nomi Erteschik-Shir

In the Department of _Foreign Literatures and Linguistics_____

Faculty of__Humanities and Social Sciences_____

Research-Student's Affidavit when Submitting the Doctoral Thesis for Judgment

I___Tatiana Philippova_____, whose signature appears below, hereby declare that:

_✓__ I have written this Thesis by myself, except for the help and guidance offered by my Thesis Advisors.

_✓__ The scientific materials included in this Thesis are products of my own research, culled from the period during which I was a research student.

___ This Thesis incorporates research materials produced in cooperation with others, excluding the technical help commonly received during experimental work. Therefore, I am attaching another affidavit stating the contributions made by myself and the other participants in this research, which has been approved by them and submitted with their approval.

Date: __22.04.2018_____ Student's name: _Tatiana Philippova____

Signature___  _____

Acknowledgments

Introduction. I love reading acknowledgements sections: they are the most heartfelt part of the dissertation never to be skipped. Now it is finally my turn to write one. Most people to be mentioned here of course know how indebted I am to them, so my purpose here is to provide a fun read for someone who is as much into acknowledgments sections as I am. Enjoy!

Movement I. First and foremost, I thank my wonderful advisors, Nomi Erteschik-Shir and Idan Landau. I have been working with Nomi for more than six years, as she was my MA supervisor as well. Besides being an excellent linguist, Nomi is a truly amazing person who finds the time and energy for everything and everyone. I have been privileged to have had her unwavering and unconditional support during the ups and downs of the research life. Nomi always believed in me even when I didn't believe in myself (hello, imposter syndrome!), which gave me a very powerful boost. Dear Nomi, I can't thank you enough and will just say that I look forward to continuing working with you (now as co-authors, I hope!).

Idan, I have been so privileged to work under your supervision. You are one of the most amazing and enigmatic linguists I know. Your way of thinking and looking at research problems is very different from mine, which I come to appreciate more and more, as it gives a whole new perspective at things. Your meticulous reading of my (and very possibly everyone else's) writings is something you find rarely in the digital era we live in. The detailed comments you gave me over the years improved both my analyses and the quality of texts.

Dear Nomi and Idan, I wouldn't have wished for a better pair of advisors!

Movement II. I thank all of my language teachers, in particular my English teachers, Irina Yurjevna Rumyantseva and Irina Germanovna Maltseva, my first French teacher Marina Georgievna Miroshnichenko, my first Japanese teacher Natalya Grigorjevna Rumak and my Hebrew ulpan teacher Berry Shalit. I would have never been as fascinated by language as I am now, were it not for your genuine and thorough interest in your subject and exceptional teaching talent. You nurtured the budding linguist in

me. My parents Tamara and Anatoly deserve a special mention here as they always encouraged my passion for languages and helped find many of the amazing language teachers that I studied with. I also thank them for letting me find my own path, rather than pushing me in the directions that were familiar to them. Your unconditional love and support was essential to my reaching this stage. I am also very much indebted to Stepan Narbeevich Matevosyan, who supported me throughout my doctoral studies and always believed in my success.

Movement III. My research life would not have been as enjoyable without my teachers, colleagues and friends at Ben-Gurion University. Thank you, Betsy Ritter, Olga Kagan, Tova Rapoport, Dorit Ben-Shalom, Aviya Hacohen, Arik Cohen, Lavi Wolf, Irena Botwinik, Roey Gafter, Sharon Taube, Irena Lawi, Kalman Sutker, Keren Flash and Zarina Levy-Forsythe. A special thank you goes to Prof. Emeritus Yishai Tobin, whose teaching had a profound effect on my professional views. I am also indebted to our amazing Head of the department Prof. Eitan Bar-Yosef and to our administrator Merav Saroosi-Peretz for all their help and advice. I have benefited a lot from undergraduate students I taught and thank them too. Beyond linguists, I would like to thank Yigal Godler and Michael N. Fried, who nurtured my interest in the history and philosophy of science.

Movement IV. I have been lucky to have had a chance to interact with many amazing junior and senior linguists in the broader Israeli community and beyond. Thank you, Elitzur Bar-Asher Siegal, Nora Boneh, Ivy Sichel, Edit Doron, Julia Horvath, Evan Gary Cohen, Svetlana Dachkovsky, Gabi Danon, Galit Weidman Sassoon, Lior Laks, Bridget Schvarcz, Tamar Lan, Odelia Ahdout, Nadav Matalon, David Pesetsky, Norvin Richards, Itai Bassi, Jason Merchant, Ming Xiang, Chris Kennedy, Karlos Arregi, Kyle Johnson, Rajesh Bhatt, Peter Alrenga, Keir Moulton, Andrew Nevins, Boban Arsenijević, Nerea Madariaga, Javier Fernández-Sánchez, Ora Matushansky, Gunlög Josefsson, Philipp Weisser, Andreas Blümel, Léa Nash, Ekaterina Lyutikova, Natalia Slioussar, Sergei Tatevosov and many, many others. Pavel Rudnev, David Erschler, Andrew Murphy, Matthew Reeve, Zsófi Schön – thank you for your scholarship and friendship!

Coda a.k.a. אחרון אחרון חביב

It will be no exaggeration to say that this dissertation owes its existence to my partner Boris Zaltzman. He was the first person to see a linguist in me and it was his remarkable vision that put me on the academic path. Since then he has been extremely supportive in every aspect of my work (and life) and helped me endure all the dire aspects of creative work. Equally important is that Boris rejoiced at least as much as I did in those rare moments of happiness that only this kind of work can bring. Thank you, Boris, for your love, for being there for me and, most importantly, for being what you are. לגמרי זכיתי!

In cherished memory of my grandfather, Sergei A. Shesterikov

TABLE OF CONTENTS

Abstract

The present dissertation deals with Russian morphosyntactic phenomena involving adpositions. Much of this dissertation is devoted to a study of the distribution of the so-called *pripredložnye* 'prepositional' pronouns in Russian. In previous research their seemingly chaotic distribution proved to be hard to account for. Especially puzzling was the fact that beyond prepositional object environments these pronouns, which I refer to as *n*-forms, appear in (and only in) the genitive standard DP of the Russian phrasal comparative, in which no preposition appears.

I propose a uniform analysis of the distribution of *n*-forms in Russian, which maintains the intuition that they are 'prepositional', i.e. occur in P-object positions. However, I refine the generalization by saying that *n*-forms are only licensed in the complement of P-heads. I thus make a crucial distinction between a vaguely defined group of prepositions and lexemes that are syntactically P-heads. The idea then is that only those 'prepositions' that have the status of a P-head will license an *n*-form in its complement position. In the main text I examine a large group of 'prepositions' in detail, providing arguments that only some of them are P-heads. The 'comparative puzzle' is solved by proposing that the phrasal comparative in Russian involves a null P-head that licenses the genitive *n*-form in the standard DP.

The study proposes that lexemes labelled prepositions in traditional grammars of Russian actually split into three classes. Class I lexemes are P-heads, the only 'true' prepositions in a sense. Class II lexemes are structurally ambiguous between N-heads or P-N combinations (with the complement being an argument of the nominal head), and a lexicalized P-head composed of syntactically inactive nominal (and prepositional) morphemes. Class III lexemes, on the other hand, are never P-heads: they are either P-N combinations, transitive adverbs/adjectives (A-heads) or gerunds (V-heads).

Let me note here that as far as I can see, the arguments for treating 'prepositions' of Class III as non-prepositions, and lexemes of Class II as ambiguous between prepositions on the one hand and nouns or prepositional phrases (active preposition-noun combinations) on the other are theoretically neutral and may be adopted to explain *n*-form distribution in various linguistic frameworks, which might make this thesis interesting to a rather wide audience of readers. However, maintaining the same

type of analysis to derive the *n*-forms in the phrasal comparative while avoiding the claim that those comparatives are actually prepositions (P-heads) is more demanding: it requires a theoretical framework that is abstract enough to allow positing syntactically active but phonetically null elements of various categories, here prepositions (P-heads). The existence of independent evidence for the null P analysis of the phrasal comparative and the explanatory power of the resulting uniform treatment of *n*-forms justify this abstract construct, which in turn suggests that theories allowing such constructs may well be on the right track.

The part of the dissertation devoted to the analysis of 'prepositional' *n*-forms has implications for the typology and morphosyntax of pronouns, Case theory and the typology of case. It may also be instructive for further studies aiming at a classification of prepositions in various languages. In addition, the classification of 'prepositions' developed here may prove relevant for understanding the non-homogeneous behavior of different prepositions in particular environments.

Another part of this dissertation delves more deeply into the Russian phrasal comparative, probing the structure behind the genitive DP-standard. Employing a variety of tests, I show that the standard of comparison DP and the DP it is contrasted with (known as the correlate DP) should belong in the same clause. This fact favors the simple, Direct Analysis of the standard DP, under which it does not involve additional (silent) structure. Next, I offer a new generalization, the Oblique Correlate Constraint, which imposes morphosyntactic restrictions on the genitive standard of comparison. I show that the constraint quite straightforwardly falls out from a more complex, Reduced Clause Analysis, positing abstract structure behind the standard. Looking at other languages, I suggest that the presence of such a constraint in their phrasal comparative may signal that they are to be analyzed as reduced clauses. This part of the thesis thus adds to the growing body of literature investigating the structure of phrasal comparatives across languages and contributes to the notorious debate on whether phrasal comparatives should be analyzed directly (the 'what-you-see-is-what-you-get' approach) or whether they have a clausal structure underlyingly (the ellipsis approach).

Keywords: syntax, morphology, Russian, adpositions, pronouns, case, comparatives, ellipsis

1. Introduction

The present dissertation deals with Russian morphosyntactic phenomena involving adpositions. I propose a uniform solution for the appearance of adpositional pronouns in the complement of prepositions and comparatives, a problem which has been elusive thus far. Additionally, the dissertation bears on the notorious debate concerning the internal structure of phrasal comparatives, in particular, on whether they have a simple DP-structure or a reduced clausal structure. The aim of this enterprise is two-fold: to contribute to the classification debate on the nature of adpositions and to build a solid basis for the investigation of phenomena involving prepositions, such as P-stranding and P-omission under sluicing, that I have touched upon in my previous work. The empirical basis of the current work comprises traditional informal acceptability judgments, commonly used in theoretical syntactic research, as well as natural data from the Russian National Corpus and acceptability judgment surveys.

Adpositions are items that typically express a location or direction in time and space, like the English in, on, to, and from. They comprise prepositions and postpositions. Prepositions appear before their complement noun phrase which specifies the location or direction, while postpositions follow their complement noun phrase. Typically, a language has either prepositions or postpositions, which correlates with their branching type. Right-branching languages in which the complement follows the selecting head usually have prepositions (e.g. English, Russian), whereas left-branching languages typically have postpositions. However, there are certain adposition-like lexical items in English and Russian that look like postpositions. For instance, the Russian nazad (lit: 'onto behind') patterns with its English counterpart ago, as can be seen from (1).

(1) Oni vstretilis' mnogo let **nazad.**
They met many years ago
'They met each other many years **ago**.'

Spustja (lit. 'bringing down') and *later* are another such pair, as evident from (2). However, unlike *later*, the Russian *spustja* turns out to be an *ambi*position (a term adopted from Hagège 2010), that is, one that can either precede or follow its complement; compare (2) and (3). complement; compare (2) and (3).

(2) Oni vstretilis' liš' mnogo let **spustja.**
They met only many years later
'They only met (again) many years **later**.'

(3) Oni vstretilis' liš' **spustja** mnogo let.
They met only later many years
'*They only met (again) **later** many years.'

Besides *ago* and *later* English has *notwithstanding* and *aside* that can be classified as postpositions, with *notwithstanding* available as either a postposition or a preposition. In Russian, *nazad* is probably the only strict postpositional adposition, but there is a considerable number of ambipositions, exemplified in (4-5).

(4) Ona posmotrela {emu **vsled** | **vsled** emu}.
She looked he.DAT in-step; in-step he.DAT
'She stared after him (watched him go).'

(5) On pobežal {ej **navstreču** | **navstreču** ej}.
He ran she.DAT on-meeting on-meeting she.DAT
'He was running toward her.'

A curious property of these ambivalent adpositions is that they are morphologically complex words, comprising a basic preposition *v(o)* 'in(to)' or *na* 'on(to)' and a nominal form. That is to say, these lexical items do not belong to the basic inventory of Russian adpositions. This is in fact also true of *nazad* 'ago' and *spustja* 'later' deriving from a P-N adverbial and a gerund respectively. The English postpositions just examined are not a part of the basic adposition inventory either: *ago* < *agone* (obsolete participle); *later* < *later* (comparative adverb); *notwithstanding* < *not withstanding* (negation + gerund); *aside* < *a(way)* + *side*.

In fact, it is not only quirky adpositions (postpositions and ambivalent adpositions) that behave unlike prepositions of the basic vocabulary. In Russian there is a vast number of morphologically complex lexemes of transparent origin that have been classified as prepositions but are quite distinct from the primary prepositions like *v* 'in', *na* 'on', *s* 'with', *k* 'to'. The fundamental question that I address in this

dissertation is whether all so-called adpositions belong to the same syntactic category as the original adpositions.

Traditional grammars typically refer to adpositions as a functional part-of-speech encoding some relation (e.g. Švedova 1980). Generative linguists focusing on the syntax of constructions involving prepositions, for example, preposition stranding (e.g. Abels 2003) and preposition stranding (omission) under sluicing (e.g. Merchant 2001), do not always explicitly delineate the set of prepositions they are considering, possibly relying on the common understanding of what constitutes a preposition. This is somewhat problematic, since lexemes traditionally labeled prepositions do not form a natural class. For instance, it has been shown that not all adpositions disallow P-stranding (omission) under sluicing in the non-P-stranding languages like Spanish, Brazilian Portuguese (Rodrigues et al 2009), Polish (Nykiel 2013), Russian (Philippova 2014).

At the same time, those generative linguists that are interested in classification issues, have observed that prepositions do not constitute a homogeneous class: some of them seem to behave as purely functional elements while others resemble lexical categories, apparently by virtue of being etymologically derived from words belonging to the lexical 'parts-of-speech', such as nouns, verbs and adjectives. This led linguists to propose a functional vs. (semi)lexical or simple vs. complex distinction within the group of prepositions in various languages. This literature includes Zwarts (1997), van Riemsdijk (1990, 1998), Yadroff and Franks (1999). Later work on the topic, while acknowledging the diversity within the category P, maintains that prepositions constitute a purely functional category (Grimshaw 2000, Baker 2003, Botwinik-Rotem 2004). Another line of research develops a complex structure for the adpositional phrase, comprising a number of projecting heads of different categories. Early work developing this approach includes Koopman (2000) and Svenonius (2003), while more recent studies of this kind are collected in Cinque and Rizzi (2010). In this family of analyses, the non-homogeneity of adpositions is captured by positing varying complexity of the phrase they project.

Notably, the two strands of research, i.e. the one focusing on constructions involving Ps and the one aimed at classifying Ps, rarely intersect. Thus, the former does not systematically discuss how the choice of a preposition affects the grammaticality of the investigated construction. In this dissertation I provide a detailed study of Russian lexical items traditionally labelled 'prepositions', arguing that not all

of them constitute P-heads, or, more generally, items of the same syntactic category. The results of my research are thus relevant to both the classification issues and to the investigation of constructions involving prepositions.

I examine the group of Russian prepositions, a large part of which are the lexemes considered prepositions in the Academic Grammar of Russian (Švedova 1980). I maintain that Russian prepositions do not constitute a homogeneous group, which is conceptually in line with Yadroff and Franks (1999), who propose to split prepositions into two groups, functional and lexical, based on a number of properties that distinguish them.

Yadroff and Franks point out that some of their diagnostics are more reliable than others, meaning that the 'reliable' diagnostics clearly split prepositions into two groups. One of the properties that they consider an *un*reliable diagnostic is the requirement by P of special 'prepositional' forms of 3rd person pronominal complements, which I refer to here as *n-forms* (exemplified in 6, 8), as opposed to (7, 9), exemplifying the default form, referred to here as *j-forms*.

(6) Maša poslala **k** *nej* Petju
Masha sent to she.DAT Petya.ACC
'Masha sent Petya to her.'

(7) *Maša poslala **k** *ej* Petju
Masha sent to she.DAT Petya.ACC
'Masha sent Petya to her.'

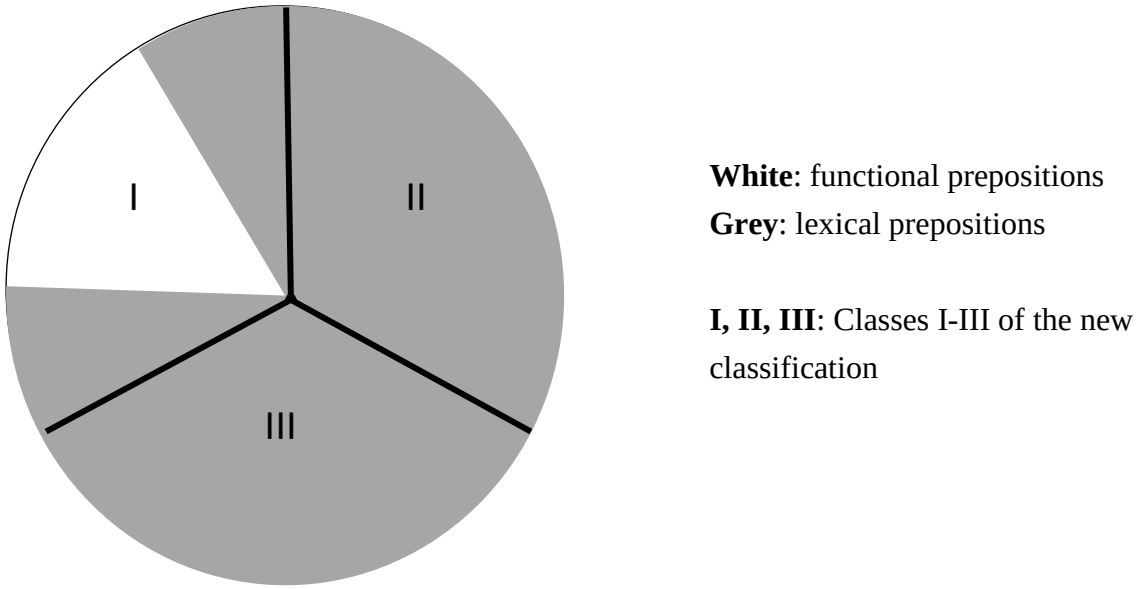
(8) *Maša poslala *nej* cvety
Masha sent she.DAT flowers.ACC
'Masha sent her flowers.'

(9) Maša poslala *ej* cvety
Masha sent she.DAT flowers.ACC
'Masha sent her flowers.'

Much of this dissertation is devoted to a study of the distribution of 'prepositional' pronouns in Russian. I show that prepositions split into three, rather than two classes with respect to taking these special forms. Thus, this property is not a useful diagnostic for classifying prepositions as functional or lexical, and the distribution of 'prepositional' pronouns requires a different account.

Let me now discuss the three classes of prepositions distinguished here and how they correspond to the functional/lexical prepositions. I argue that Class I lexemes are simple P-heads, the only 'true' prepositions in a sense. They constitute a more or less closed class (like other, uncontroversially functional categories, such as particles and coordinators) and properly contain the set of Yadroff and Franks' functional prepositions. Class II lexemes are structurally ambiguous between N-heads or P-N combinations (with the complement being an argument of the nominal head), and a lexicalized P-head composed of syntactically inactive nominal (and prepositional) morphemes: [P [N DP]]/ [N DP] vs. [$_{PP}$ P$_{P\text{-}N}$ DP] Class III lexemes, on the other hand, are never P-heads: they are either P-N combinations, constituting PPs with an adverbial function, transitive adverbs (or rather, short-form predicative adjectives) or gerunds (V-heads). All Class II and III lexemes, as well as a subset of Class I lexemes belong to lexical prepositions, according to the diagnostics specified in Yadroff and Franks (1999). The correspondence between the old binary classification and the new ternary classification just described is also schematically presented in Figure 1.

Figure 1. The binary vs. the ternary classifications of Russian prepositions



The key property underlying the proposed classification of prepositions is their behavior with respect to the N-factor (a term adopted from Hill 1977), i.e. their in/ability to license the so-called *pripredložnye* 'prepositional' pronouns *(n-forms*) on their internal arguments. Class I prepositions require n-forms, Class II prepositions

allow *n*-forms on a par with *n*-less forms and Class III prepositions disallow *n*-forms, as shown in (10)-(12) respectively.

(10) Maša bežala ***k*** {✓nemu; *emu}
Masha ran to 3MSG.DAT[+n] 3MSG.DAT [-n]
'Masha was running to him.'

(11) Maša bežala ***pozadi*** {✓neë; ✓eë}
Masha ran behind 3FSG.GEN[+n] 3FSG.GEN[-n]
'Masha was running behind her.'

(12) Maša bežala ***navstreču*** {*nim; ✓im}
Masha ran on-meeting 3PL.DAT[+n] 3PL.DAT [-n]
'Masha was running to meet them halfway.'

I show that the N-factor elucidates the existence of the three classes of adpositions. Furthermore, the ability of 'lexical' prepositions to license *n*-forms correlates with their case assignment properties, as well as their ability to be postposed and stranded. The classification of prepositions is revisited based on these clusters of properties.

To explain the distribution of *n*-forms, I analyze them as syntactically licensed by P-heads, overt or null. I propose that the initial [n] on these pronouns has evolved from a person agreement marker associated with P into a P-case marker on the pronoun. It follows from the account that if a 3rd person pronominal DP may not take an *n*-form in a particular environment, it is not a complement of a P-head in this environment. If the analysis is on the right track, it follows that none of the oblique case-marked DPs that can be pronominalized but cannot surface as *n*-initial pronouns are licensed by P-heads. This implies that the property of licensing oblique case cannot be restricted to P-heads. Thus, analyses that posit null P-heads to account for oblique case-marking on bare nominals (e.g. Pesetsky 2013 for Russian) would not be directly compatible with the current proposal.

N-forms, i.e. the 'prepositional' forms of pronouns, also do not quite fit into the existing pronominal typologies, such as those proposed in Cardinaletti and Starke (1999) and Déchaine and Wiltschko (2002) respectively. This is because they look at more general aspects of pronominal syntax and semantics and the prepositional/non-prepositional distinction is orthogonal to the distinctions they are concerned with. The existence of the peculiar, 'prepositional', series of pronouns in Russian, as well as

Ukrainian and West Slavic, shows that these formal typologies of pronouns are not exhaustive. This is to be taken into account when approaching 'ill-behaved' pronouns in various languages.

The dissertation is organized as follows. **Chapters 2-4** concern the distribution of the adpositional and regular forms of third person pronouns, which leads to building a new classification of Russian prepositions. **Chapter 2** lays the background for this study, first discussing the origin of the n-initial forms and the way different Slavic languages have accommodated these pronouns. In one type of language, represented by Russian, Ukrainian and West Slavic languages, n-forms are distinguished from others as adpositional forms and I suggest that there is a functional reason for the language to make this distinction, which is related to the three-way ambiguity of GEN/ACC 3rd person pronominal forms. Next, previous literature dealing with n-initial forms is reviewed and finally, an overall picture of the distribution of n-forms is presented, based on a detailed corpus study.

A licensing mechanism for *n*-forms and *j*-forms in the prepositional complement is proposed in **Chapter 3.** It also provides more details about particular prepositions and divides them into three classes, which derives the distributional facts.

The morphosyntactic status of n-forms and j-forms is further clarified in **Chapter 4.** It explores the issue of what form 3rd person pronouns take when appearing in a coordinated complement of a preposition. The data comes from a corpus study and a large-scale acceptability judgment survey.

Chapter 5 deals with "the comparative puzzle". A curious fact about n-forms is that they are only found in one environment besides the prepositional complement, namely on the standard of comparison in the Russian genitive of comparison. As the reader will see from the review of the literature, this has been a long-standing problem. I propose a solution that is conceptually in line with the solution for prepositions.

In **Chapter 6** I continue the investigation of phrasal comparatives by probing into the internal syntax behind the genitive DP-standard. Employing a variety of tests, I show that the standard of comparison DP and the DP it is contrasted with (known as the correlate DP) should belong in the same clause. This fact favors the simple, Direct Analysis of the standard DP, under which it does not involve additional (silent) structure. Next, I offer a new generalization, the Oblique Correlate Constraint, which imposes morphosyntactic restrictions on the genitive standard of comparison. I show that the constraint quite straightforwardly falls out from a more complex, Reduced

Clause Analysis, positing abstract structure behind the standard. Looking at other languages, I suggest that the presence of such a constraint in their phrasal comparative may signal that they are to be analyzed as reduced clauses.

Chapter 7 concludes the study, discusses its theoretical implications and outlines directions for further research.

2. *N*-forms and *j*-forms of 3rd person pronouns: Background

2.1. Historical development of n-forms (based on Seliščev 1951, Hill 1977)

Originally, in Common Slavic (Proto-Slavic), the nasal segment of the present-day 'prepositional' pronouns only appeared when the three basic prepositions *vъn* 'in(to)', *sъn* 'with', *kъn* 'towards' were followed by a word beginning with a vowel or glide: [n] was the final segment of these prepositions, subject to a phonological deletion rule when preceding a consonant. It is reasonable to assume that the rule operated on the domain of a prosodic word[1] and had the effect of deleting the syllable-final [n] preceding consonants while making the pre-nasal vowel nasalized. This is informally illustrated in (13).

(13) *Reconstructed phonological rule*: $(C)VC_{[+nasal]} \rightarrow (C)V_{[+nasal]} / [\ldots [\ \_\ ]_{\sigma} [C\ldots]_{\sigma} \ldots]_{\omega}$

Historic third person pronouns, like the modern n-less forms (j-forms), began with [i] or [j] and thus conditioned the realization of the nasal segment. Since most words begin with consonants (both due to the higher probability of a consonant, stemming from their larger repertoire and due to the prototypical/unmarked nature of syllables with an onset), the final [n] must have been consistently realized with third person pronouns, which are furthermore among the most frequent lexemes. It must have been due to this as well as the general preference for open syllables[2] and syllables with an onset (CV), that the final [n] of these prepositions was reanalyzed as the initial [n] of third person pronouns.[3] (14) depicts a reconstruction of how an additional set of third person pronouns emerged in the language.[4]

(14) Vъn+jь→Vъn^{j}ь→ Vъ-n^{j}ь (in + 3MSG.ACC) - 'into him'

[1] P-PRN collocations likely formed one prosodic word, given that the palatalization rule, a PWd-internal process applied to them, (cf. (14).

[2] Around the 6th century AD the language came to only tolerate open syllables (Hill 1977: 308-309; Seliščev 1951).

[3] Interestingly, such reanalysis, even involving specifically [n] changing from the coda of the preceding word and becoming the onset of the following one also happened in English and Swedish (Campbell 2013: 103). Campbell also points out that "reanalysis is one of the most important mechanisms of syntactic change".

[4] A similar process led to the formation of new lexical items of other categories as well. For example, the noun *nedra* 'bowels (of the earth), depths (metaphorical)' was formed from the *n* of the preposition *vъn* 'in(to)' and the noun **jedro* > *jadro* 'core, nucleus'. Compare also *sned'* 'food/edible things' and *s''edobnyj* 'edible': the root of both is *ed*, found e.g. in *eda* 'food', while *sn* of the former is the historic variant of the prefix/preposition *s* 'with' and *s''* of the latter is the synchronic form of this prefix.

Vъn+jemь→Vъn^{j}emь→Vъ-n^{j}emь (in + 3MSG.LOC) - 'in him'

Sъn+jimь→Sъn^{j}imь→Sъ-n^{j}imь (with + 3MSG.INS) - 'with him'

Kъn+jemu→ Kъn^{j}emu→ Kъ-n^{j}emu (towards + 3MSG.DAT) – 'to(wards) him'[5]

The grammar of Proto-Slavic can be said to have accommodated these new, n-initial pronouns as "prepositional" forms: according to Seliščev (1951), they came to be used with a number of other prepositions (*za n'imь* 'behind/after him.INS', *o n'otь* 'about him.LOC') at an early stage. If this is correct, the distribution of [n] in prepositional environments had ceased to be conditioned phonologically long before the present day. Thus, the notions of epenthesis and prosthesis, sometimes employed to refer to the appearance of [n] on third person pronouns (cf. Yadroff & Franks 1999), are synchronically misleading: no adequate (phonological) insertion analysis[6] can be posited, especially in view of the fact that n-form licensing is sensitive not only to the particular category of third person pronouns but also to their syntactic position. Thus, [n] only appears on the actual 3rd person pronominal complement of a P, rather than on any adjacent 3rd person pronoun. Compare (15) to (16-17):

(15) Maša vljubilas' v {*ego; ✓**n**ego}.

Maša fell-in-love in 3MSG.ACC$_{[-n]}$ 3MSG.ACC$_{[+n]}$

'Masha fell in love with him.'

(16) Maša vljubilas' v {✓ego; *__n__ego} brata

Maša fell-in-love in 3MSG.GEN$_{[-n]}$ 3MSG.GEN$_{[+n]}$ brother.ACC

'Masha fell in love with his brother.'

(17) Maša vljubilas' v {✓emu; *__n__emu} nenavistnogo parnja

Maša fell-in-love in 3MSG.DAT$_{[-n]}$ 3MSG.DAT$_{[+n]}$ hateful.ACC guy.ACC

'Masha fell in love with a guy he hates.'

[5] Symbols 'ъ' and 'ь' stand for the Proto-Slavic extra short vowels [ŭ] and [ĭ] respectively. Subscript [j] marks palatalized consonants.

[6] Despite this, Bulygina (1977) still attempts a phonological explanation, which is easily shown to be not viable (cf. Daniel 2015).

In fact, adjacency between the P and its pronominal complement is not even required for the n-form to be licensed. (18) and (19) show that an intervening modifier or an enclitic particle do not preclude the appearance of [n] on the pronoun in Contemporary Russian; (20) shows that an n-form is also licensed on the second complement of a preposition, separated from the P by its first complement and a conjunction.

(18) Maša vljubilas' vo vsex {?ix; ✓**n**ix}

Maša fell-in-love in all.ACC 3PL.ACC**[-n]** 3PL.ACC**[+n]**

'Masha fell in love with all of them.'

(19) Radi že {??ego; ✓**n**ego}

For-the-sake-of TOP 3MSG.GEN**[-n]** 3MSG.GEN**[+n]**

Maša gotova na vsë

Maša ready on everything.ACC

'For him, Masha would do anything.'

(20) Meždu Mašej i {✓im; ✓nim}

Between Maša.INS and 3MSG.INS**[-n]** 3MSG.INS**[+n]**

ustanovilis' družeskie otnošenija

set-up friendly relations

'Masha and he became friends.'

The only linear constraint on n-forms seems to be that the n-form should follow its governing head. For instance, although comparative adjectives and adverbs allow both n-forms and n-less forms on the standard of comparison argument that follows them, only n-less forms may occur *preceding* the comparative:

(21) Ty {✓ix; *nix} lučše {✓ix; ✓nix}

You 3PL.GEN**[-n]** 3PL.GEN**[+n]** better 3PL.GEN**[-n]** 3PL.GEN**[+n]**

'You are better than them.'

This contrast is particularly nicely illustrated in the following minimal pair, courtesy of Pavel Rudnev (p.c.), which was naturally produced by his daughter.

(22) Ja na god starše neë.
1SG.NOM on year.ACC older 3FSG.GEN**[+n]**.
Ja eë na god starše.
1SG.NOM 3FSG.GEN**[-n]** on year.ACC older
'I am a year older than her. *I am (than) her a year older.'

To summarize, the licensing of n-forms in Contemporary Russian should be conceived of in morphosyntactic rather than phonological terms.

2.2. N-forms across Slavic

The 'adpositional' n-forms developed in Old Slavic followed different paths in individual Slavic languages.

In Russian, as well as Ukrainian and West Slavic languages (Polish, Czech and Slovak), n-forms largely preserved association with the P-object position (cf. Comrie & Corbett 1993), expanding to prepositional phrases, headed by 'secondary' prepositions. They are thus in opposition to n-less or j-forms, which I call unmarked in Table 1. However, there is a difference between Russian and Ukrainian on the one hand and West Slavic on the other: the latter also have clitic forms for certain pronouns: GEN/ACC and DAT for 2nd (and 1st person) singular pronouns as well as the for 3rd person masculine. Thus, they display a three-way distinction for the GEN/ACC and DAT of the 3rd person masculine, as shown in Table 2. In addition, Russian and Czech, Contemporary Slovak, Polish, Ukrainian do not have a j-form/n-form opposition in the instrumental case – n-forms appear across the board, as can be seen from Table 2. This makes the INSTR case similar to the PREP (LOC) case in these languages. It is tempting to hypothesize that INSTR has become a P-case, implying that instrumental n-forms are always governed by a P-head (overt or null) in these languages. Developing (or rejecting) this idea needs further research.

In South Slavic, on the other hand, n-forms have evolved into phonologically strong pronouns standing in a paradigmatic opposition to pronominal clitics. Still, these strong pronouns do occur in P-object positions as well, but that is apparently due to prepositions being phonologically weak (proclitic) by and large.

Finally, Contemporary Belorussian and some East Slavic dialects lack *n*-pronouns altogether (c.f. Šaxmatov 1957: 310-311). It is not clear how exactly they lost them, but it is estimated to have happened not earlier than in the 17th century (Hill 1977).

The functional reason behind the survival of the n-form/j-form distinction as the 'adpositional'/regular form distinction in Russian, Ukrainian and West Slavic languages is likely to be its ambiguity resolution potential. There is a striking correlation between the preservation of n-forms as 'adpositional' variants and the lack of designated possessive forms for 3rd person pronouns (in contrast to 1st and 2nd person pronouns); compare the 1st person singular and 3rd person singular masculine/neuter pronominal paradigms for languages of this type, illustrated by Russian in Table 1.

Table 1. 1st person and 3rd person pronominal paradigms in the 3 types of Slavic languages

1SG							
CASE	**Russian**			**Serbo-Croatian**		**Belorussian**	
	Clitic	*Non-clitic*		***Clitic***	***Non-clitic***	*Clitic*	*Non-clitic*
		unmarked	**adpositional**				
NOM	-	ja	-	-	Ja	-	ja
GEN	-	**menja**	-	Me	**Mene**	-	**mjane**
ACC	-	**menja**	-	Me	**Mene**	-	**mjane**
POSS	-	moj	-	-	Môj	-	moj
DAT	-	mne	-	Mi	Meni	-	mne
INSTR	-	mnoj(u)	-	-	mnom, mnome	-	mnoj(u)
PREP	-	mne	-	-	Meni	-	mne
3SG MASC/NEUTER							
CASE	**Russian**			**Serbo-Croatian**		**Belorussian**	
	Clitic	*Non-clitic*		***Clitic***	***Non-clitic***	*Clitic*	*Non-clitic*
		unmarked	**adpositional**				
NOM	-	on (ono)	-	-	on (ono)	-	ën/janó
GEN	-	**ego**	nego	**Ga**	**njega**	-	**jahó**
ACC	-	**ego**	nego	**ga**, nj	**njega**	-	**jahó**
POSS	-	**ego**	-	-	njegov	-	**jahó** > jahóny
DAT	-	emu	nemu	mu	njemu	-	jamu
INSTR	-	im	nim	-	njim, njime	-	im
PREP	-	-	nëm	-	njemu	-	im

Table 2. Three-way distinction for the 3rd person masculine pronoun in West Slavic

CASE	**Czech (3SG.MS.AN)**			**Polish (3SG.MS)**		
	Clitic	*Non-clitic*		*Clitic*	*Non-clitic*	
		unmarked	**adpositional**		**unmarked**	**adpositional**
NOM		on			on	
GEN	ho	**jeho**	něho	go	jego	niego
POSS		jeho			jego	
DAT	mu	**jemu**	němu	mu	jemu	niemu
INSTR		jím	nim		**nim**	
ACC	ho	**jeho**	neho/něj	go	jego	niego
PREP			něm			**nim**

By contrast, in South Slavic languages, exemplified by Serbo-Croatian in Table 1, there do exist adjective-like possessive pronouns in the 1st, 2nd and 3rd person pronominal paradigm alike. At the same time, n-forms are not 'adpositional' – they are phonologically strong forms of pronouns, contrasting with clitics, typically 2nd position (Wackernagel) clitics.

In Russian, Ukrainian and West Slavic languages adjective-like possessive pronouns are only available for 1st and 2nd person pronouns: the 3rd person possessor function is fulfilled by the genitive j-form, whereas n-forms, in turn, are conceived of as 'adpositional' forms of pronouns, rather than prosodically strong forms. Each 3rd person pronoun in Russian displays the GEN-ACC syncretism and, on top of it, these GEN-ACC forms are used in the possessive function, so *ego/eë/ix* 'he/she/they$_{GEN=ACC}$' are three-way ambiguous out of context. Serbo-Croatian, on the other hand, only has GEN-ACC ambiguity for 3rd person masculine and plural pronouns (*ga-njega/ix-nix)* and no such ambiguity for the feminine pronoun *je/nje* 'she.GEN' vs. *ju/njû* 'she.ACC', whereas possessive pronouns have a distinct form, e.g. *njegov* 'his.M'*; njên, njezin* 'her.M'*; njihov* 'their.M'. In this respect, Polish, Czech, Slovak and Ukrainian are similar to Russian, whereas Slovenian, Macedonian and Bulgarian are like Serbo-Croatian.

Belorussian and certain East Slavic dialects are a third type of language: as can be seen from Table 1, they lack n-forms altogether. Since all Slavic languages have a common ancestor which did have n-forms, it means that n-forms in the third group of languages have been lost. Indeed, n-forms used to exist in Belorussian as 'adpositional' forms and occur as late as in the 18th century documents (cf. Hill 1977). Given that

standard Belorussian has developed adjective-like possessive forms in the third person pronominal paradigm (e.g. *jahony* 'his'), it is reasonable to connect the late loss of n-forms with the emergence of these designated possessive pronouns. In other words, in this respect Belorussian now bears certain affinity to the South Slavic languages. The crucial difference is that in Belorussian n-forms disappeared at a late stage when the pronominal clitics had long been lost (cf. Jakobson 1971 [1935] on the timing of this loss). Under these circumstances, the n-forms could not have been reanalyzed as strong pronouns as there were no clitics to create such an opposition. At the same time, n-forms became redundant as possessive and genitive/accusative case forms in the P-complement position had already become distinct. Since j-forms had always been used in a wider variety of contexts than n-forms (as a default or unmarked form), it is the n-series that was lost. Interestingly, the reduction of the paradigm also led to a "uniquely Belorussian innovation[n] […namely] the extension of the initial /j/ element of the other cases to the nominative" (Mayo 1993: 905), cf. Table 1.

To summarize, n-forms in the Slavic languages either serve as phonologically strong 3rd person pronouns (in South Slavic) or as 'adpositional' forms (in Russian, Ukrainian and West Slavic). The reviewed East and West Slavic facts strongly suggest that adpositional/non-adpositional distinction persists because it helps resolve the GEN/ACC vs. possessive ambiguity found with 3rd person pronouns. This, however, does not explain why they are only found on P-objects and in the Russian phrasal comparative, but not elsewhere, so we still need uncover a formal mechanism of their licensing.

2.3. Previous treatments of n-forms: their findings and empirical and analytical limitations

The distribution of n-forms in Russian has gained some attention in the literature. However, to the best of my knowledge, there exists only one generative analysis of n-forms, that of Yadroff and Franks (1999). Other literature discussing 'prepositional' pronouns – Hill (1977), Daniel (2015) – is non-generative. In this section I review these three studies. We will see that Yadroff and Franks (1999) cannot account for the appearance of n-forms in the complement of most 'lexical' prepositions: the syntax of lexical prepositions that they propose is incompatible with n-forms, if their account of n-form licensing is to be maintained. This problem appears to stem from the fact that

the set of prepositions they consider is too restricted: although they do acknowledge that there exist exceptional 'lexical' prepositions that take "prothetic *n*-", they seem to be unaware of how abundant these exceptions are. Hill (1977) and Daniel (2015) describe the distribution of n-forms more precisely, but by virtue of being structuralist-functionalist in spirit, do not propose any formal mechanism of n-form and j-form licensing that one could adopt and develop. A formal account of the respective distribution of n-forms and j-forms is therefore still lacking.

2.3.1. Yadroff and Franks (1999)

Yadroff and Franks' (1999) study is not devoted solely to n-forms. Rather, the paper is a proposal to split Russian prepositions into the functional and the (semi-)lexical class, in the spirit of van Riemsdijk's (1990, 1998) proposal for Germanic prepositions. Class membership of a given prepositional lexeme is determined based on a number of criteria that "span the gamut of linguistic modules" (Yadroff and Franks 1999: p.4), that is, that have to do with phonological, morphological, syntactic and semantic properties. Licensing of n-forms or taking of "the prothetic *n*- before 3rd person pronouns" as they put it (Y&F: p.7), is one such criterion. Their generalization is that functional prepositions require the "prothetic *n*-" whereas "newer lexical additions to the group [of prepositions – *T.P.*] usually do not take this *n*-" (Y&F: p.7). They illustrate the contrast using the following examples:

(23)

Functional prepositions	Lexical prepositions
*v {nej; *ej}* 'in her'	*navstreču {*nemu; emu}* 'towards him'
*ot {nego; *ego}* 'from him'	*vopreki {*nej; ej}* 'despite her'
*u {nix; *ix}* 'at their place' [lit. 'at them']	*blagodarja {*nim; im}* 'thanks to them'

(adapted from Yadroff and Franks 1999: ex. 9)

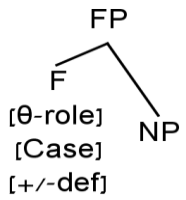
At the same time, Yadroff and Franks express their doubts concerning the reliability of this diagnostic for Contemporary Russian, since there are (lexical) prepositions allowing either pronominal form, like *othositel'no* 'with respect to' and since "the prothetic *n*-" can also appear after comparatives (*ne men'še nego* 'no less than him').

Although at the first sight the distribution of *n*-forms indeed seems chaotic, we will see in further sections that there is in fact a system behind it. There are a very

large number of morphologically derived and therefore '(semi-)lexical' prepositions and even P-N combinations, which behave like *otnositel'no*. Crucially, all of them are associated with genitive case assignment. Only a much smaller group of derived prepositions that govern case other than genitive (dative, instrumental or accusative) behave as Yadroff and Franks's 'lexical' prepositions, i.e. do not admit "the prothetic *n*-".

Despite doubting the reliability of the prothetic *n*- diagnostic, Yadroff and Franks still attempt to derive the facts in (23) from their general analysis of functional and lexical prepositions. The main idea of their paper is that functional prepositions do not exist from the perspective of syntax – rather, they are a product of morphology that realize certain functional features associated with NPs, located in the F-node of a generalized functional FP-projection. Under this approach, noun phrases have the following representation in the narrow syntax:

Figure 2. Noun phrase structure in narrow syntax (adapted from Yadroff and Franks, p.14: ex. 28)



At the level of Morphological Structure (MS) syntactic feature bundles are manipulated and mapped onto PF. As a result, the FP projection may be split, i.e. undergo the "fission" operation, when F-features are distributed among a determiner and a preposition; otherwise, if they are encoded on bound morphemes of the noun, the FP layer undergoes "pruning". To illustrate, Yadroff and Franks propose that the goal argument in (24) instantiated by a dative NP in Russian and by a PP in English has the same syntax in both languages, namely the one in Figure 2, with F carrying [+def], [DAT] and [Goal] features.

(24) Ja rasskazyval ženščinam drug o druge
I spoke woman.PL.DAT friend about friend
'I spoke to the women about each other.'

At MS, this FP in English undergoes fission into a DP (with D realizing the [+def] feature) and a PP (with P realizing the [Goal] and [DAT] features), cf. Figure 3. In Russian, the FP is pruned to NP, as in Figure 4, since neither feature has phonological realization, according to Yadroff and Franks (1999).[7]

Figure 3

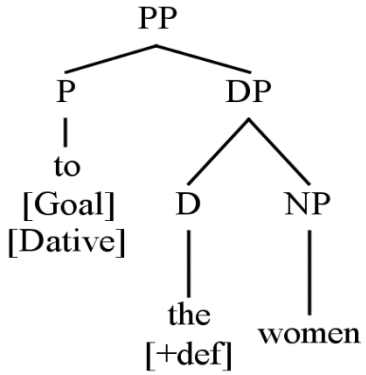


Figure 4

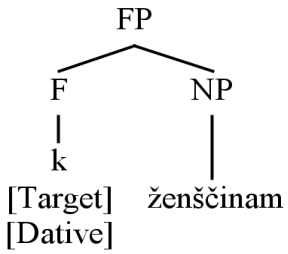
The FP in Russian will not always be pruned since "many case/theta-role combinations do of course require lexical instantiation, parallel to English… These are the functional prepositions…which we claim are simply realizations of F°" (Yadroff and Franks 1999: p.15). An example provided by Y&F is *k ženščinam* 'to(wards) the

[7] This point is not uncontroversial, since nominal inflection depends on the case that the noun bears, although there is indeed no special morpheme just for case.

women', an argument occurring in sentences like (25), which apparently fulfills a Target theta-role.

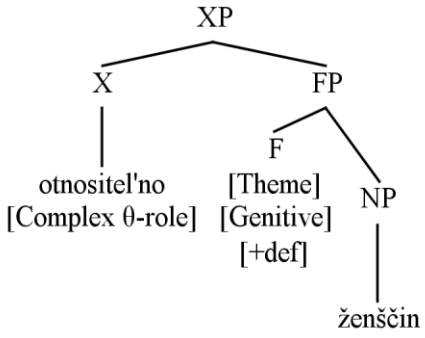
(25) Ivan pošël k ženščinam
Ivan went to woman.PL.DAT
'Ivan went to (the) women.'

Figure 5



'Lexical' prepositions, on the other hand, do project in the syntax. Their only difference from the lexical categories they are derived from (i.e. N, V and A) is that they are "functionally bleached" (Y&F 1999: 16), i.e. lack functional superstructure. The structure of the phrases headed by such prepositions is illustrated in Figure 6 for *otnositel'no ženščin* 'regarding/with regard to women'.

Figure 6



(Yadroff and Franks, p.16: ex. 32)

Now let us examine Yadroff and Franks' proposal concerning n-forms. Yadroff and Franks suggest that "Russian pronouns ordinarily reflect fusion of F^0 and N^0 in MS. That is, *ego* [and n-less forms in general – *T.P.*] is realized in F^0" (Y&F 1999: p.17). N-forms appear when a third person pronoun is forced to be realized in N. This happens whenever F^0 is *overt*, i.e. whenever there is a functional preposition occupying the F-node. This account, however fails empirically. As mentioned above, n-forms do not necessarily appear following functional prepositions: a large subset of "lexical" prepositions (including *otnositel'no*) and comparative adjectives and adverbs allow n-forms as well. To maintain Y&F's analysis for these cases, one would have to postulate a null functional preposition in F that would morphologically force the pronominal complement to be realized under NP. Conceptual problems aside, in order to generate both n-forms and n-less forms in the complement of *otnositel'no* in Figure 6, we would need to posit a silent morpheme occupying the F-node in the former and no such morpheme in the latter case. This, however, would be a vacuous distinction that furthermore would have to be postulated for most but not all 'lexical' prepositions. Although, as I mentioned, the contrast between genitive- and non-genitive-assigning prepositions is systematic, it seems impossible to accommodate in Yadroff and Franks' analysis.[8]

Yadroff and Franks (1999) face several other conceptual and empirical problems. Since they are not directly connected to the n-form licensing issue, let me just mention one of the most problematic implications, which concerns Case licensing. If the lexemes that Yadroff and Franks classify as functional prepositions are a product of morphology, they cannot assign Case to the noun phrase that is their complement. This means that morphological case of the NP/DP as well as the actual prepositional morpheme have to be determined by some other element in the clause. Whereas it might be made to work when a PP instantiates a theta-role-bearing argument of a particular predicate, it is not clear how to deal with PP-adjuncts or PP-predicates headed by functional prepositions. These are usually not considered to bear any theta-role, so neither the case of the NP, nor the choice of the preposition can be thematically motivated. The only option that I can see is that some abstract null element is what licenses a particular prepositional phrase but in that case, we would need a very elaborate system of semantically rich null heads. A more natural account

[8] It is also not clear how their analysis would accommodate the fact that either n-form or j-form may appear in P-complement when there is an intervening modifier, cf. (18).

of such structures would involve giving prepositions some syntactic status, in particular, the ability to assign case to their complement NP/DP.

2.3.2. Daniel (2015)

Daniel (2015) is the most recent treatment of n-forms in Russian. His data comes from search queries in the Russian National Corpus (RNC).[9] Since it is not a quantitative corpus study examining individual prepositions in detail, Daniel (2015) does not aim to provide a comprehensive picture of n- and n-less forms distribution, rather, he attempts to make major qualitative generalizations. The observations he makes are the following:

1) n-forms are obligatory with "primary" prepositions and optional or impossible with "secondary" prepositions;[10]
2) n-forms appear optionally rather than obligatorily after comparative adjectives and adverbs;
3) n-forms are fully ungrammatical when the pronoun depends on a verb;
4) adjacency to P does not necessarily trigger an n-form;
5) n-forms do not have to be adjacent to P[11];
6) non-P-adjacent third person pronouns may appear in the n-less form;
7) despite (4-6) adjacency does have a certain effect on n-forms' occurrence: n-forms are generally less obligatory when non-adjacent to a preposition.

We will see below that while observations (3)-(6) always hold true, (1)-(2) are slightly more problematic. In particular, (1) should be formulated more strictly, since the primary/secondary distinction is notoriously vague (see Hill 1977). As to (2), while it is generally true, we do find comparative adverbs such as *prežde* 'earlier' that only occur with n-forms in the 1980-2015 subpart of the RNC, which I consider to be

[9] Available online at http://ruscorpora.ru/.
[10] 'Primary' and 'secondary' are to be understood as morphologically derived and underived. This distinction corresponds to the lexical/functional distinction of Yadroff and Franks.
[11] Daniel considers cases involving modifiers *ves'/vse* 'whole/all' and *sam* 'self (emphatic)' that may intervene between a preposition and an n-form and also n-forms in the second coordinated complement of the preposition *meždu* 'between' that requires a plural or a coordinated complement. He actually says that in the RNC (but not in general) n-forms in the second conjunct position are restricted to the complements of *meždu* 'between'. This, again, is not strictly the case: I also managed to find one example for *krome* 'except' and *dlja* 'for' each, although the latter involved a disjunction *ili* 'or', cf. Section 4.1.

representative of Contemporary Russian. Since the RNC stretches back to the 18th century texts and Daniel (2015) does not restrict his queries to the modern period, the overall picture gets more obscure.

The fact that the realization of n-forms is sensitive to the linear distance between P and the pronoun and that the degree of obligatoriness of n-forms depends on a lexical item, leads Daniel to propose, following Zaliznjak (2002 [1967]), that n-forms are mere *variants* of specific case forms, rather than the realization of a value of a separate morphosyntactic category of "adpositionality" – an option entertained and rejected in both Daniel (2015) and Zaliznjak (2002 [1967]). In particular, "to account for the morphosyntactic variation, both across types of contexts [e.g. across prepositional lexemes – *T.P.*] and across texts [meaning more or less across speakers – *T.P.*], I [Michael Daniel – *T.P.*] suggest substituting the notion of controller that *requires* a certain value of a *morphosyntactic category* with the notion of a trigger that *allows* to *favors* to *requires* a *non-standard realization* of a form." (Daniel 2015: p. 102). In this framework, *nemu* '3SG.M.DAT' is a contextually determined "*adpositional variant* of the dative value of the case category" rather than a separate value of a distinct category, i.e. adpositional singular dative, as opposed to *emu* '3SG.M.DAT', which is a non-adpositional singular dative.

As far as I can see, if we recast Daniel's (2015) proposal in generative terms, the prepositions that require n-forms would need to be specified as such "triggers" in the lexicon, as "the alternation seems to be partly governed by lexical properties of the controllers" (Daniel 2015: p. 99) and those prepositions that disallow n-forms would not be so specified. Although this means that n-form-taking is idiosyncratic (which we would want to avoid ideally) and it is hard to formalize how exactly prepositions should be specified as such triggers, this idea seems realizable in the generative approach. More serious problems begin when we need to specify a vast number of prepositions (and comparative adjectives and adverbs) as weak triggers and especially if we want to capture adjacency effects, i.e. that an n-form is not obligatory on a P-complement pronoun that is not adjacent to P. This, of course, might mean that the generative approach is on a wrong track. However, even if we judge Daniel's (2015) approach on its own terms, it still faces problems. The lexical specification approach implies that the set of lexical items allowing/requiring n-forms is idiosyncratic. It does not provide a principled explanation for why n-forms are possible with e.g. *naščët* 'about, concerning' but impossible with *nazlo* 'to spite' – morphologically very

similar prepositions that differ in terms of which morphological case appears on their notional complement (genitive in the former and dative in the latter). In other words, we are not able to accommodate the genitive/non-genitive generalization under this approach: saying that only prepositions specified for [GEN] trigger n-forms would be inaccurate, since prepositions triggering n-forms assign the full range of cases available to prepositions.

We run into deeper problems when we look at preposition *vzamen* 'in exchange for' that 'triggers' n-forms when its notional complement bears genitive case, as in (28) but not when it bears dative as in (26).[12]

(26) Ničego cel'nogo i xudožestvenno ubeditel'nogo *vzamen* **emu**
Nothing holistic and artistically convincing in-replacement he.DAT
[naturalističeskomy teatru] sozdano ne bylo
naturalistic.DAT theatre.DAT created NEG was
'Nothing holistic and artistically solid was created to replace it [the naturalistic theatre].'

[Yu. Elagin. 1998. *Temnyi genij*]

(27) *Vzamen* **ego** [fakul'teta] byl otkryt otdel'nyj institut
In-replacement he.GEN faculty was opened separate institute
'A new institute was established to replace it [the faculty].'

[I.M. Djakonov. 1995. *Kniga vospominanij*]

(28) On perestal byt' idolom a *vzamen* **nego**
He ceased be.INF idol.INS CONJ in-exchange he.GEN
nekogo bylo postavit'
no one was place.INF
'He ceased to be an idol and there was no one to suggest in his replacement.'

[A. Laskin. Angel letjaščij na velosipede. *Zvezda*. 2001]

A way out would be to suggest that this preposition is lexically ambiguous: e.g. *vzamen*$_1$ [GEN], [n-trigger] vs. *vzamen*$_2$ [DAT], but to the best of my knowledge, solutions of this kind are not common practice in functionalist linguistic approaches.

[12] Examples (26-28) come from the Russian National Corpus.

To be fair, in his concluding remarks, Daniel (2015: p. 103) mentions that "in the long run, this may lead to a situation where the trigger of a variant becomes a true controller of a morphosyntactic value, the initial stem alternation turning into a full-fledge [sic] inflectional category". Eventually, I will argue that a morphosyntactic approach that is similar in spirit can derive the full range of synchronic data, once we look at individual prepositions more closely and distinguish between prepositions as a functionally defined part-of-speech and the syntactic category P.

2.3.3. Hill (1977)

Hill (1977) is a structuralist treatment of n-forms' distribution based on a thorough historical corpus study, spanning 10 centuries from 1000 to 1964 with a volume of 2,16 mln words.[13] Hill documents the appearance of n-forms with all types of prepositions and with a number of comparative adjectives and adverbs. A "preposition in the broadest sense is defined as any item, other than the inflected forms of a verb or an adjective, which controls an object…in an oblique case form and which in combination with its object forms a phrase capable of functioning as an adverbial modifier" (Hill 1977: pp. 50-51).

Hill further develops a clear, fine-grained, morphology-based classification of the lexical items falling in the group of "prepositions in the broadest sense". This classification is much more useful and objective than the more traditional and vague primary/secondary, underived/derived and functional/lexical distinctions, each of which has been used by many researchers for the description of overlapping but not strictly co-extensive sets of prepositions, as Hill (1977) himself discusses in detail. A slight modification of Hill's classification will prove very useful for the statement of several generalizations in Section 2.4, so I provide its original version here.

Hill (1977) distinguishes the following 9 types of prepositions (the terms are Hill's and the descriptions – mine):

[13] This is the volume of the basic corpus, that is, the sources that were thoroughly examined by the author and does not include additional instances that he came across on an individual basis.

- ***Simple*:** morphologically simple, underived lexemes having prefix counterparts[14]

  The three originally *n*-final prepositions *k, s, v* 'to, with/off, in(to)' and various, predominantly monosyllabic, prepositions like *iz* 'from', *za* 'behind, for', *pod* 'under', *p(e)red* 'before, in front of'.

- ***Double*:** a combination of two simple prepositions

  Only two such prepositions are common in Contemporary Russian, namely *iz-za* 'because; from behind' and *iz-pod* 'from under'.

- ***Root*:** prepositions originally derived from roots, that are synchronically found in other lexemes.

  Čerez 'through, across', *skvoz'* 'through', *protiv* 'opposite, against', *bliz* 'near', *mež* 'between', *sred'* 'among'.

- ***Suffixed*:** prepositions originally derived from an inflected, non-prefixed noun.

  Dlja 'for', *radi* 'for the sake of', *meždu* 'between', *sredi* 'among', *krome* 'except', *mimo* 'past'.[15]

- ***Prefixed*:** prepositions originally derived from a simple preposition and a noun (inflected or not).

  Vozle, podle 'near, close to', *posle* 'after', *okolo* 'near, around', *vrode* 'like', *vokrug* 'around, round'.

- ***Combination*:** preposition-noun collocations, which are basically the same type as prefixed prepositions, but less lexicalized, which is reflected in their orthography.

[14] Only preposition *k* 'to(wards)' lacks a prefix counterpart.

[15] Root prepositions *mež* and *sred'* and suffixed prepositions *meždu* and *sredi* share a root and are nearly equivalent. In Contemporary Russian the latter pair is used much more frequently, whereas the former two Ps are perceived as archaic.

V ramkax 'in the framework (of)', *v čest'* 'in honor/on the occasion (of)', *po povodu* 'as to, on the occasion of' etc.

- ***Comparative:*** prepositions derived from the comparative forms of adverbs and adjectives.[16]
  Prežde 'earlier/before', *ran'še* 'earlier', *pozže* 'later', *bliže* 'closer', *dal'še* 'further', *lučše* 'better' etc.

- ***Adverbial*:** prepositions derived from and having the morphological form of adverbs/neuter short-form adjectives.
  Otnositel'no 'in relation to, regarding', *podobno* 'similarly (to)', *soglasno* 'according (to)' etc.

- ***Gerund*:** prepositions derived from and having the morphological form of gerunds.
  Blagodarja 'owing (to)', *vključaja* 'including', *isključaja* 'excluding' etc.

Hill further introduces the notion of an 'actual' preposition, which applies to those "prepositions in the broadest sense" that can trigger an n-form.

"The presence of epenthetic ***n-*** after a certain preposition **MARKS** that item positively as an actual preposition; the absence of ***n-*** here means that the item is **UNMARKED** - or, more accurately, is marked negatively as a non-actual [or apparent – *T.P.*] preposition" (Hill 1977:47).

"When the object is a 3PP [third person pronoun – *T.P*] beginning with *n-* in any given chronological period, the preposition is defined as ACTUAL, at that time. If the 3 PPO lacks *n-* in any period, the preposition is defined as APPARENT, at that time." (Hill 1977:51)

Another important contribution of Hill's study is the empirical generalizations that he draws. Since most of them are not found in Daniel (2015) or Yadroff and Franks (1999) reviewed above, I review and discuss them here.

[16] Treating comparatives as prepositions is problematic, as will be discussed in Chapter 5.

Hill (1977), similarly to Daniel (2015), observes that interposed modifiers may inhibit the occurrence of *n*, see (29). Importantly, he lists such modifiers as *odin* 'alone/only', *oba* 'both' that occur as interveners besides *ves'* 'whole, all', *vse* 'all.PL' *sam* 'X.self, even'. Hill notes that these modifiers used to completely block the n-forms up until the 20th century, but in the 1923-1962 period both the n-forms and the n-less forms occur in this environment with about equal frequency. Taking into account similar questionnaire results from Blažev's (1962), Hill concludes that both options are acceptable in modern Russian with n-form-taking prepositions.[17]

(29) Maša vljubilas' vo vsex {?ix; ✓**n**ix}
Maša fell-in-love in all.ACC 3PL.ACC**[-n]** 3PL.ACC**[+n]**
'Masha fell in love with all of them.'

Again, like Daniel, Hill (1977) discusses the other environment where the third person pronoun is not adjacent to a preposition, namely the second coordinated complement of *meždu* 'between':

(30) Meždu Mašej i {✓im; ✓nim} ustanovilis' družeskie otnošenija
Between Maša.INS and 3MSG.INS**[-n]/[+n]** set-up friendly relations
'Masha and he became friends.'

Since his corpus contains only 6 such instances from the period in which *meždu* could already be used with n-forms, with only one involving an n-form, he concludes that "[t]he fact that two words intervene between *meždu* and its slot B object seems to exert a greater resistance to the occurrence of n-, than does the single interposed-modifier word. However, we must not exclude the possibility that within another fifty or hundred years the epenthetic nasal might be occurring regularly after both these constructions" (Hill 1977: 235). More than fifty years have passed since the times that

[17] Interestingly, the facts also accord well with the prescriptive reference grammar of Bylinskij & Rozental' (2011), who note that *n-* may or may not appear when the pronominal object is preceded by a modifier: *u vsex nix* and *u vsex ix* 'at all (of) them' are both possible, the second being "considered more modern [sic!]". In addition, Hill (1977: 231) reports that sometimes pronominal modification is accompanied by preposition repetition, e.g. *u nix u vsex* type (12 occ.) and *u vsex u nix* type (4 occ.). Notably, only simple prepositions show this repetition in Hill's corpus: most occurrences are with *u* 'at' – 10, while *v* 'in', *o* 'about', *s* 'with', *na* 'on' account for the remaining 6 examples. This is reminiscent of Yadroff and Hill's observation that only functional prepositions can be doubled in colloquial language.

the most recent examples in Hill (1977) go back to. As we shall see from the corpus and questionnaire data in Chapter 4, n-forms have indeed become much more acceptable and for some speakers obligatory in non-P-adjacent positions.

Another intervener type discussed by Hill (1977) but not by Daniel (2015) is interposed enclitic particles, like *že* in (31).

(31) Radi že {??ego; ✓**n**ego} Maša gotova na vsë
For-the-sake-of TOP 3MSG.GEN**[-n]** 3MSG.GEN**[+n]** Maša ready on everything.ACC
'For him, Masha would do anything.'

The observation that 'secondary' but not 'primary' prepositions can be separated from their complements by enclitic particles goes back to Vostokov (1874).[18] Vostokov, however, did not say anything about third person pronominal complements in this type of environment. Hill (1977) in his turn does find a number of relevant examples from different periods, all of which involve n-less forms. He rightfully says that these data are not very insightful since the prepositions in all these instances did not even take n-forms adjacently in the period these *P enclitic pronoun* examples go back to. Hill's mention of this kind of an intervener is important, since n-forms are possible in this environment in Contemporary Russian, cf. (31). To be fair, Yadroff and Franks (1999) do point out, citing Vostokov, that lexical, but not functional prepositions allow intercalating particles. However, they do not discuss the issue of which pronominal form surfaces in such cases.

Another data point of Hill's would have been in order in Daniel's (2015) discussion of how the n-form triggering historically changed from being a purely (morpho)phonological process to becoming a morphosyntactic one. Hill observes that immediately right-adjacent third person pronouns which depend on the nominal P-object rather than on the P itself exclusively appeared with an initial *n* in Old Church Slavonic and Old Russian, i.e. in 1000-1300 [although there aren't particularly many instances of such kind because possessively used third person pronouns were typically post-positioned with respect to the nominal head – *T.P.*]. An example of this is (32) with [n] on the possessor 3rd person pronoun, ungrammatical in Contemporary Russian.

[18] Interestingly, Vostokov has an example with *dlja že* X 'for TOP smth' from 1608, which is ungrammatical in Contemporary Russian.

(32) iz nego doma
From he.GEN[+n] house.GEN
'From his house'

(Hill 1977: 314)

In Middle Russian (1377-1753) these pronouns still appeared with *n* more often than without it and only began to uniformly appear without an initial *n* from the beginning of the 19th century (i.e. already in Modern Russian).

Hill (1977) also puts forward the following two generalizations related to n-form occurrence, both of which are missing in Daniel (2015) and Yadroff and Franks (1999):

(33) **When a P is *postposed* to a 3rd person pronoun, the initial n- never appears.** Notably, however, all Ps that can be postposed actually do not appear with n-forms even when they are preposed, so, as Hill (1977) correctly points out, we cannot claim that it is postpositioning per se that causes the absence of *n*.

(34) **"GOVERNMENT OF A CASE OTHER THAN GENITIVE PRODUCES VERY STRONG RESISTANCE TO ACTUAL STATUS"** (Hill 1977:300): secondary prepositions occurring with non-GEN notional complements never take n-form complements in Hill's corpus.

Similar but not identical observations have been made by several Russian linguists and philologists. Bylinskij and Rozental' (2011: §77, pp. 182-3)[19] in their prescriptive editorial manual were apparently the first to state that *n-* does not appear with secondary prepositions governing **DATIVE** case. Despite that, the dative-governing prepositions are not considered a special group – rather, they are treated on a par with the P-N combinations like *v otnošenii* 'in relation to' (Hill's *combination* prepositions). The absence of n- with both groups of prepositions is attributed to their being "prepositions of latest origin". This kind of account would not work for Contemporary Russian, where we find a clear-cut contrast in *n*-form taking between the dative-governing adverbial, prefixed, gerund and combination prepositions and genitive-governing

[19] The first edition of this manual goes back to 1957 and is cited in Hill (1977). I only have the access to the third, most recent edition, which I cite instead. The discussion of n-forms' use is the same in both editions.

prepositions of any type, including the less lexicalized combination prepositions, discussed by Bylinskij and Rozental'.

Another, less strict generalization made by Hill is that the tendency of a P to occur without an overt complement is in inverse proportion to its tendency toward the actual status (occurrence with n-forms). Hill juxtaposes simple, double, root and suffixed prepositions on the one hand and the five remaining categories on the other. Incidentally, this split nicely accords with the split I eventually propose – the former always project P-heads, whereas the latter may instantiate heads of a different category (e.g. V or A) or even be P-N combinations. For Yadroff and Franks (1999) obligatory object-taking as well as n-form-taking are two of the properties characterizing functional, but not lexical prepositions, so in a way they do capture the tendency that Hill is talking about.

Let me emphasize the importance and novelty of Hill's observations. Coupled with the unbiased morphological classification of prepositions that he develops, his monograph constitutes a significant contribution to the study of Russian prepositions, furthering our understanding of the 'N-factor'.

2.3.4. Concluding remarks

In this section I reviewed one generative and two non-generative treatments of n-forms. I have shown that the generative analysis of Yadroff and Franks (1999) turns out to be neither descriptively, nor explanatorily adequate. Whereas the non-generative treatments of Daniel (2015) and especially Hill (1977) provide a much more comprehensive and accurate description of n-forms' distribution, they also do not give a principled explanation for why n-forms are only licensed with a subset of prepositions. Although the n-form distribution data provided in Hill (1977) are quite precise, his corpus is much more limited in size than the currently available Russian National Corpus and his most recent examples date back to 1964. Given the ongoing n-form expansion wave documented by Hill himself, we need to collect more up-to-date data on n-form/n-less form distribution before attempting a synchronic analysis. I present detailed corpus data in the next section.

2.4. Corpus data for individual prepositions: P *n*-form vs. P *j*-form in Contemporary Russian

The core data I gathered concerns the form 3rd person pronominal complements take with various prepositions.

The sample of prepositions examined was obtained as follows: a nearly exhaustive[20] list of Russian prepositions was gathered from various sources (mainly Švedova 1980, Hill 1977, Lyashevskaja and Sharoff 2009) and those Ps that may in principle take a pronominal complement were selected.[21]

The combinations of each of the prepositions with all 3rd person pronominal forms were then searched for in the Russian National Corpus (specifically, in the Main, Newspaper and Speech subcorpora)[22] and the number of documents they occurred in was registered. The search for combinations involving highly frequent prepositions governing genitive/accusative case was restricted to sentence-final positions: since an *n*-less form in GEN/ACC is ambiguous between a P-complement and a possessive modifier of a P-complement, as discussed above, the simple search returns too many results, making it virtually impossible to count the relevant instances. As possessive modifiers typically do not occur sentence-finally, the aforementioned restricted search allows circumventing the issue and obtaining an approximate picture for these prepositions.

The searches were restricted to documents dating between 1980 and 2015, approximately to one generation, yielding the total volume of 341.4 mln words (104.3 mln in the Main sub-corpus; 228.5 mln in the Newspaper sub-corpus and 8.6 mln in the Speech sub-corpus - 8,6 mln). This was done to obtain an adequate synchronic picture of the distribution of 3rd person pronominal forms with different prepositions. It turned out to be impossible to discern any pattern in the data before the 18th – 21st century corpus was restricted such a way: as documented by Hill (1977), n-forms have undergone a significant expansion starting from the XIX century. Thus, many Ps (and

[20] Despite the fact that the class of prepositions is quite closed, new members, sometimes called phrasal adpositions (Norde 2000) which usually have the P-N form may rather freely enter it; this seems to be cross-linguistically true. This makes the borders of the category quite fluid which makes it hard to determine whether a P-N collocation has already lexicalized into a complex P. The list of prepositions described in the text could not possibly include all P-N collocations of this sort, but it does include a sample of them.

[21] It's known that complements of certain Ps cannot be pronominalized, e.g. *spustja god* '(a year) later', *porjadka 10 čelovek* 'around (10 people)'.

[22] http://ruscorpora.ru/en/index.html, December 2015 version.

probably all synthetic comparatives) that could not license *n*-forms at the end of the 19th – beginning of the 20th century now admit and for many speakers even require *n*-forms.

A table with full quantitative data on individual prepositions is given in Table A1 of the Appendix. Here I provide a qualitative summary of their distribution. Based on the corpus data, I divide prepositions into three major classes:

I. Prepositions **requiring** *n*-forms (percentage of *j*-forms: 0-1%)
II. Prepositions **allowing** *n*-forms
    a. Prepositions *strongly favoring* n-forms (percentage of *j*-forms: 1-10%)
    b. Prepositions allowing both forms (percentage of *j*-forms: 10-94%)
    c. Prepositions *disfavoring n*-forms (percentage of *j*-forms: 95-99%)
III. Prepositions **blocking** *n*-forms (percentage of *j*- forms: 100%)

Within the three classes I distinguish subgroups based on the morphological nature of prepositions. I adopt Hill (1977)'s fine-grained classification, reflecting the morphological composition and the origin of lexical items. Although this classification was already presented in the previous section, I provide it here again, for convenience.[23]

- ***Simple*:** morphologically simple, underived lexemes having prefix counterparts[24]

  The three originally *n*-final prepositions *k, s, v* 'to, with/off, in(to)' and various, predominantly monosyllabic, prepositions like *iz* 'from', *za* 'behind, for', *pod* 'under', *p(e)red* 'before, in front of'.

- ***Double*:** a combination of two simple prepositions

  Only two such prepositions are common in Contemporary Russian, namely *iz-za* 'because; from behind' and *iz-pod* 'from under'.

[23] I only exclude the comparative type here, since they cannot be treated as prepositions, as I argue in Chapter 5.

[24] Only preposition *k* 'to(wards)' lacks a prefix counterpart.

- ***Root*:** prepositions originally derived from roots, that are synchronically found in other lexemes.
  Čerez 'through, across', *skvoz'* 'through', *protiv* 'opposite, against', *bliz* 'near', *mež* 'between', *sred'* 'among'.

- ***Suffixed*:** prepositions originally derived from an inflected, non-prefixed noun.
  Dlja 'for', *radi* 'for the sake of', *meždu* 'between', *sredi* 'among', *krome* 'except', *mimo* 'past'.[25]

- ***Prefixed*:** prepositions originally derived from a simple preposition and a noun (inflected or not).
  Vozle, podle 'near, close to', *posle* 'after', *okolo* 'near, around', *vrode* 'like', *vokrug* 'around, round'.

- ***Combination*:** preposition-noun collocations, which are basically the same type as prefixed prepositions, but less lexicalized, which is reflected in their orthography.
  V ramkax 'in the framework (of)', *v čest'* 'in honor/on the occasion (of)', *po povodu* 'as to, on the occasion of' etc.

- ***Adverbial*:** prepositions derived from and having the morphological form of adverbs/neuter short-form adjectives.
  Otnositel'no 'in relation to, regarding', *podobno* 'similarly (to)', *soglasno* 'according (to)' etc.

- ***Gerund*:** prepositions derived from and having the morphological form of gerunds.
  Blagodarja 'owing (to)', *vključaja* 'including', *isključaja* 'excluding' etc.

[25] Root prepositions *mež* and *sred'* and suffixed prepositions *meždu* and *sredi* share a root and are nearly equivalent. In Contemporary Russian the latter pair is used much more frequently, whereas the former two Ps are perceived as archaic.

Although the three classes distinguished based on the *j*-form proportion are not strictly homogeneous, the following generalizations are robust:

(35) **All simple, double and root prepositions belong to Class I**, i.e. require n-forms;

(36) **GEN-assigning prefixed prepositions are all in Class I-II, whereas DAT-assigning prefixed prepositions are restricted to Class III**. In other words, prepositions like *nasčët* 'concerning', *vmesto* 'instead', *nakanune* 'on the eve of' do occur with *n*-forms, whereas prepositions like *navstreču* 'towards (lit. 'on meeting') or *vsled* 'following (lit. 'in footstep') never occur with *n*-forms. The contrast, therefore, is categorical;

(37) **Class III prepositions are all non-GEN assigning**. This is reminiscent of the generalization of Hill's formulated in (34), stating that prepositions governing case other than the genitive resist becoming 'actual' prepositions, that is, taking the n-forms. The two generalizations are not logically equivalent, however: (34) states that *non-GEN-assigning → n-form excluded*, whereas (37) states that *n-form excluded → non-GEN-assigning.* Note also that (34) holds only for the prefixed, combination, adverbial and gerund prepositions, whereas (37) is a generalization over the whole group of 'prepositions'. *GEN-assignment is thus a sufficient, but not a necessary condition for a preposition to license n-forms.*

(38) **Gerund prepositions only appear in Class II(c) and Class III**, that is, deverbal prepositions virtually do not allow *n*-forms: the only exception is *blagodarja* 'thanks/owing to' which I will discuss separately.

(39) **Suffixed prepositions all belong to Class I-II**, i.e. no suffixed preposition completely disallows *n*-forms;

(40) **Adverbial prepositions are mostly confined to Class III**, i.e. virtually do not allow *n*-forms. The two exceptions are *soglasno* 'according (to)' of Class II(c) and *otnositel'no* 'regarding' of Class II(b), to be discussed below.

We can also make the following statement:

(41) **Prefixed prepositions are the most versatile: they either require, allow or block n-forms**.

No other morphological type appears in all three categories. Combination prepositions, which are the less lexicalized prefixed prepositions are expectedly found in all classes except Class I – none of them requires *n*-forms.

Table 3 Correspondence between Hill's and present classification of prepositions

	Class I	**Class II**	**Class III**
Simple			
Double			
Root			
Suffixed			
Prefixed	GEN-assigning	GEN-assigning	DAT-assigning
Combination		GEN-assigning	DAT-assigning
Adverbial		*otnositel'no* 'in relation to' *soglasno* 'according to'	
Gerund		*blagodarja* 'owing to'	

The versatility of the prefixed/combination type is at least partially due to their being most numerous. The latter fact, in turn, follows from their being a truly open class – more and more frequent P-N combinations lexicalize into a morphologically complex preposition. For some reason P-N combinations are the only active source for new prepositions: simple, double and root prepositions instantiate a completely closed class, with most recent members dating back to the second half of the 19th century (cf.

Hill 1977: 107); suffixed, adverbial and gerund classes are potentially open, but are admitting new members at a much slower pace.

It is also crucial to distinguish *simple, double, root* and *suffixed* prepositions on the one hand from *prefixed/combination, adverbial* and *gerund* prepositions on the other: the members of the former group either require or allow *n*-forms (i.e. never prohibit *n*-forms) while some members of the latter disallow *n*-forms. In other words, each of the latter three classes comprises both the lexemes that admit *n*-forms and those that do not admit them, whereas members of the former four groups admit and predominantly require *n*-forms. All this can be seen from Table 3 above.

To account for the distribution of *n*-forms I will capitalize on the following, slightly idealized generalization:

(42) Prepositions whose morphological composition transparently reflects their origin (i.e. prefixed/combination, adverbial and gerund prepositions) license *n*-forms iff they govern genitive case.

3. Licensing of *n*-forms and *j*-forms with prepositions: Disentangling prepositions and P-heads

3.1. Syntactic licensing of n-forms and j-forms in a nutshell

I propose that n-forms are syntactically licensed by P-heads. In particular, 'n' is the morphological reflex of the privative [Prepositional Case] super-feature which a P-head values on its complement in addition to a specific case feature. N-less forms are licensed when a nominal phrase receives case not from a P-head. Thus, the contrast between n-forms and n-less forms concerns their feature make-up:

(43) *nego:* [3], [SG], [MASC], [ACC], **[P-Case]**

ego: [3], [SG], [MASC], [ACC]

In the spirit of the Distributed Morphology approach, I assume late, post-syntactic Vocabulary insertion. The most specified form available is the one that gets inserted in the morphological component (idea known as the Subset Principle):

"Vocabulary insertion […] requires only that the feature bundle of the Vocabulary item be nondistinct from the features of the terminal node at MS [Morphological Structure – T.P.] that serves as the site of insertion. The competition among different Vocabulary items nondistinct from the features of the terminal node at MS ensures that the Vocabulary item that matches the most features of the node will be inserted." (Halle & Marantz 1993: 121-122).

Thus, n-forms must be inserted whenever they are licensed, that is, every time a 3rd person pronoun is an immediate complement of a P-head and receives case from it. [P-case] is <u>always</u> valued on P-complements, but for historical reasons only a handful of Vocabulary entries (n-forms of 3rd person pronouns) are specified for that feature.

This proposal essentially goes back to the traditional description of *n*-forms as *pripredložnye* 'prepositional' and seems to inherit its problems at the first sight. Thus, the following questions arise:

(44) Why aren't *n*-forms licensed with non-GEN-assigning prefixed/combination, adverbial and gerund prepositions?

(45) How can the same preposition license both forms, given the Subset Principle?

In this section I will demonstrate how the analysis works by considering different groups of prepositions in detail.[26] To anticipate this detailed discussion, let me first give short answers to the questions raised above.

First, I propose that non-genitive-assigning prefixed/combination, adverbial and gerund prepositions are generally not P-heads: some of them are transitive adverbs/adjectives, licensing theta-related, inherent dative case; others are transitive gerunds, displaying the same pattern of case-licensing as the corresponding verbs and the remaining prefixed/combination lexemes are not P-heads but rather P-N combinations.

Second, almost all prepositions admitting both forms (Class II) belong to the prefixed/combination or the suffixed type, which means that they contain a nominal morpheme. This morpheme is the source of genitive case assignment. I propose that lexemes of this morphological type oscillate between fully lexicalized P-heads and N-heads or preposition-noun combinations. *N*-forms obligatorily appear in the former case and *j*-forms surface in the latter, when the genitive case is assigned by the nominal head which preserves certain properties of a full-fledged noun.

Thus, I show that the apparent problems that my proposal encounters are resolved if we look at the non-conforming prepositions more closely and reconsider their actual syntactic/categorial status on independent grounds. To recap, *n*-forms are indeed *pripredložnye* 'adpositional', as traditionally described. However, maintaining an analysis built on this intuition requires answering the rather complex question of which elements fall in the syntactic category of P-heads.

3.2. Class I prepositions: P-heads

The core of the prepositional category is represented by lexemes that are unambiguously P-heads. I propose that all prepositions, not having independently active ancestors from other categories that are discernable for a naïve speaker, are P-

[26] *N*-form licensing in the phrasal comparative will be treated separately in Chapter 5.

heads. Thus, simple prepositions that only have prefix cognates and did not have such ancestors already at the stage of Common Slavic are definitely P-heads. This group includes lexemes like *u* 'at', *v* 'in(to) *s* 'with, off', *o(b)* 'about', *pod* 'under', *iz* 'from' etc.

Double prepositions are essentially combinations of two simple prepositions (*iz-za* 'from behind; because of'; *iz-pod* 'from under') that inherit case-assigning properties of the first simple P and thus also fall in this category. Added to them are root prepositions whose cognates can be found in the roots of words of different lexical categories, but which do not occur as independent nominal lexemes. This is illustrated in Table 4.

Table 4 Root prepositions as morphemes found in words of other categories

Root preposition	**Words containing the P as a morpheme**
čerez 'through, across'	***čeres****čur* 'too much/excessively' ***čeres****polosica* 'strip farming' ***črez****mernyj* 'excessive'
skvoz' 'through'	***skvoz****njak* 'draught'; ***skvoz****noj* – 'end-to-end, straight-through'
protiv 'against'	***protiv****nik* 'opponent'
bliz 'near, close to'	***bliz****kij* 'near, close (Adj)' ***bliz****ost'* 'intimacy'
mež 'between, among'	***meža*** 'boundary, border'

A number of suffixed and prefixed prepositions are perceived today as indeclinable, monomorphemic forms by Russian speakers: their origin is not transparent, and they cannot be analyzed as a combination of syntactically active members of different categories. This fact is supported by their occurrence as morphemes in words of other categories, cf. Table 5.

Table 5 Suffixed and prefixed prepositions as morphemes found in words of other categories

Suffixed/prefixed preposition	**Words containing the P as a morpheme**
meždu 'between, among'	***mežду****narodnyj* 'international' ***mežдо****usobica* 'intestine war'
sredi 'among'	***Sredi****zemnoe more* 'Mediterranean Sea'
krome 'except' (also non-	***krome****šnaja (t'ma)* 'pitch darkness'
mimo 'past'	***mimo****xodom* 'in passing' ***mimo****lëtnyj* 'passing, ephemeral, short-lived'
vne 'out of'	***vne****šnij* 'external' ***vne****klassnyj* 'extracurricular'
posle 'after'	***posle****obedennyj* 'afternoon (Adj)'
okolo 'by, around, about'	***okol****'nyj* 'indirect, circuitous' ***okolo****istoričeskij* 'quasi-historical'

Two suffixed prepositions – *dlja* 'for' and *radi* 'for the sake of' – and two prefixed prepositions – *vozle*, *podle* 'near' – are no less obscure but are not found as morphemes of other words.[27]

[27] *Vozle* 'near, beside' is reported by Hill (1977:146) with reference to Vondrak [1928: 300-19] to consist of the prefix *v''z* and the root **dl'jg*; the former does not exist as a preposition in Contemporary Russian and the latter noun also doesn't exist. According to the Vasmer Online Etymological Dictionary, the original nominal morpheme in *vozle* is **dьl'a*, related to the word *dlina* 'length'. Whichever hypothesis is correct, naïve speakers of Contemporary Russian do not decompose *vozle* 'near, next to' into two meaningful independent elements. *Podle* is reported to come from **po-dəlg* by Hill (1977:154) and from **podьlja/ *podьlě* of *dьlja 'length' in Vasmer. Again, the crucial thing is that neither nominal component exists as a noun in Contemporary Russian.

Finally, it is reasonable to include in the category of P-heads the following four prefixed prepositions: *promež & promeždu* 'between' (archaic), *naprotiv* 'opposite', *pomimo* 'besides', *posredi* 'among, in the middle of'. Although Hill (1977) does not distinguish them from other prefixed prepositions, they differ from the remaining prefixed prepositions in that they are composed of a simple preposition ('prefix') and a root/suffixed preposition, rather than a nominal morpheme. They also inherit the case-assignment properties of the root/suffixed preposition they derive from.

Table 6 'Prefixed' prepositions of type P-P: their composition and case-assignment properties

P+P prefixed prepositions	**1st P-morpheme**	**2nd P-morpheme**
promež 'between' (GEN-assigner)	*pro* (ACC-assigner) Simple P	*mež (GEN/INS-assigner)* Root P
Promeždu 'between' (GEN-assigner)	*pro* (ACC-assigner) Simple P	*mež (GEN/INS-assigner)* Root P
naprotiv 'opposite' (GEN-assigner)	*na* 'on(to)' (LOC/ACC-assigner) Simple P	*protiv* 'against, opposite' (GEN-assigner) Root P
pomimo 'besides' (GEN-assigner)	*po* 'along', 'by' (DAT-assigner) Simple P	*mimo* 'past' (GEN-assigner) Suffixed P
posredi 'amidst' (GEN-assigner)	*po* 'along', 'by' (DAT-assigner) Simple P	*sredi* 'among' (GEN-assigner) Suffixed P

Given that both elements they comprise fall in the category of P-heads and these morphologically complex lexemes are not found as words of other categories, it is reasonable to treat them as P-heads as well.

To summarize, all simple, double and root prepositions (all closed groups) constitute P-heads. They are joined by a limited number of suffixed and prefixed prepositions, which do not involve independently syntactically active nominal lexemes.

If the lexemes discussed here constitute simple P-heads, we expect them to always take n-forms. Most members of this class indeed fall in our corpus-based Class I, that is, the percentage of n-less forms in their third person pronominal complement is less than or equal to 1%. However, for 5 prepositions the picture is messier: *pomimo* 'besides' is 2%, *naprotiv* 'opposite, in front of' – 3.6%, *posredi* 'amidst' – 6.4%, *promež* 'between' – 18%, with the most alarming percentage being registered for the suffixed *vne* 'outside (PLACE)' – 96% of n-less forms. The first three fall into Class II(a), that tends to Class I, so we might want to treat them on a par. *Promež* is quite archaic and very infrequent (11 occurrences overall): I suggest that it is mostly used to create an archaic/dialectal effect, leading to the unexpectedly high percentage of n-less forms; we thus may disregard it. However, the much more frequent *vne* 'out of/outside.LOC' (183 occurrences overall) deserves an explanation. Given that n-forms are highly disfavored rather than unacceptable with *vne*, I argue that *vne* does constitute a P-head for some speakers, but for most speakers it preserves the properties of an N-head it is derived from.[28] Why is that?

I suggest that there are three major factors precluding the complete prepositionalization of *vne*. First, unlike the simple monosyllabic prepositions, which are proclitics (cf. Gribanova 2009) *vne* constitutes an independent prosodic word. It is evident from three facts 1) its only vowel does not undergo phonetic reduction; 2) it can occur without an overt complement; 3) it is able to host enclitic particles like *že* (particle marking a topicalized phrase)[29] and *li* (question particle).

(46) Vne že sten alma mater Katju stol' že
Outside TOP walls.GEN alma mater Katja.ACC equally
neščadnomučil KGB…
unmercifully tortured KGB

[28] According to Vasmer's dictionary, the origin of *vne* 'outside.LOC' is the Old Indic vánē 'forest.LOC (in a forest)'. Likewise, the adverb *von* 'outside.DIR' is hypothesized to derive from the Old Indic vánam 'forest.NOM/ABS'.

[29] Note that *že* is phonologically a second position clitic, that is, it attaches at the right edge of the first phonological word of the topicalized phrase it modifies, cf. e.g. Valova and Slioussar (2016). Hence, examples like (46) show that *vne* 'out of' indeed projects a PWd of its own.

‘Outside of her alma mater Katya was equally unmercifully tortured by the KGB.’

[N. Klimontovič. Dalee – vezde (2001)]

Second, unlike other prepositions that form new prepositions when combined with a simple preposition, *izvne* ‘from outside’ and *vovne* ‘outside (directional)’ that are derived from *vne,* are used only adverbially (i.e. are intransitive).

(47) {Vne; *izvne; *vovne} doma
Outside.LOC from-outside outside.DIR house.GEN
‘Outside; from outside; outside the house.’

Third, complete prepositionalization entails obligatory *n*-form-taking. In the case of *vne* the insertion of a genitive n-form results in a sequence of identical or nearly identical syllables: v**ne ne**go/**ne**ë/**ni**x ‘out of/outside it.M/F/PL’, which violates the Obligatory Contour Principle.[30] With *vne* preserving the N-head status, n-forms are not licensed hence this problem does not arise.

Finally, I would like to note that treating *vne* on a par with Class II prepositions is also justified because it behaves similarly to the four other prefixed/combination prepositions ending in *-ne: posredine*/*poseredine/v seredine* ‘in the middle of’ and *nakanune* ‘on the eve of’, which take *j*-forms in 83-100% of all cases they occur with a third person pronominal complement. The reason for this unusual preference of *j*-forms on the part of *-ne* final ‘prepositions’ must be a constraint like the Obligatory Contour Principle. The plausibility of an account along these lines is emphasized by the existence of the following minimal pair: the prepositions *posredi vs. posredine* both meaning ‘in the middle of, amidst’. If the /-ne/ ending is indeed what precludes *n-form*s with *posredine*, the prediction is that *posredi* will normally occur with them. This prediction is borne out: as can be seen from Table A1 of the Appendix, this P strongly favors *n-forms*.

Table 7 provides the repertoire of P-heads, with prepositions divided by Hill’s morphological classification and frequency in the Russian National Corpus.

[30] The OCP (Obligatory Contour Principle) prohibits adjacent identical elements, e.g. tones, segments, features, syllables and even morphemes (cf. e.g. Rose 2000).

Table 7 Repertoire of P-heads in Russian

Simple (n=20)	**Double (n=4)**	**Root (n=6)**	**Suffixed (n=6)**	**Prefixed (n=8)**	**Total (n=44)**
Prepositions ranking among the first hundred of Russian lexemes by frequency					
v 'in, into, to' *na* 'on, onto, to' *s* 'with/off' *po* 'along/for' *k* 'to(wards)' *iz* 'from, of' *u* 'at, by' *za* 'behind, for, instead' *ot* 'from, of' *o/ob* 'about/against' *do* 'up to, before' *pri* 'at, by' *pod* 'under' *bez* 'without'			*dlja* 'for'	*posle* 'after'	N=19
Prepositions with frequency rank between 101 and 2000					

pered ‘before, in front of’ *nad* ‘above’ *pro* ‘about’	*iz-za* ‘due to; from behind’ *iz-pod* ‘from under’	*čerez* ‘through, via’ *protiv* ‘against’ *skvoz’* ‘through’	*meždu* ‘between, among’ *krome* ‘except’ *sredi* ‘among’ *radi* ‘for the sake of’ *mimo* ‘past’	*okolo* ‘by, around, about’ *vozle* ‘near’ *pomimo* ‘besides’	**N=16**
Prepositions with frequency rank below 2000[31]					
	po-nad ‘above (path)’ (#41085) *po-za* ‘behind (path)’ (#48208)[32]	*mež* ‘between, among’ (#4598) *bliz* ‘near, close to’ (#7846) *sred’* ‘amidst’ (13604)[33]		*posredi* ‘amidst, in the middle of’ (#3341) *naprotiv* ‘opposite’ (#4416) *podle* ‘near’ (#11864)[34]	**N=9**

[31] The number appearing in parentheses next to the preposition is its rank in the frequency list of Russian lexemes based on a 100-million subcorpus of the Russian National Corpus (Lyashevskaya and Sharoff 2009).
[32] Both *po-nad* and *po-za* are archaic and possibly dialectal.
[33] Archaic/poetic: 4.8 ipm - overall, 8.3 ipm – Main subcorpus (mostly fiction), 2.6 – newspaper texts. (Ipm = instances per million words).
[34] Archaic/poetic: 5.8 ipm – overall, 10.4 ipm – Main subcorpus (mostly fiction), 3.8 ipm – newspaper.

				promež ‘between’ (#23907)	

3.3. Class II prepositions: P-heads vs. N-heads/P-N combinations

All combination prepositions, several prefixed prepositions and certain suffixed prepositions (*tipa* ‘like’, *posredstvom* ‘by means of’, *krugom* ‘around, round’ – allowing 3rd person pronominal complements; *putëm* ‘by way of’, *porjadka* ‘in the order of, approximately’ – disallowing 3rd person pronominal complements, since their complements are not referential) involve a nominal element that may occur independently in the language. Most of these prepositions govern genitive case, which is likely inherited from the nominal they contain, as genitive case is typically assigned by nouns to their dependents. I propose that until such an inflected noun/combination is fully prepositionalized (lexicalized), it remains structurally ambiguous between a P-head and an N-head (for suffixed prepositions) or between a P-head and a P-N combination (for prefixed and combination prepositions). The phrase comprising the ‘preposition’ and its notional object looks as follows:

(48) a. Preposition = P-head: [P DP]
b. Preposition = N-head: [N DP]
c. Preposition = P-head + N-head: [P [N DP]]

N-pronouns will surface in structure (48a), i.e. when a suffixed noun/preposition-noun combination is analyzed as a P-head by a given speaker, whereas n-less forms will appear in structures (48b,c), that is, when the nominal morpheme of the ‘preposition’ still projects an N-head and assigns genitive case to the complement DP.

Since the genitive 3rd person possessive modifiers of nouns precede them in the unmarked case (*ego brat* vs. ?*brat ego* ‘his brother’), similarly to true possessive pronouns (*moj brat* vs. ?*brat moj* ‘my brother’) and unlike DPs fulfilling the same function (??/**Miši brat* vs. *brat Miši* ‘Misha’s brother’), a useful test for prepositionalization of an inflected noun or a P-N combination is whether it is possible to place an n-less form *preceding* the nominal element without a substantial change in

meaning. If it is, it suggests that the nominal element is still syntactically active (behaves as an independent noun, to a certain extent) and the combination/suffixed preposition has not completely prepositionalized.

Most genitive-assigning combination prepositions in my sample (17 out of 24) allow both *n*-forms and *j*-forms following the nominal element. 15 of them also allow interposition of the genitive n-less form between the prepositional and the nominal element. Thus, *v otnošenii (n)ego* and *v ego otnošenii* 'in his respect', *v ramkax (n)ego* and *v ego ramkax* 'in its framework' are nearly equivalent, as illustrated by the corpus examples in (49-54).

(49) Vozbuždenie *v otnošenii* *ego* ugolovnogo dela.
Initiation in relation he.GEN criminal.GEN case.GEN

(50) Vozbuždenie *v* *ego* *otnošenii* ugolovnogo dela.
Initiation in he.GEN relation criminal.GEN case.GEN

(51) Vozbuždenie *v* *otnošenii* ***n****ego* ugolovnogo dela.
Initiation in relation he.GEN criminal.GEN case.GEN
'Initiation of a criminal case with respect to him.'

(52) *V ramkax* *ego* sostojatsja sorevnovanija…
In framework.LOC he.GEN will be held competitions
po 12…vidam sporta.
for 12…types of sport
'In its [the festival's] framework 12 different sports competitions will be held.'
[E. Vasilyeva, Novyj Region 2, 2009.08.20]

(53) *V ego* *ramkax* sostojalis' dve naučnye konferencii.
In he.GEN framework.LOC were-held two scientific conferences
'In its [the forum's] framework two scientific conferences were held.'
[Transportnyj forum v Sokol'nikax 2003, Logistika 2003.03.24]

(54) *V ramkax* ***n****ego* budet rasprostranën
In framework.LOC he.GEN will be-distributed
okončatel'nyj variant… položenij…
final variant regulations.GEN

'In its [the forum's] framework the final regulations document will be distributed.'

[Kalendar' sobytij 11 sentjabrja – 12 oktjabrja. RIA Novosti, 2009.09.09]

Na protjaženii 'during, throughout' and *po slučaju* 'on the occasion (of)' (1 instance only) only occur with interposed j-forms in the corpus, i.e. in the order shown in (50) and (53).

Another group of combination prepositions disallows interposition of j-forms but admits both forms after the P-N. Examples of this type are *v tečenie* 'in the course of, during' and *vo vremja* 'during'.

Finally, five items, including *v terminax* 'in terms (of)', *v svete* 'in light (of)' allow interposition and postposition of genitive n-less forms but **do not occur with n-forms** in the corpus.

It is reasonable to propose that these four types of combination prepositions reflect distinct stages of prepositionalization. More specifically, combinations like *v tečenie* allowing both forms but blocking interposition of n-less forms are the most lexicalized items, while combinations like *na protjaženii* only allowing interposition of n-less forms, i.e. the structures shown in (50) and (53) reflect early stage of lexicalization.

To summarize, there are three options for how and where the pronominal complement of combination prepositions is realized (and whose dependent it actually is):

(55) a. postposed n-form: [[P-N] *n-pron*]
b. postposed n-less form: [P [N *j-pron*]]
c. interposed n-less form: [P [*j-pron* N]]

As was said above, adnominal 3rd person pronouns that depend on the N-head typically appear preceding the head noun (unlike nominal dependents)[35] and always take the GEN/ACC syncretic *j*-form. It is therefore reasonable to propose that combination 'prepositions' only allowing the (55c) pattern, are not P-heads: the P-morpheme and the N-morpheme project separate heads in the syntax.

[35] This difference may be due to their different prosodic status: a nominal dependent would typically project a PWd of its own, while *j*-form pronouns may be prosodically dependent on a PWd (cf. Rappaport 1988).

When a given P-morpheme and a given N-morpheme begin to occur next to each other frequently, they eventually become a frozen expression, i.e. are reanalyzed as a single word. At the early stage of this process the *j*-pronouns begin to appear following the noun, as in (55b); the two configurations, (55c) and (55b), are both possible at this stage.

The more the original compositional meaning of the P-N complex is bleached, the more preposition-like it becomes, as the complex is being reanalyzed into a transitive, indeclinable lexeme. When the etymological source of the P-N element becomes obscure to naïve speakers the prepositionalization process is completed. We have seen prefixed prepositions that originally evolved from P-N combinations, but whose nominal element is not connected to any independent noun in Contemporary Russian. We thus said that they are unambiguously P-heads and as such require n-forms.

By contrast, all combination 'prepositions' in our database comprise a nominal morpheme that can be an independent noun in Contemporary Russian. This and their comparatively low frequency (below 2000 ipm) has precluded their complete prepositionalization. Thus, the most lexicalized items among combination 'prepositions' still allow both postposed n-forms and postposed *j*-forms, i.e. configurations in (55a) and (55b).

Thus, a PP projected by an apparent combination preposition, i.e. a frequently used P-N combination requiring a DP-complement may take three different forms when the DP-complement is a 3rd person pronoun, which have been schematically shown in (55). Which of the three options are allowed for a given P-N complex depends on how lexicalized the complex is.

I thus distinguish 5 stages of lexicalization for P-N combinations, from Stage 0 (lexicalization not yet initiated) to Stage 4 (lexicalization completed). Table 8 shows the behavior of a P-N combination with respect to 3rd person pronominal complement that characterizes these stages. In our sample of combination prepositions, we have P-N complexes representing each stage of lexicalization, except Stage 4, standing for complete lexicalization (i.e. becoming a member of Class I).

Table 8 Stages of lexicalization of P-N complexes

	[P [*j-pron* N]]	[P [N *j-pron*]]	[[P-N] *n-pron*]	Example	No. of 'prepositions'
Stage 0	✓			*na protjaženii* 'during, throughout'	2
Stage 1	✓	✓		*v svete* 'in light (of)'	5
Stage 2	✓	✓	✓	*v ramkax* 'in the framework (of)'	15
Stage 3		✓	✓	*v tečenie* 'in the course of, during'	2
Stage 4 (final)			✓	--	0

Thus, prepositions going through Stage 2 and 3 are analyzed as ambiguous between morphologically complex P-heads and P-N combinations with active P- and N-heads.

With suffixed prepositions involving clear independently active nouns we also find that n-less forms are available. *Posredstvom* 'by means (of)' (frequency = 4729 ipm) behaves similarly to the combination prepositions going through Stage 2: it allows n-less forms on the left and both n- and j-forms on the right, as shown in (56-58).

(56) Sroki dostavki deneg *eë* *posredstvom* –
Period delivery money she.GEN[-n] by-means

ot	četyrëx	časov	do	sutok
from	four	hours	till	24 hours

'Money transfer by its [this system's] means takes between 4 and 24 hours.'

[Sistemy denežnyx perevodov. Komsomol'skaja Pravda, 2003.11.03]

(57)

...posredstvom	*eë*	Vovka	govorit
By-means	she.GEN[-n]	Vovka	speaks
s	posetiteljami	za	dver'ju
with	visitors	behind	door

'By its [receiver's] means Vovka speaks to visitors behind the door.'

[È. Limonov. Podrostok Savenko (1982)]

(58)

Posredstvom	*nego*	živët specifika
By-means	he.GEN[+n]	lives specificity
nacional'noj	duxovnoj	kul'tury
national	spiritual	culture

'By its [the language's] means the specifics of the nation's nonmaterial culture are preserved.'

[B. Xamčiev. Obščnost' sudeb. Žizn' nacional'nostej, 2002.10.14]

The 'preposition' *tipa* 'like' is much less frequent overall (36835 ipm), so only two occurrences of n-forms and no occurrences of n-less forms were found following *tipa* in the investigated subpart of the RNC. Instances of the latter type are thus particularly hard to come by: the example presented in (60) was found via search across blogs at www.livejournal.com and even there only one example of this type was found. At the same time, genitive n-less forms preceding the nominal *tipa* 'like' are found in the RNC, with (59) as one of those instances. The *j-form + (nominal) preposition* phrase in (59) functions and is distributed similarly to *(nominal) preposition + n-form* in (61): both instantiate a post-nominal modifier.

(59)

On…	znal	čto	nado	ženščinam	*eë*	*tipa*
He	knew	what	necessary	women	she.GEN[-n]	type

'He knew what women of her type needed.'

[E. Kozyreva. 2001. Damskaja oxota]

(60) Zakazat’ Našej Raše ili komu-to *tipa* *ix*
order.INF Nasha.DAT Rasha.DAT or someone.DAT type they.GEN[-n]
fil’m v stile “Policejskoj Akademii”
film in style Police Academy
‘…to commission “Nasha Rasha” or someone like them to produce a film in the style of *Police Academy*’.
[I. Bigdan. http://ibigdan.livejournal.com. 2011.09.19]

(61) Možet li porjadočnyj, vernyj, mužčina *tipa* *nego*
Can Q decent loyal man type he.GEN[+n]
ljubit’ srazu dvux ženščin?
love.INF at-once two women
‘Can a decent and loyal man like him love two women at the same time?’
[I. Saxnovskij. Oktjabr’ 2003]

Thus, *tipa* and *posredstvom* behave alike and are to be distinguished from other, more obscure suffixed prepositions discussed in Section 3.2 on Class I prepositions. I will thus treat these two suffixed prepositions as ambiguous between P-heads and N-heads.

Krugom ‘around (lit. circle.INS)’ seems to be different: it originated as a complement-less noun in the adverbial function, always used in transitive sentences. With time the complement (the direct object) of a transitive predicate became associated with this adverbial noun, becoming its complement and making *krugom* ‘around’ preposition-like. Examples with *krugom* are very few and all come from the Main subcorpus. The only example with an n-less form is stylistically non-neutral, which might have driven the j-form choice. In general, prepositional *krugom* is very infrequent: in Contemporary Russian it is still mostly found as an adverbial, like in (62).

(62) Oni …vyšli iz doma i obošli ego krugom
They out-went from house and went-around it round
‘They went out of the house and walked around it.’
[A. Marinina. 1996. Illjuzija grexa]

To conclude, the availability of both *n*- and *j*-forms in the complement position of a large number of GEN-assigning combination and suffixed prepositions is due to the nominal element remaining partially syntactically active in that it retains the ability to license genitive case on the notional complement of the complex 'preposition'.

3.4. Class III: not P-heads

3.4.1. A-heads and V-heads

Most adverbial/deadjectival[36] and gerund[37] prepositions are indistinguishable from their source adjectives/adverbs and gerunds (verbs) as far as their case-assignment properties are concerned.

The four adverbial (deadjectival) 'prepositions' are associated with dative case assignment (63), just like the corresponding transitive adjectives (64) and, in some cases, verbs of the same root (65). This dative case might well be associated with theta-role assignment and thus be an instance of inherent case as originally defined by Chomsky (1986).

(63) Ètot stress podobno lavine obrušivaetsja na okružajuščix…
This stress similar.N avalanche.DAT falls on surrounding
'This stress falls on people around you like an avalanche.'
[Izvestija, 2006.02.08]

(64) Ètot process podoben snežnoj lavine.
This process similar.M snow.DAT avalanche.DAT
'This process is like an avalanche.'
[D. Gluxovsky. 2005. *Metro 2033*]

(65) Takoj vnešnij vid prosto ne podobaet juristu
Such exterior look simply not becomes lawyer.DAT
'Such a look just doesn't befit a lawyer.'
[Forum: *Škol'naja forma: Za i protiv*, 2007-2010]

[36] *podobno* 'like, similar to', *sorazmerno* 'in proportion to', *sootvetstvenno* 'according to', *soobrazno* 'conforming to', but not *otnositel'no* 'regarding' and *soglasno* 'according to'.
[37] All except *blagodarja* 'thanks/owing to'.

Furthermore, the adverbial ‘preposition’, just like the source adjective, only occurs with *j*-forms, as exemplified below.

(66) Učilsja ja podobno **ej**, na odni pjaterki
Studied I similar.N she.DAT$_{[-n]}$ on single.PL fives
‘I studied like her, only getting A’s.’
[Yu.M. Nagibin. 1980-1983. *Školnyj al’bom*]

(67) Naša ljubov’ k čeloveku dolžna byt’ podobna **ej**
Our love to man.DAT must be similar.F she.DAT$_{[-n]}$
‘Our love for man should be like it [our love for God – *T.P.*].’
[Mitropolit Antonij. 1981.*Razgovor s knižnikom*]

Substitution of *j*-forms with *n*-forms renders both these sentences ungrammatical:

(68) *Učilsja ja podobno ***n*ej**, na odni pjaterki
Studied I similar.N she.DAT$_{[+n]}$ on single.PL fives
‘I studied like her, only getting A’s.’

(69) *Naša ljubov’ k čeloveku dolžna byt’ podobna ***n*ej**
Our love to man.DAT must be similar.F she.DAT$_{[+n]}$
‘Our love for man should be like it.’

Based on these data, I propose that these adverbial ‘prepositions’ are just transitive adverbs (or short form adjectives in the neuter – the two are indistinguishable), that is, A-heads. Therefore, we do expect them not to license *n*-forms. Indeed, for none of the four adverbials do we find *n*-forms in the investigated subpart of the Russian National Corpus – only *j*-forms (56 all in all). We should bear in mind though, that in a larger corpus we might have found n-forms, as there seem to be no preposition – *n*-form combinations “that are completely un-googleable”, to

quote Daniel (2015: 96). Given the similarity of the considered adverbs to prepositions – both are indeclinable; transitive adverbs/adjectives assigning lexically specified case are a rare phenomenon; their meaning denotes a *relation* of similarity or correspondence between two objects; their dative object follows the adverb/adjective – it is not unlikely that certain speakers might treat them as P-heads, but this type of grammar is marginal in Contemporary Standard Russian. I suggest that what precludes complete prepositionalization of these adverbs is their relatively low frequency - 1.2 - 1.5 ipm for *sootvetstvenno* 'according to, in correspondence with' and *soobrazno* 'conforming to' (no data for *sorazmerno* 'in proportion to') and around 30 ipm for *podobno* 'like, similar to'.

Let us now turn to gerund 'prepositions'. In Hill (1977) this group includes lexemes having the form of a gerund or a negation particle *ne* (NEG) plus a gerund. My list contains 11 lexemes of the former type and 4 lexemes of the latter type (*ne doxodja*, *ne doezžaja* 'not reaching'; *ne sčitaja* 'not counting/besides', *ne projdja* 'not having passed'). In addition to these 15, we find gerund-simple P complexes (7 items on my list, e.g. *nesmotrja na* 'despite'; *sudja po* 'judging by'). I leave out the latter here, since their behavior with respect to the N-factor is indistinct from that of the simple 'auxiliary' Ps they comprise – namely, they require n-forms.

Let us start with one-word gerund prepositions. Two of them, namely *spustja* 'later' and *pogodja* 'after (waiting)', require temporal complements which cannot be pronominalized: compare (70) and (72) with (71) and (73) respectively. They are therefore irrelevant to the present discussion.

(70) {nedelju spustja | spustja nedelju}
week.ACC later later week.ACC
'a week later'

(71) {*spustja (n)eë | *(n)eë spustja}
later she.ACC[+n/-n] she.ACC[+n/-n] later
'*it later.'

(72) pogodja neskol'ko minut
Waiting few.ACC/NOM minutes.GEN
'after a few minutes'

(73) *pogodja (n)ix
Waiting they.ACC[+n/-n]
'*After them (a few minutes)'

The gerund *prišed* 'having reached (lit. come)', is archaic and does not function as a preposition in the examined subpart of the RNC. *Projdja* 'having passed' and *sčitaja* 'taking into account' are indistinguishable from verbal gerunds in their function and distribution and thus could hardly be considered prepositions. Incidentally, they do occur with 3rd person pronominal complements which unsurprisingly only take the j-form (as verbal objects do).

Of the remaining 6 gerund lexemes of this kind only three occur with pronominal complements in the RNC: the exceptional *blagodarja* 'owing/thanks to' that I consider later; *vključaja* 'including' – 26 documents (n-less forms only) and *isključaja* 'excluding' – 1 document with an n-less form. The INS-assigning *načinaja* 'starting (with)' and *končaja/zakančivaja* 'ending/finishing (with)' do not occur with 3rd third person pronouns in the RNC but appear to only be acceptable with n-less forms, according to my own judgment and data from the Twitter microblogs and Livejournal blogs.

Vključaja 'including' and *isključaja* 'excluding', in both their gerund and prepositional disguise, take complements marked accusative. The corresponding verbs are also associated with accusative case, which appears to be structural (therefore, assigned by the light *v* under standard assumptions) since the direct object in the negated sentence can bear either accusative or genitive case. This genitive, known as the genitive of negation, GenNeg, is commonly considered to alternate only with structural cases, i.e. obey a Non-Obliqueness Restriction (Babby 1978, 1986; Pesetsky 1982). (74) and (77) show that in positive clauses only accusative case-marking is available on the direct object of verbs *vključat'* 'include; turn on'[38] and *isključat'* 'exclude', whereas in negative clauses either accusative or genitive case-marking shows up on their direct objects, as (75-76) and (78-79) demonstrate.

(74) Ja *vklučil* v knigu {pritči | *pritč}.

[38] The second meaning is not a part of the preposition's meaning.

I included into book parables.**ACC** parables.**GEN**

'I included parables in the book.'

[V. Šaxidžanjan. 1999. *1001 vopros pro ETO*]

(75) Opisanie eë on *ne* *vključil* v knigu…

Description.**ACC** she.GEN he NEG included into book

'He didn't include the description of it into the book.'

[O. Trojanovskij. 1997. *Čerez gody i rasstojanija*]

(76) Ètogo ja v donesenie *ne* *vključil.*

This.**GEN** I into report NEG included

'This, I didn't include in the report.'

[I.M. D'jakonov. 1995. *Kniga vospominanij*]

(77) Istočnik *isključil* {verojatnost' | *verojatnosti}

Source excluded probability.**ACC** probability.**GEN**

podobnogo [razvitija sobytij]

such.GEN development.GEN events.GEN

'The source excluded the possibility of such developments.'

[S. Barsukova. RBC daily, 2005.07.01]

(78) Žirinovskij takže *ne* *isključil* verojatnost'

Zhirinovsky also NEG excluded probability.**ACC**

takogo scenarija sobytij

such.GEN scenario.GEN events.GEN

'Zhirinovsky also did not exclude the possibility of such a scenario.'

[A. Vinogradov. RBC daily, 2005.03.18]

(79) [On] *ne* *isključil* verojatnosti takogo scenarija

He NEG excluded probability.**GEN** such.GEN scenario.GEN

'He did not exclude the possibility of such a scenario.'

[V. Teterin. 2014. *Èkspert*]

The accusative/genitive alternation is also found in the corresponding gerund-headed adjuncts as shown in (80-82) and (83-85). True gerund adjuncts are the ones

that denote an eventuality concomitant with or preceding the eventuality of the matrix clause and are subject-oriented, that is, the null external argument of the gerund predicate necessarily co-refers with the matrix subject. They behave similar to temporal gerund adjuncts discussed in Landau (2013: Chapter 6), displaying the Obligatory Control signature. I therefore assume they instantiate non-finite clauses with a PRO subject.

(80) Pisatel' postupil nepravil'no
Writer acted incorrectly
vključiv ètot èpizod v p'esu
including.PRF this.**ACC** episode.**ACC** in play
'It was wrong of the author to include this scene in the play.'

(81) Pisatel' postupil nepravil'no
Writer acted incorrectly
ne *vključiv* ètot èpizod v p'esu
NEG including.PRF this.**ACC** episode.**ACC** in play
'It was wrong of the author not to include this scene in the play.'

(82) Pisatel' postupil nepravil'no
Writer acted incorrectly
ne *vključiv* ètogo èpizoda v p'esu
NEG including.PRF this.**GEN** episode.**GEN** in play
'It was wrong of the author not to include this scene in the play.'

(83) Žjuri sprovocirovalo skandal *isključiv* ètot film
Jury provoked scandal excluding.PRF this.**ACC** film.**ACC**
iz programmy festivalja
from program festival
'The jury provoked a scandal, having excluded this film from the festival's program.'

(84) Žjuri sprovocirovalo skandal (tak i) *ne* *isključiv*
Jury provoked scandal so and NEG excluding.PRF
ètot film iz programmy festivalja
this.**ACC** film.**ACC** from program festival
'The jury provoked a scandal, having failed to exclude this film from the festival's program.'

(85) ?Žjuri sprovocirovalo skandal (tak i) *ne* *isključiv*
Jury provoked scandal so and NEG excluding.PRF
ètogo filma iz programmy festivalja
this.**GEN** film.**GEN** from program festival
'The jury provoked a scandal, having failed to exclude this film from the festival's program.'

More surprisingly, the gerunds *vključaja* 'including' and *isključaja* 'excluding' show a similar pattern even when used in the 'prepositional' sense (i.e. with the meaning of 'including' and 'except', typically modifying/specifying a noun phrase rather than introducing an additional event). Like the 'true' gerunds, they require accusative marking on their object in the regular case, as shown in (86) and (89) and display the genitive/accusative alternation when a negation particle is attached to the gerund phrase, as the naturally occurring examples in (87-88) and (90-91) demonstrate.

(86) My zainteresovany v priobretenii aktivov na vsej
We interested in acquisition assets.GEN on whole
territorii byvšego SSSR *isključaja* {strany|*stran} Baltii
area former USSR excluding countries.**ACC|*GEN** Baltia
'We are interested in acquiring assets all around the former USSR except the Baltic states.'

[S. Zaiceva. RBC daily, 2006.04.05]

(87) Ljubaja strana…*ne* *isključaja* samyx bol'šix i moščnyx
Any country NEG excluding most.**GEN** big.**GEN** and powerful.**GEN**
zavisit ot svoego geografičeskogo položenija
depends from self's geographic position

‘Any country, not excluding the biggest and strongest ones, depends on its geographic location.’

[S. Artobolevskij et al. Prostranstvo i razvitie Rossii…*Vestnik RAN*, 2009]

(88) Nevežestva konečno xvatalo... kak i vezde
Ignorance of course sufficed as and everywhere
(*ne* *isključaja* samye prosveščennye strany)
NEG excluding most.**ACC** enlighted.**ACC** countries.**ACC**
‘There was enough ignorance of course, just like everywhere (the most progressive countries included).’

[I. Volgin. 2005. Sokrušaemyj Karfagen. *Oktjabr’*]

(89) V [èti] otrasli… planiruetsja vložit’ okolo 4 trln rub,
In these sectors plan.REFL invest around 4 trillion roubles
vključaja {rasxody | *rasxodov} na NIOKR.
including expenses.ACC/*GEN on R&D
‘4 trillion roubles, including the R&D expenses, are planned to be invested in these sectors.

[RBC daily, 2007.11.22]

(90) Kontrol’ nad biznesom... byl ocenen v 20 mlrd rub,
Control over business was estimated in 20 bln rubles
ne *vključaja* dolg seti
NEG including debt.**ACC** chain.GEN
‘Taking control over the business was estimated to require 20 billion roubles not including the chain’s debt.’

[N. Kopeičenko & E. Perceva RBC daily, 2011.02.04]

(91) Vnedrenie… novoj koncepcii potrebuet…4.6 mln. doll. investicij,
Introduction new concept will-require 4.6. mln doll investment
ne *vključaja* rasxodov na… reklamu
NEG including expenses.**GEN** on advertising
‘Introducing the new concept will require $4.6 million in investment, not including the advertising costs.’

[E. Želobanova RBC daily 2005.10.20]

The availability of the genitive/accusative alternation in the examples above provides evidence in support of the verbal status of 'prepositions' *vključaja* 'including' and *isključaja* 'excluding, except'. Why is that so? If the ACC/GenNeg alternation is indeed only possible when the accusative is structural and if structural accusative is licensed by the light *v*, it follows that the gerund 'prepositions' are of the category V. And if they retain a V-head status we do not expect n-forms to be licensed with them. Indeed, n-forms are neither found with them in the investigated subpart of the RNC, nor are they possible in my idiolect. Again, few occurrences of *n*-forms with *vključaja* 'including' and *isključaja* 'excluding, except' can be found in Twitter microblogs and in Livejournal blogs, as well as elsewhere on the web. Notably, all these occurrences involve the prepositional function of these gerunds, which suggests that individual grammars treating them as P-heads might exist (although they are clearly substandard).

There seems to be a problem with this reasoning, however: *v*P is generally considered to introduce an external argument of a verb and is thus projected with unergative and transitive verbs, but not with unaccusative verbs. A true gerund *vključaja* 'including' or *isključaja* 'excluding' projecting a temporal adjunct phrase that denotes an event occurring concomitantly with or preceding the event of the matrix clause (depending on the aspect morphology) requires the understood agent of this event to co-refer with the subject of the matrix clause. The existence of such an understood agent gives us some grounds to introduce at least a *v*P-layer in the structure that would involve a PRO in the subject position, which would allow us to understand why the accusative/genitive case alternation is observed in gerund phrases as well as in matrix clauses. However, positing anything more extended than a VP and, actually even a VP itself, is problematic for the 'prepositional' cases, since no event or agent is involved, and the phrase does not seem to be predicated of any DP in the matrix clause. Should we then say that *vključaja* 'including' or *isključaja* 'excluding' in (87) and (90) are actually prepositionalized items, and the accusative case of their complements has now been reanalyzed as an instance of lexical case that these Ps are lexically specified for? This looks like the right analysis for those rare speakers who produce n-forms with these gerund prepositions. Notably, none of these marginal occurrences involves a preceding Neg particle (which is suggestive but does not imply its unacceptability), but if we had access to such speakers or sentences we would expect Neg there to be treated as constituent negation, not inducing genitive of negation on the object. In my idiolect, however, n-forms are illicit with *vključaja*

'including', *isključaja* 'excluding' in either gerund or prepositional function and the genitive/accusative alternation obtains when a negative particle is added. Both facts speak against a P-head analysis for them. How do we make sense of all this then?

I propose that the gerunds develop into prepositions in the following way: they start as non-finite adjunct clauses with PRO-subjects obligatorily controlled by the matrix subject:

(92) Vključaja ego v naši rjady my pomogaem emu perevospitat'sja
Including him in our ranks we help him.DAT re-bring-up.REFL
'(Through) including him in our ranks, we are helping him rebuild himself.'

Next, they evolve into clauses with non-obligatorily controlled PRO-subject that may have arbitrary [+human] reference or be controlled by e.g. an implicit [+human] argument in the matrix clause. Finally, they become genuine PP-structures, lacking any external argument. So far, the only gerund preposition that is reaching this final state is *blagodarja* 'thanks to/owing to' to be discussed in Section 3.5. Other gerund 'prepositions', in particular, *vključaja* 'including' and *isključaja* 'excluding' have an intermediate status – they project non-finite adjunct clauses with a non-obligatorily controlled [+human] PRO subject.

(93) **P-Lexicalization path for gerunds**: Head of a non-finite adjunct clause (PRO is OC-ed)→ Head of a non-finite adjunct clause (PRO is NOC-ed) → P-head

The instrumental case assigned by preposition-like gerunds *načinaja* 'starting (with)' and *končaja/zakančivaja* 'ending (with)' is also assigned in environments with truly verbal instances of these lexemes: the instrumental-marked DPs in these cases however are not obligatory – they indeed introduce an optional instrument DP. Since instrument-denoting INSTR-marked DPs may optionally be introduced in many sentences of Russian, never appearing in the *n*-form when pronominal, we do not expect to find *n*-forms when *načinaja* and *končaja/zakančivaja* are used prepositionally either. The most adequate analysis for these INSTR-assigning gerund prepositions would be not to consider them prepositions at all and treat them as verbal elements (gerunds projecting a non-finite clause with NOC PRO) or frozen idioms: after all, these only function prepositionally in one construction, *načinaja X.INS i končaja/zakančivaja Y.INS* 'from X to Y [lit. 'starting (from) X and finishing (with) Y'] – that is, these antonyms have to be always used together. Interestingly, *načinaja*

can be used independently and function as a preposition, but in that case, it requires a simple auxiliary P *s* 'from' and the gerund+P complex has the meaning 'starting/beginning from'. This supports the analysis of instrumental case found with *načinaja* and *končaja/zakančivaja* as licensed by whatever mechanism licenses it[39] on a non-subcategorized instrument DP in general or as fixed in the lexicon as part of the idiomatic expression. Instrumental case assignment is not a lexical property of these gerund lexemes, which might be one of the reasons precluding lexicalization into P-heads assigning INS: they resist it so much that no reliable instances involving n-forms can be found.

Finally, let us examine the gerund prepositions of the NEG+gerund type, those that apparently lexicalize the negation particle (suggested, in particular, by the widespread erroneous one-word spelling of these items). Very little data have been found for them in the investigated part of the Russian National Corpus: 1 occurrence of a *j*-form with *ne doezžaja* 'on the way to (lit. 'not arriving (at)') and 2 occurrences of a *j*-form with *ne sčitaja* 'excluding (lit. – 'not counting')' with *n*-forms not attested in the corpus with these 'prepositions'. Such combinations as *ne doxodja* 'on the way to' (lit. 'not reaching (some place) by foot') and *ne projdja* 'not having passed' were not found with 3rd person pronominal complements in the corpus at all.

The meaning of the gerund in these cases is rather bleached, similarly to the gerund+simple P complexes like *nesmotrja na* 'despite (lit. 'not looking at'). Since their complements bear genitive case and the corresponding gerunds are not genitive case-assigners, it is reasonable to assume that this case is inherited/comes from the NEG particle. Notably, *ne doezžaja*, derived from *ne doezžat'/doexat'* has different selectional requirements: it takes a bare GEN complement, while its (modern) verbal counterpart requires a PP-complement headed by *do* 'to, up to, till'. So, in this case the NEG particle somehow can only assign GenNeg when *ne doezžaja* is used prepositionally.

3.4.2. P-N combinations associated with dative arguments

We now turn to the last remaining group of prepositions in our data set – prefixed/combination 'prepositions' associated with dative case government, i.e. those whose notional complement DPs bear dative morphology.

[39] Possibly, the default VP-internal case of Pereltsvaig 2007.

Within this group two clear semantic classes of lexemes can be distinguished:

1) 'prepositions' introducing the Goal of motion:

    - *navstreču* 'towards (an object moving in the opposite direction)'
    - *v(o)sled, vdogonku* 'following (a moving object)'
    - *napererez* 'crossing the trajectory (of a moving object)/cutting across to'

2) 'prepositions' typically introducing a Maleficiary or a Beneficiary (Experiencer) into the event structure:

    - *vopreki* 'in spite'[40]
    - *naperekor* 'to spite, against the will of'
    - *nazlo* 'to spite'
    - *v piku* 'to spite'
    - *v otmestku* 'in revenge'
    - *na radost'* 'to the delight of'

The remaining 6 prepositions *v protivopoložnost'* 'contrary to/as opposed to', *ne v primer* 'unlike', *v protivoves* 'in contrast to (lit. 'as a counterweight')', *na smenu/vzamen* 'as a substitution to/to replace', *srodni* 'akin to' form a heterogeneous group.

A central question regarding these 'prepositions' is the locus of dative case assignment. The simple P-morphemes that these complex lexemes comprise do not govern dative case and only some of the independent nominal correlates of the internal nominal morpheme may be argued to license dative case.

The 16 'prepositions' differ in whether the nominal morpheme they comprise is an independently active noun and if such a noun is associated with dative DPs (possessor datives).

Only 2 'prepositions' of the first group have a clear noun inside – *vstreča* 'meeting' in *navstreču* 'towards a moving object' and *sled* 'footstep, trace' in *v(o)sled* 'following', but these nouns do not occur with associated dative arguments, either preceding or following them:

[40] From *preky/pereki* 'against', generally regarded as an old instrumental plural form of the nominal *prěk"* (Hill 1977:134). These lexemes showed adpositional use in Old Russian but no longer exist.

(94) *{emu} sled {emu}
he.DAT[-n] footstep he.DAT[-n]
‘*Him footstep;*footstep him.’

(95) *{emu} vstreča {emu}
he.DAT[-n] meeting he.DAT[-n]
‘*Him meeting; meeting him.’

However, sentences involving verbs of the same root do license dative DPs, which might suggest thematic licensing.

(96) Ona vo vsem sledovala emu.
She in everything followed he.DAT
‘She followed his example in everything.’

(97) Ona vstretilas’ emu po doroge domoj.
She meet.F.PST.REFL he.DAT along road.DAT home
‘He came across her on his way home.’

There are four such prepositions in the second group:

- *zlo* ‘evil’ > *nazlo* ‘to spite’;
- *pika* ‘pike’ > *v piku* ‘to spite’;
- *otmestka* ‘revenge’ > *v otmestku* ‘in revenge to’;
- *radost’* ‘joy’ > *na radost’* ‘to the delight of’;

Only the latter two involve nominals that can be associated with dative arguments, however

(98) Èto – moja otmestka emu
This my revenge he.DAT
‘This is my revenge to him.’

(99) Moja pečal – emu radost’.
My sorrow.NOM he.DAT joy.NOM
‘My sorrow brings joy to him.’

Although, compare:

(100) Ja pričinil emu mnogo zla
I did he.DAT much evil
'I have done him much evil.'

Notably, the third, heterogeneous, group comprises lexemes and collocations whose nominal morphemes do independently occur with associated dative DPs, as shown below:

(101) Ona - polnaja emu *protivopoložnost'*
She.NOM complete he.DAT opposite.NOM
'She is his complete opposite.'

(102) Pioner – vsem rebjatam *primer*
Pioneer.NOM all.DAT guys.DAT example.NOM
'A scout is an example for all guys.'

(103) My dolžny sostavit' emu *protivoves*
We.NOM must provide he.DAT counterbalance.NOM
'We must present a counterbalance to him.'

(104) Vot i podrosla nam *smena*!
Here and grew-up we.DAT replacement.NOM
'The younger generation has grown up to replace us.'

(105) My uže našli emu *zamenu*
We.NOM already found he.DAT substitution.NOM
'We already found a replacement for him'

(106) Oni mne *rodnja* po materi
They.NOM I.DAT family along mother.DAT
'They are my relatives on my mother's side.'

The reason is perhaps their more compositional (less lexicalized) nature: it might well be the case that those 'prepositions' whose nominal morphemes are not associated with dative DPs used to behave in the same way when they were only undergoing lexicalization (with the P-head). This idea is supported by the fact that *vstreču* + DP_{DAT} *'to meet'*, *pereky* + DP_{DAT} 'against' were attested at the earlier stages of Russian (cf. Hill 1977: Chapter 6). I tentatively suggest that the dative case associated with the third group of prepositions is licensed by an adnominal Appl(icative) head.

In addition to dative case on their notional complement, there are several properties that sets this group of 17 prepositions apart from all the prepositions considered earlier:

1) Resistance to taking the *n*-forms, even by fully lexicalized P-N complexes (zero occurrences in the RNC, only paucal occurrences in Twitter and Livejournal)

2) Ability to appear as a postposition (except maybe *ne v primer* and *v protivipoloznost':* this seems to depend on how phonologically heavy the 'preposition' is):

(i) Ja bežal emu navstreču
I.NOM ran he.DAT on-meeting
'I was running towards him.'

3) Strandability:

(ii) Komu ty bežal navstreču?
who.DAT you.NOM ran on-meeting
'?Who were you running towards?'

Note that apart from that group there are practically no postpositions in Russian (except maybe …*spustja* 'later', *…nazad* 'ago/before'; *radi* 'for the sake of' is also cited as a postposition in e.g. Podobryaev 2009 but its postpositional use is actually restricted to idioms – Itkin and Tolkacheva 2015 – and it cannot be stranded). Likewise, P-stranding by *wh*-movement is by no means a characteristic property of the

prototypical Russian prepositions. Note also that only DAT-assigning postpositions can be stranded.

It is tempting to link the obligatory n-less form taking and postpositionality/strandability to explain the extraordinary behavior of this set of prepositions. How can it be done?

Recall that the appearance of n-forms in our analysis is licensed by a governing P-head. Therefore, there are two options to account for the current group of prepositions:

a) posit that they are not P-heads, but heads of a different category that may license dative case
b) say the DAT on their notional complement is not assigned by the P-N lexeme at all and is licensed through some other mechanism

Proposal. both explanations are relevant, reflecting the two etymological sources of these 'prepositions'.

Source 1: a transitive noun (or *root*, more generally) with lexically (inherently) specified dative case government:

- *otmestka* + DP_{DAT} 'revenge to smb' > *v otmestku*+ DP_{DAT} 'in revenge to smb'; cf. *otomstit' emu* 'to revenge him';
- *protivopoloznost'*+ DP_{DAT} 'the opposite of smb/smth' > *v protivopoloznost'*+ DP_{DAT} 'as opposed to smb/smth';
- *protivoves* + DP_{DAT} 'counterweight to smth/smb' > *v protivoves* + DP_{DAT} 'in contrast to/to balance out smth/smb';
- *?rodnja* + DP_{DAT} 'relatives, kin to smb' > *srodni* + DP_{DAT} 'akin to smb/smth' Cf. *rodstvennyj emu* 'related to him';
- **preky/pereki* + DP_{DAT} 'against smb' > *naperekor, vopreki, poperek* + DP_{DAT} 'in spite/to spite/against the will of smb'. Cf. *perečit' emu* 'to go against his will'.

These dative DPs therefore bear inherent, non-prepositional case.

Source 2: an adnominal applicative construction (in the sense of Pšexotskaja 2012), where the dative DP fills the Spec, ApplP position and the prefixed/combination 'preposition' is a PP which the Applicative is merged to.

(107) [$_{ApplP}$ [$_{Spec,\ Appl}$ DP_{dat} [$_{Appl'}$ $\emptyset_{Appl}$ [$_{PP}$ na smenu/vzamen/na radost'/nazlo/v piku/ne v primer]]]]

- in this structure DAT is clearly not assigned by a P-head, hence we predict that *n*-forms will not be licensed.

- the unusual DP_{dat} – *'preposition'* word order and the strandability of these 'prepositions' is derived: nothing should preclude *wh*-movement of the dative DP from Spec, ApplP to Spec, CP. Movement of dative DP + postposition is also predicted to be possible since they form a constituent.

It might seem trickier to derive the *preposition-dative DP* order, but this might be governed by some further phonological re-ordering/scrambling. Alternatively, the P-DAT order is a newer one, emerging as these combinations become more frequent leading to their prepositionalization: the *n*-form is still not used, because both orders are possible and the prepositionalization is therefore not complete.

Lexicalization paths:

From a PP in the Appl-construction to a true dative-assigning P. As specific P-N combinations begin to occur exceptionally frequently in the applicative structure, they get more and more lexicalized → the dative DP begins to be associated with the 'preposition' → is placed to the right rather than to the left of the P-N combination → reanalysis of the PP as a P-head governing DAT case assignment.

This prepositionalization process, however, is not complete in the standard language yet: it might be that as long as the *dative DP - 'preposition'* ordering remains productive, complete prepositionalization is precluded. This is what explains the lack of *n*-forms on the notional complements of these 'prepositions' in Standard Russian.

From a PP with inherent-DAT-assigning noun to a DAT-assigning P. The explanation for the other set of lexemes, whereby DAT-assignment is the lexical property of the internal nominal morpheme is even simpler. In fact, it is parallel to the one for prefixed/combination prepositions associated with GEN-assignment. As the preposition-noun collocation gets lexicalized, the emerging P-head inherits the properties of its components – case government of the nominal element (here – inherent dative case) and the P-Case feature of the simple prepositional morpheme. When the prepositionalization process is complete, the new, morphologically complex

P-head should be able to license *n*-forms on the complement to which it assigns dative case.

3.5. Exceptions

In the section on distribution of n-forms and n-less forms across prepositions I formulated a generalization, which I reproduce below:

(108) Prepositions whose morphological composition transparently reflects their origin (i.e. prefixed/combination, adverbial, gerund and certain suffixed prepositions) license *n*-forms iff they govern genitive case.

The sections on the groups of prepositions that are not (necessarily) P-heads have shown why this generalization obtains. This generalization, however, is not ideal since two prepositions, the dative-assigning gerund-based *blagodarja* 'thanks to' and the dative-assigning adverbial *soglasno* 'according to' do marginally occur with *n*-forms (in less than 3% of all cases). Informal judgment queries suggest that not all speakers allow n-forms with these two.

I suggest that the marginal *n*-form taking ability of *blagodarja* and *soglasno* distinguishing them from other non-GEN-assigning morphologically transparent prepositions is due to several factors.

First, these lexemes are strictly *pre*positional and do not admit positioning of the dative DP to the left, unlike those just discussed in Section 3.4.2.

Second, when used prepositionally they must occur with an overt dative argument, that is, they cannot be used intransitively/adverbially and cannot have an implicit complement.

Third, their case-governing properties are synchronically distinct from the related gerund and adverb/short adjective respectively. Let us elaborate on that point. The prepositional instance of *blagodarja* requires a dative DP-complement whereas the true gerundival *blagodarja* 'thanking' in Contemporary Russian requires a direct object that is marked **accusative**, just like other instances of the verb *blagodarit'* 'to thank', although in the 18th century all forms of that verb were attested with either accusative[41] or dative-marked complements (my study of the relevant subpart of the

[41] The accusative on the internal argument of *blagodarit'* 'to thank' is apparently inherent (it does not seem to allow Genitive of Negation).

RNC). It might well be the case that the proliferation of the metaphoric (prepositional) use of the gerund was a factor in the change of the case-government of the verb. Note that the related noun *blagodarnost'* and adjective *blagodarnyj* also take a dative complement, whose case is probably thematically licensed, that is, is inherent. I suggest that those speakers who grammatically associate the 'prepositional' instances of gerund *blagodarja* with the noun *blagodarnost'* and adjective *blagodarnyj* treat the dative case on their complement DP as an instance of inherent, theta-related case, rather than P-Case (i.e. *blagodarja* + DAT is a NOC gerund, a bleached verbal category rather than a P-head) and therefore do not accept *n*-forms.

On the other hand, those speakers who clearly distinguish the verbal and the prepositional instances of *blagodarja* are likely to treat the dative case associated with the latter as a special, lexically specified property of the P-head *blagodarja*. These speakers are predicted to license *n*-forms in its dative complement.

Turning now to preposition *soglasno* 'according to', while I managed to find one instance of the related adjective *soglasnyj* 'agreeing (with)' with a dative complement, such sentences are archaic and impossible in my idiolect, for instance Rather, the verb, noun and adjective based on the same stem now all require a PP-complement headed by the simple preposition *s* 'with':

(109) Soglasie |soglašat'sja| soglasnyj {*emu| *nemu| s nim}
Consent |to agree| agreeing he.DAT$_{[-n]/[+n]}$ with he.INS$_{[+n]}$
'Agreement/to agree/to be in agreement {*to him | with him}.

Thus, dative case government is again specifically associated with the prepositional instance of *soglasno* 'according to'. I conjecture that its semantic affinity to the simple dative-assigning preposition *po* 'along' might have boosted its prepositionalization with the retainment of the archaic dative case government.

The third property that sets *blagodarja* and *soglasno* apart from other complex dative-assigning 'prepositions' and unites them with core prepositions is their remarkable frequency. Lyashevskaya and Sharoff (2009) give the figures of 94.5 and 91.9 ipm, making them the 1293rd and the 1335th most frequent Russian lexeme respectively. This is the highest frequency attested for lexemes of the 4 "derived" classes (prefixed, combination, adverbial, gerund): other non-GEN-assigners range between 1.21 (*sootvetstvenno* 'according to') and 58.6 (*vključaja* 'including') ipm. Comparable in frequency are such basic prepositions as the root preposition *skvoz'*

'through' – 110.5 ipm; suffixed prepositions *radi* 'for the sake of' (109.2 ipm) and *mimo* 'past' (75.9 ipm); prefixed prepositions *vozle* 'near, close to' (111 ipm) and *vrode* 'like' (78.5 ipm) and the double preposition *iz-pod* 'from under' (86.8 ipm) – notably, all belonging to Class I, i.e. P-heads, requiring n-forms. Note that *blagodarja* and *soglasno* turn out to be the only non-Class I prepositions in this frequency range and are in fact even more frequently used than the least frequent Class I prepositions such as *iz-pod* 'from under', *vrode* 'like' and *mimo* 'past'.

To sum up, *blagodarja* 'owing to' and *soglasno* 'according to' have the potential to become fully lexicalized as prepositions in the near future. As the marginal but sizeable number of *n*-forms occurrences demonstrate, their prepositionalization is very much under way even in the standard variety of Contemporary Russian. The factors hindering their complete prepositionalization are a) dative case assignment (not typical for basic prepositions but very typical for morphologically complex ambivalent adpositions considered in the previous subsections) and b) the transparent etymological source, verbal and adverbial respectively.

Finally, there is one additional (adverbial) preposition that deserves mention, namely *otnositel'no* 'with regard to'. Unlike all other adverb-based prepositions it is associated with genitive, rather than dative case assignment. Furthermore, the adverb/adjective it derives from as well as the corresponding verb/deverbal noun do not take a DP-complement directly, but require a PP introduced by the simple dative-assigning preposition *k* 'to, towards':

(110) Delo *otnositel'no* ètogo čeloveka
Case relating this.GEN man.GEN
'a case concerning this man'

(111) *Delo *otnositel'no* (k) ètomu čeloveku
Case relating to this.DAT man.DAT
'a case relating to this man'

(112) Delo *po* *otnošeniju* {k ètomu čeloveku | *ètogo čeloveka}
Case with respect to this.DAT man.DAT this.GEN man.GEN
'a case with respect to this man'

(113) Delo otnositsja {k ètomu čeloveku | *ètogo čeloveka}
Case relates.REFL to this.DAT man.DAT this.GEN man.GEN
'The case concerns this man.'

Where does this genitive assignment come from then? The 18th century data on *otnositel'no* from the Russian National Corpus include instances like *otnositel'no* + DP_{GEN}, *otnositel'no do* + DP_{GEN} 'concerning smb/smth' and *otnositel'no k* + DP_{DAT} 'with respect to smb/smth'. In the 19th century the modern variant without an auxiliary P is already predominant and n-forms become licit in the genitive complement on a par with n-less forms (n-forms: n-less forms = 16:86). The prepositionalization of *otnositel'no* progresses in 1900-1980, with n-forms and n-less forms attested equally frequently (n-forms: n-less forms = 31:28) and virtually completes in the 1980-2015 period, in which the proportion is already 45:10, where only 8 instances of n-less forms are genuinely modern, all coming from the Main corpus (predominantly consisting of fiction), which might be affected by stylistic considerations.

I propose that *otnositel'no* 'regarding' borrowed the genitive case government from the auxiliary simple preposition *do* '(up) to; till' it used to occur with in the 18th-19th century. The prepositionalization of *otnositel'no* might also have been boosted by the influence of the semantically close sub-standard preposition *kasatel'no* 'concerning (lit. touching)', whose GEN-assignment properties are lexically specified, and probably the related combination preposition *v otnošenii* +DP_{GEN} 'in relation, with respect to' (see Hill 1977). In any event, *otnositel'no* was prepositionalized already in the 19th century, thanks to the peculiar, unpredictable case government properties as well as lack of a corresponding transitive adjective/adverb. I therefore treat it as a P-head. The minor deviations from the constraint on n-less forms that it exhibits might be attributed to dialectal, idiolectal and stylistic variation.

4. *N*-forms and *j*-forms in a coordinated P-complement. Fine-tuning of the analysis

4.1. Corpus data

It has been observed in the literature, especially in Blažev (1962), Hill (1977) and Daniel (2015), that a third person pronominal P-complement has a weaker tendency to take the n-form if separated from the P by other material. Specifically, two such environments have been discussed: instances with an interposed modifier and the case of preposition *meždu* 'between' with the complement instantiated by two coordinated DPs, exemplified by (114) and (115) respectively.

(114) Maša vljubilas' vo vsex {?ix | ✓**n**ix}
Maša fell-in-love in all.ACC 3PL.ACC**[-n]** 3PL.ACC**[+n]**
'Masha fell in love with all of them.'

(115) Meždu Mašej i {✓im | ✓nim}
Between Maša.INS and 3MSG.INS**[-n]** 3MSG.INS**[+n]**
ustanovilis' družeskie otnošenija
set-up friendly relations
'Masha and he became friends.'

In addition, Kholodilova (2013) observes that it is almost impossible for pronouns to occur as a second conjoined complement of a preposition, i.e. in the [P [X & pron]] configuration with the possible exception of preposition *meždu* 'between' that accepts pronouns in second position more freely. However, Kholodilova only considers Russian 1st and 2nd person pronouns, noting that the coordination of 3rd person pronouns is further complicated by the *N*-issue.

This and the following section focus on sentences like (115), involving various prepositions and not just *meždu* 'between'. In this section I register and discuss the few instances of n-forms and j-forms in the second conjunct position attested in the investigated subcorpus of the RNC. I show that the scarcity of corpus examples does not enable establishing the grammatical status of sentences of this type. The next sections discuss the results of an acceptability judgment survey probing native speakers' intuition concerning these infrequent sentences.

It is easy to see that in the RNC *n*-forms occur in the second conjunct position very rarely. The search for *{i; ili} {nego; neë; nix…}* '{and; or} {n-form}' returned very few results:

(116) *Krome Vladimira Viktoroviča i **n**eë*
Except Vladimir Viktorovich.GEN and she.GEN**[+n]**
'Except Vladimir Viktorovich and her'

(117) *Dlja nego ili **n**eë*
For he.GEN**[+n]** or she.GEN**[+n]**
'for him or her'

(118) *Meždu {DP; 1,2 pron}* *i* *{**n**im; **n**ej(u); **n**imi}*
Between {DP; 1,2 pron}$_{INS}$ and {he; she; they}$_{INS}$**[+n]**
'Between X/me/us/you and him/her/them.'

There were 17 instances of type (118) with preposition *meždu* 'between' and only one with *krome* 'except' and *dlja* 'for'. Based on such scarce data we cannot say much about the status of n-forms in the second conjunct position: whereas (116) and (117) might be performance errors, (118) suggests a regular pattern. Given that the semantics of preposition *meždu* 'between' is special in that it requires a semantically plural internal argument, either a plural DP or two conjoined DPs, it is conceivable that this PP-internal coordination may have a distinct structure, which would make this P exceptional in allowing n-forms in the second conjunct. Alternatively, it might be that all three prepositions are exceptional and that the scarcity of examples for *krome* 'except' and *dlja* 'for' is a consequence of the limited corpus size and their relatively less frequent occurrence with a conjoined complement in general. Finally, it is also possible that there is no constraint on the n-form in the second conjunct position and what we see is again due to relative infrequency of these combinations. In particular, *n*-forms may occur relatively infrequently in the second conjunct position due to the fact that there exist alternative, perhaps, more natural strategies to express roughly the same meaning:

a) using a plural pronoun;

b) placing the 3rd person pronoun in the first conjunct position;

c) employing the comitative strategy – P DP/pron with *n*-form (not always available for *meždu* ‘between’);

d) repeating the preposition after the conjunction (not available for all Ps and not always yielding the same interpretation);

e) using a proper name/other full DP instead of the pronoun.

These strategies are illustrated below as alternatives to (116).

(119)

a. *Krome **n**ix*
Except they.GEN**[+n]**
‘Except them’

b. *Krome **n**eë* *i* *Vladimira Viktoroviča*
Except she.GEN**[+n]** and V.V.GEN
‘Except her and Vladimir Viktorovich’

c. *Krome **n**ix* *s* *Vladimirom Viktorovičem*
Except they.GEN**[+n]** with V.V.INS
‘*Except them with Vladimir Viktorovich’

d. *Krome Vladimira Viktoroviča* *i* *krome **n**eë*
Except Vladimir Viktorovich.GEN and except she.GEN**[+n]**
‘Except Vladimir Viktorovich and except her’.’

e. *Krome Vladimira Viktoroviča* *i* ***Marii Ivanovny***
Except Vladimir Viktorovich.GEN and Maria Ivanovna.GEN
‘Except Vladimir Viktorovich and Maria Ivanovna.’

The strategy that is minimally different from the original one is (119b), that is, putting the 3rd person pronoun in the first conjunct, adjacently to P. Note that all the alternative strategies avoid the use of an *n*-form non-adjacently to P. The sentences of (119b) type abound in the corpus, so their number was not registered. We thus conclude that *n*-forms are good in the first conjunct but may be bad in the second conjunct.

A natural question to ask at this point is whether *j*-forms behave differently. It would seem that they do: we find few examples in the corpus, but they involve a wider

range of prepositions. For instance, *pered* 'before, in front of', *pomimo* 'besides' and *nasčët* 'regarding' never occur with an n-form in the second conjunct position but each occurs once with a *j*-form in that position. The example with *pomimo* is particularly notable as it involves the coordination of a [+n] and a [-n] form. *Krome* 'except' and *meždu* 'between', the apparently 'exceptional' prepositions that allow *n*-forms in the 2nd position, also allow and possibly favor n-less forms in that position: thus, 5 occurrences of [P X & n-less form] were found for *krome* and 43 such occurrences for *meždu*, compared to 1 and 17 occurrences of *n*-forms respectively.

To summarize, the corpus data for preposition *meždu* 'between' give us the following picture. Both forms in principle can appear in either conjunct but:

(1) *N*-forms are much more frequent than *j*-forms in the 1st conjunct position
(2) *N*-forms are less frequent than *j*-forms in the 2nd conjunct position

Of the 4 possible combinations of 3rd person pronouns only [+n] & [-n] and [+n] & [+n] are attested, the former being more prevalent.

Based on the data presented above, we can formulate the following tentative generalizations:

(120) Most basic, Class I prepositions allow neither *n*-forms nor *j*-forms in the second coordinated complement; some may only exceptionally allow *j*-forms (*pered* 'before', 'in front of').

(121) When two 3rd person pronouns are coordinated, 'n' can appear either on the adjacent conjunct or on both conjuncts; two other logically possible combinations are not found for Class I/Class II prepositions.[42]

However, the empirical basis of these generalizations is very poor. To recap, the relevant data come mostly from preposition *meždu,* so we cannot conclude whether it is due to its possible structural exceptionality mirroring its special semantics or merely a consequence of its occurring more frequently with a coordinated complement. Therefore, to test these generalizations, we need to elicit native speakers' judgments concerning the acceptability of the two forms of pronouns in the second complement position of different prepositions.

[42] Note though that the [+n] & [+n] combination is only found for *meždu* 'between' (and *dlja* 'for' with disjunction), so it may be a quirk of preposition *meždu.*

4.2. Acceptability judgment survey

4.2.1. Questionnaire make-up

4.2.1.1.Choice of prepositions

In order to test the corpus-based generalizations about the legitimate positions of n- and n-less forms within the conjoined complement of Russian prepositions, a representative sample of prepositions to be used in the target sentences had to be obtained. Several factors lay behind the choice of the 10 prepositions provided in Table 9:

- Preposition *can* take an n-form (belongs to Classes I-II);
- Preposition takes a 3rd person pronominal complement frequently enough (> 100 examples in the RNC, cf. Table A1 of the Appendix);
- Preposition can take a [+human] complement (to ensure its felicitous coordination with proper name DPs);
- Preposition may head a sentence-final PP;
- The potentially exceptional prepositions *krome* 'except', *mežду* 'between', *pered* 'in front of' and *dlja* 'for' are included;

Given that for prepositions *krome* 'except', *dlja* 'for' and *pered* 'in front of' there exist synonymous prepositions *pomimo* 'besides', *radi* 'for the sake of' and *vperedi* 'in front of' that may exhibit a different behavior with 3rd person pronouns, it was decided to include them for comparison. Along the same lines, another Class I/Class II minimal pair, *za/pozadi* 'behind' was added to the list. In addition to their having Class II minimal pairs, *za* and *pered* are similar to the potentially exceptional preposition *mežду* in that they are Class I locative Ps governing instrumental case; thus, it is also instructive to compare them to each other. Since there is no preposition in Russian that is semantically close to *mežду* and belongs to a different class, a frequent Class II preposition *naprotiv* 'opposite' was 'paired' with it. Table 9 presents the 5 selected pairs.

Table 9 Prepositions used in the target sentences[43]

'Class I'	**'Class II'**
Krome 'except' (GEN)	*Pomimo* 'besides' (GEN)
dlja 'for'(GEN)	*Radi* 'for the sake of'[44] (GEN)
Za 'after, behind' (INS)	*Pozadi* 'behind' (GEN)
Pered 'before, in front of' (INS)	*Vperedi* 'in front of' (GEN)
Meždu 'between' (INS)	*Naprotiv* 'opposite' (GEN)

4.2.1.2.Target sentences

The structure of the target sentences was dictated by the aim to test the respective acceptability of n- and n-less forms in the second conjunct in the complement of various Class I and Class II prepositions.

It was also decided to test whether the type of the nominal in the first conjunct within the PP affects the acceptability of n- and n-less forms in the second conjunct. Thus, for each preposition 8 similar sentences were composed. These sentences involved combinations of 4 types of noun phrase: proper names (coded as DP); second person singular pronoun (2SG); n-form of 3rd person pronouns, [+n]; n-less form of 3rd person pronouns, [-n]. These are presented as minimal pairs in Table 10.

Table 10. Types of coordinated P-complements

DP & [+n]	DP & [-n]
2SG & [+n]	2SG & [-n]
[-n] & [+n]	[-n] & [-n]
[+n] & [+n]	[+n] & [-n]

[43] Abbreviations in the parentheses indicate the morphological case of the P-complement.

[44] *Radi* 'for the sake of' was classified as Class I rather than Class II based on the corpus data. However, intuitively *dlja* 'for' is more basic and never tolerates the *j*-form (in our grammar), while the *j*-form with *radi* 'for the sake of' is marginally possible. Also, both these prepositions were originally postpositions (Hill 1977), but only *radi* 'for the sake of' can occur as such in certain cases in Contemporary Russian (Itkin and Tolkacheva 2015), which certainly sets it apart from other Class I prepositions.

This yielded 80 target sentences.

The linear order was very similar across the target items. It can be represented as follows:

(122) 1SG Verb (NP_0) [P NP_1 & NP_2]

In other words, all sentences had the 1st person singular pronoun as subject (this enables a neutral Information Structure as this pronoun is always a topic/given) and the target PP in the sentence-final position. However, the function of PPs was not the same across sentences. *Krome* 'except' and *pomimo* 'besides' always project adjunct PPs; in our case the PPs they headed modified the direct object of the verb, that is, they were put in the version of (122) with an object NP. Similarly, *dlja* 'for' and *radi* 'for the sake of' headed adjunct PPs modifying events signified by transitive verbs.

The remaining 6 prepositions were all used in their locative meaning, which enabled constructing very simple and highly similar sentences, all involving a stative verb *sidet'* 'sit' or *stojat'* 'stand' with the PP specifying the location of the subject DP referents's sitting or standing. These PPs appear to have a more predicative function, differing from the first group in this respect.[45]

Finally, 6 verbs were used, all in the past tense, with a feminine or masculine gender inflection; 48 sentences involved a masculine 3rd person pronoun in the target, 2nd conjunct position, whereas 32 sentences involved a feminine one: the semantically paired prepositions were matched in terms of these two features for each sentence type. Below is a sample of sentences used as targets in the questionnaire.

(123) Ja pozval vsex *krome* tebja i ego — 2SG & [-n]
I called**M** all$_{GEN}$ except you$_{GEN}$ and **he**$_{GEN}$
'I invited everyone, except you and him.'

(124) Ja soveršila èto *radi* Paši i neë — DP & [+n]
I committed**F** this for-the-sake-of Pasha$_{GEN}$ and **she**$_{GEN}$
'I committed this for her and Pasha's sake.'

(125) Ja sidela *za* ej i im — [-n] & [-n]
I sat**F** behind she$_{INS}$ and **he**$_{INS}$

[45] Still, these PPs specifying the location are not s-selected by the verb and thus cannot be considered their true complements.

'I was sitting behind her and him.'

(126) Ja stojal *pered* im i nej [-n] & [+n]
I stood$_{M}$ in-front-of he$_{INS}$ and **she**$_{INS}$
'I was standing in front of him and her.'

4.2.1.3.Control sentences

To control for the potential reluctance of speakers to give extreme judgments for the target items, 40 filler sentences were composed. These involved coordination of verbal, rather than prepositional complements. 20 of these were composed so as to be completely unacceptable (in my own grammar) as they contained *n*-forms which are not licensed in the verbal context, whereas the other 20 were composed so as to be as good as possible. In addition, half of the control sentences contained accusative case-marked arguments (appearing in the same morphological form as the genitive assigned by most prepositions in the sample) whereas the other half involved arguments with instrumental case morphology (like the complements of the second part of the prepositions). The particular types of coordinated arguments used in these sentences are given in Table 11.

Table 11. Types of coordinated V-arguments

2SG & [-n]	2SG & [+n]
[-n] & DP	[+n] & DP
DP & DP	DP & [+n]

The linear order in the fillers was similar to that of the target sentences:

(127) 1SG [V NP_1 & NP_2]

To distract the participants, the control sentences were constructed in such a way as to be more varied in terms of the verb choice (n=14) and tense (past or present). As a result, there were fewer genuine minimal pairs than within the target sentence group. It was assumed that these properties of the verb should not affect the acceptability of sentences. Below is a sample of control items (judgments are mine):

(128) Ja zval tebja i ego
I call.PAST.**M** you.ACC and he.ACC**[-n]**
'I called you and him.'

(129) *Ja ždala tebja i nego
I waited.PAST.**FEM** you.ACC and he.ACC**[+n]**
'I waited for you and him.'

(130) Ja dorožu toboju i im
I value.PRES you.INSTR and he.INSTR**[-n]**
'I care for you and him.'

(131) *Ja gorzhus' toboj i nim
I **proud**.PRES you.INSTR and he.INSTR**[+n]**
'I am proud of you and him.'

4.2.2. Methods, participants and notes on data collection

The coordination questionnaire was prepared and run using the eSurv on-line survey platform. The 80 target and 40 control sentences were distributed over 16 distinct pages. Each page featured 5 target sentences involving different prepositions and 2-3 control sentences. The eSurv platform automatically randomized the order of items on each page.

To avoid direct comparison of various minimal pairs on the part of the informants I developed and adhered to the following principles. First, the set of the prepositions appearing on each page was unique and always involved two Class I and three Class II prepositions or vice versa, but the semantically paired prepositions never appeared on a same page. Second, target sentences on each page involved distinct types of coordinated elements: for instance, at most one of the five sentences involved coordination of two *n*-forms in the P-complement position. Finally, each page always featured two sentences with *n*-forms in the second P-complement and three sentences

with n-less forms in this position or vice versa. As for the fillers/controls, no specific strategy was used to distribute them over pages.[46]

4.2.2.1.Participants

57 people completed the questionnaire, 55 of whom also completed a short questionnaire probing the *n*-form/*j*-form distribution in the single complement case, for the same 10 prepositions.[47] I focus on this latter group to be able to juxtapose the options available for a simple and a coordinated complement for individual speakers. Three speakers specified Tatar, Azerbaijani or a Belarussian dialect as their native language and were excluded from the analysis. I will thus consider data from 52 speakers.

The sociological profile of the group is the following: 25 male and 27 female informants aged 19-70 (mean = 33.5, SD = 15.7; median = 26); 13 linguists and 39 non-linguists.

Almost a half (n=25, 48%) of the group spent their childhood and early adolescence years in the Moscow region of Russia/USSR; 2 speakers grew up in the former Soviet Republics (Belarus or Latvia); 23 participants were raised in other regions of Russia/Russian part of the USSR and 2 speakers didn't specify the relevant region of Russia/USSR.

4.2.2.2. Remarks on the data

To enable a quantitative analysis of the data as well as its more convenient presentation, I converted the verbal judgments into numbers, with the lowest judgment corresponding to '1' and the highest to '4'.

All control sentences described in preceding sections were obtained from all speakers.

As for the target sentences, in the case of 4 prepositions (*naprotiv*, *vperedi*, *krome*, *dlja*) I obtained judgments for 7 and not 8 types of sentences, with one of these

[46] Instructions and the judgment options were the same as in the short questionnaire (cf. A3 of the Appendix).
[47] See A3 of the Appendix.

7 judged twice by the speakers. This unfortunate error happened during the transfer of sentences to the online questionnaire. Table 12 below summarizes for which sentences judgments were not obtained (0) and which were judged twice (2).

Table 12

	Krome	***dlja***	***naprotiv***	***vperedi***
DP & [+n]	1	1	1	1
2SG & [+n]	1	1	2	1
[-n] & [+n]	1	1	1	0
[+n] & [+n]	2	0	1	1
DP & [-n]	1	1	1	1
2SG & [-n]	0	1	1	2
[-n] & [-n]	1	2	0	1
[-n] & [+n]	1	1	1	1

This unintended error, however, gives us a chance to assess how stable the judgments were and thus serve as an additional control. Thus, on an individual, qualitative level:

- 14 speakers did not change their judgments for any preposition
- 20 speakers changed it once or twice but not radically
- 7 speakers changed their judgments from unacceptable ("1") to border-line acceptable ("2") once or twice
- 3 people radically ("3"/ "4"→ "1") changed their judgment for 1 P each and didn't change any other judgment
- 6 speakers radically ("3"/ "4"→ "1") changed their judgment for 1 P each and non-radically changed other judgment(s)
- 2 speakers radically ("3"/ "4"→ "1") changed their judgment for 2 P each and non-radically changed judgments for 1 or 2 Ps. (in two cases – radically)

A paired t-test was performed to compare the mean judgments for the two occurrences of the same sentence, which yielded that for neither preposition this difference was statistically significant ($p > 0.17$, one-tail). For further analysis, I decided to take the maximal judgment of each informant.

4.2.3. Overall results

Overall results are summarized in Table 13, from which we can see that, like in the single complement case (cf. the rightmost column: the first value is for single n-less form complements and the second is for n-forms), in the second conjoined complement of a preposition, *n*-forms were judged more acceptable than *n*-less forms. However, the contrast between n-forms and n-less forms turned out to be less clear-cut than in the single complement case. I suggest that the contrast is blurred for two reasons:

(132) *Lower acceptability of n-forms* in the second conjunct, compared to the single complement case

(133) *Higher acceptability of j-forms* in the second conjunct, compared to the single complement case[48]

Table 13 Means and SD for each sentence type pair for all speakers

Second conjunct	*First conjunct*				*Single complement*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.86 (1.08)	2.07 (1.17)	1.86 (1.08)	1.45 (0.87)	1.29 (0.68)
& [+n]	2.80 (1.08)	2.92 (1.13)	2.82 (1.13)	1.40 (0.83)	3.76 (0.61)
	$**p<<0.001$, one-tail t-test	$**p<<0.001$	$**p<<0.001$	$p=0.32$	$**p<<0.001$

The contrast between n-forms and n-less forms in the second conjunct position was found for each of the three sentence pairs with every preposition (the data on individual prepositions can be found in A2 of the Appendix). It was statistically significant for all pairs except the [-3] & [+n] vs. [-3] & [-n] sentence pair in the case of preposition *pered* 'in front of'.

[48] The two contrasts proved highly significant for all pairs, with $p << 0.001$, one-tail t-test.

As can be seen from Table 13, when two 3rd person pronouns are coordinated, [+n] & [+n] combination turns out to be the most acceptable option, [+n] & [-n] – borderline acceptable and the two other options almost unacceptable overall. At the level of individual prepositions, the hierarchy is the same, but the different options are not always significantly different (see A2 of the Appendix).

These findings, if we assume that acceptability judgments align with actual production, allow us to predict the patterns that we would find in a corpus that is large enough. However, they do not tell us much about speakers' individual grammars. For instance, we cannot really know whether there are individual grammars that allow all sentence types to a certain extent. Also, are there individual grammars in which non-P-adjacent *n*-forms are as good as P-adjacent ones and non-P-adjacent *j*-forms as bad as in the single complement case? To answer questions of this kind let us look at individual response patterns.

4.2.4. Data on individual speakers

Obviously, the 2 factors in (132) and (133) do not have to hold for each speaker. Subsequently, neither does the n-form > n-less form hierarchy need to be true for the second conjunct position – the contrast might be neutralized or even reversed for certain speakers.

Table 14 shows that 9 potential grammars are logically possible, depending on whether the acceptability of n-forms and n-less forms in the second conjunct plunges, drops or stays the same, compared to the single complement case.

Table 14

	X&[-n] = [-n]	**X&[-n]>[-n]**	**X&[-n]<[-n]**
X&[+n]=[+n]	1a. no +/-adjacency effects *8 speakers* *(no linguists)*	2a. -adjacency for [-n] *6 speakers (no linguists)*	
X&[+n]<[+n]	1b. +adjacency for [+n] *19 speakers* *(4 linguists)*	2b. +adjacency for [+n] –adjacency for [-n] *19 speakers (9 linguists)*	
X&[+n]>[+n]			

In fact, however, only 4 patterns - the ones consistent with the overall picture - are attested in our sample. Let us discuss them in detail.

(1a) Neither the acceptability of j-forms, nor that of n-forms significantly changes in the second conjunct position compared to the single P-complement case. Thus, descriptively speaking, the coordination condition (perhaps boiling down to adjacency/non-adjacency of the pronominal complement to the head P) does not significantly affect the choice of a pronominal form. 8 non-linguist speakers admitting both *n*- and *j*-forms with up to 3/10 prepositions in the single complement case displayed this pattern. Two grammars presented as tables below instantiate different manifestations of pattern (1a).

Speaker AF

Second conjunct	*First conjunct*				*Single*[49]
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.0 (0)	1.0 (0)	1 (0)	1.0 (0)	1.0 (0)
& [+n]	3.6 (0.97)	3.4 (1.27)	3.0 (1.5)	1.0 (0)	3.4 (0.97)
	*p<<0.001, one-tail	*p<<0.001, one-tail	*p<<0.001, one-tail	p=0.5, one tail	*p<<0.001

[49] This column presents mean acceptability of n-less and n-forms respectively in the single complement case (Auxiliary questionnaire data, found in A2 of the Appendix).

Speaker TF

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.9 (1.45)	1.22 (0.44)	1 (0)	1.11 (0.33)	1.2 (0.42)
& [+n]	3.8 (0.63)	4 (0)	4 (0)	1(0)	3.8 (0.63)
	*p=0.001, one-tail	*p<<0.001, one-tail	*p<<0.001, one-tail	p=0.17, one tail	*p<<0.001

(1b) *N*-forms are significantly degraded in second conjunct whereas the acceptability of n-less forms remains unchanged. Descriptively speaking, n-forms' acceptability is sensitive to adjacency to the P-head. This pattern is found for 19 speakers (4 linguists) admitting both n- and n-less forms with up to 7/10 prepositions in the single complement case). Three grammars (of speakers with stable judgments) presented below instantiate different manifestations of pattern (1b). Depending on the extent of n-form degradation and n-less form acceptability in the simple case, this may result in the persistence of the n-form > n-less form contrast (16 speakers) or in neutralization of this contrast (3 speakers).

For 6 speakers that only accepted n-forms in the single complement case the contrast remained categorical, as illustrated by data for Speaker EM.

Speaker EM

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1 (0)	1 (0)	1 (0)	1.0 (0)	1.0 (0)
& [+n]	2.7 (0.68)	2.9 (0.74)	3.11 (1.05)	1.0 (0)	3.6 (0.7)
	*p<<0.001, one-tail	*p<<0.001, one-tail	*p=0.0002, one-tail	p=0.5, one tail	*p<<0.001

Speaker IS pattern is interesting. Although there is a significant contrast between n-forms and n-less forms in the second conjunct following an n-form and overall, we can see that the degradation of n-forms is so prominent that there seems to be no

categorical contrast between the two forms in coordination. Their grammar thus obeys quite a strong constraint on non-P-adjacent n-forms in the P-complement domain.

Speaker IS

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1 (0)	1.44 (1.01)	1.1 (0.32)	1.56 (1.13)	1.0 (0)
& [+n]	1.2 (0.42)	1.8 (1.32)	1.67 (0.87)	1.56 (0.73)	4.0 (0)
	p=0.17, two-tail	p=0.52, two-tail	*p=0.04, one-tail	p=1, two-tail	*p<<0.001

Speaker MCh represents another pattern, which I deem reliable as their judgments were stable and they displayed a clear-cut contrast in control sentences. They accepted both n-forms and n-less forms with 6/10 prepositions in the simple case. Note that this lack of a categorical contrast between n-forms and n-less forms in the single complement case is even more prominent in coordination: the contrast is only border-line significant for just one sentence type and overall. *MCh* can be said to admit both forms in coordination.

Speaker MCh

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.8 (0.63)	2.44 (0.73)	2.7 (0.48)	1.89 (0.93)	2 (1.15)
& [+n]	2.3 (0.67)	2.9 (1.1)	2.78 (0.44)	1.78 (0.83)	3.8 (0.42)
	*p=0.05, one tail	p=0.15, one tail	p=0.36, one tail	p=0.4, one tail	*p=0.0004

(2a) N-less forms are significantly more acceptable non-adjacently, whereas the acceptability of n-forms remains unchanged. Descriptively speaking, non-adjacency to P relaxes the constraint on n-less forms in P-complement. As in (1b), n-form > n-less form contrast may either remain significant (4/6 speakers, exemplified by *Speaker MM*) or almost dissolve (exemplified by *Speaker V*).

Speaker MM

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.9 (0.57)	2 (0)	1.9 (0.32)	1.33 (0.5)	1.2 (0.42)
& [+n]	3.3 (0.67)	3.8 (0.42)	3.22 (1.3)	1.78 (0.97)	3.9 (0.32)
	*p<<0.001, one tail	*p<<0.001, one tail	*p=0.008, one tail	p=0.12	*p<<0.001

Speaker V

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	3.2 (0.79)	3.22 (0.97)	3.4 (0.97)	2.78 (1.3)	1.6 (0.7)
& [+n]	3 (1.05)	2.8 (0.92)	3.22 (1.09)	2.22 (1.2)	3.4 (0.84)
	p=0.32, two tail	p=0.17, one tail	p=0.35, one tail	p=0.18, one tail	*p<<0.001

(2b) N-forms are significantly degraded *and* n-less forms significantly more acceptable in the second conjunct. Descriptively, n-forms' licensing as well as the constraint on n-less forms in P-complement are sensitive to adjacency to their P-heads. The radical manifestation of this pattern is the reversal of the initial n-form > n-less form pattern, found for 5 speakers and most prominent for *Speaker SB* (linguist), presented here.

Speaker SB$_{ling}$

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	3.6 (0.97)	3.56 (0.53)	2.9 (1.29)	1.78 (0.44)	1.7 (0.68)
& [+n]	1.1 (0.32)	1 (0)	2 (0.71)	1 (0)	3.9 (0.32)
	*p<<0.001, one tail	*p<<0.001, one tail	*p=0.04, one tail	*p=0.0004, one tail	*p<<0.001

For 5 other speakers, the contrast is merely neutralized (*Speaker PP*) and for the remaining 9 speakers in the group it is preserved (*Speaker EL*).

Speaker PP$_{ling}$

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	3.4 (0.97)	3.44 (0.88)	3.5 (0.53)	1.67 (0.71)	2.2 (1.03)
& [+n]	3.7 (0.48)	3.9 (0.32)	3.89 (0.33)	1.89 (0.78)	4.0 (0)
	p=0.2, one tail	p=0.09, one tail	*p=0.04, one tail	p=0.27, one tail	*p=0.0002

Speaker EL

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.4 (0.7)	2.0 (1.12)	1.3 (0.95)	1 (0)	1 (0)
& [+n]	2.9 (0.74)	3.5 (0.71)	3.22 (0.83)	1 (0)	4 (0)
	*p=0.0002 (two tail)	*p=0.003 (two tail)	*p=0.0002, two-tail	P=1, two-tail	*p<<0.001

4.2.5. Summary

The *n*-form > *j*-form contrast was observed for each individual speaker in the single P-complement case.[50] Judgments for *n*-forms ranged from 2.9 to 4 and for *j*-forms – from 1 to 2.2, with the smallest difference recorded for *Speaker KD* (2.9 vs. 2), which that nevertheless proved statistically significant.

In the second conjoined complement, however, we find all possible patterns:

- n-forms > n-less forms (37/52 speakers);
- n-forms ≈ n-less forms (9/52 speakers);
- n-less forms > n-forms (6/52 speakers).

All 52 speakers displayed a clear-cut contrast between 'good' and 'bad' controls (t-test, $p<0.001$, 1-tail), the individual mean judgment range being [2.55; 4] and [1; 2.1] respectively, which allows us to rely on their judgments of the target sentences.

In the next section, we will discuss how these patterns can be accounted for and what they reveal about the nature of *n*-forms.

[50] n-forms (mean = 2.9-4) were significantly more acceptable than n-less forms (mean = 1-2.2) for each individual speaker: this difference was highly significant (t-test, $p<0.001$, 1-tail) for 47 speakers and significant (t-test, $0.001<p<0.05$, 1-tail) for 5 speakers (*OM, TYa, DCh, IM, KD* – the first four characterized by rather unstable judgments).

4.3. Accommodating new data in the analysis: The evolution of [n] from a person agreement marker associated with the P into a P-case marker on the pronoun

I propose that there are two clear types of grammar that can be distinguished among the speakers of Standard Contemporary Russian:

i. *N*-forms are licensed in the P-complement, irrespectively of adjacency to P;
ii. *N*-forms are licensed in the P-complement only under adjacency with P.

Other patterns either tend to Grammar (i) or Grammar (ii) or combine the properties of both (depending on a preposition, perhaps). The important factor is whether n-forms and j-forms are sensitive to adjacency when in a coordinated P-complement. I propose that in those individual grammars that require or at least allow n-forms both adjacently and non-adjacently, the [n] is morpho-phonologically a part of the pronominal case form.

If a grammar only admits *n*-forms adjacently, the [n] is phonologically a part of the pronoun (for historical reasons, cf. Chapter 2.1),[51] but belongs to the preposition in the morphosyntactic sense, namely, is a person agreement marker on the P. Thus, there are speakers who treat [n] as part of the pronoun and speakers who treat it as a person agreement marker on the preposition. I suggest that this reflects the continuing reanalysis of [n] in the language, schematically represented in (134).

(134) **Prothetic n** [Common Slavic]

⇓

3rd person agreement marker on P [Old Russian to Modern Russian][52]

⇓

P-case marker [Contemporary Russian]

Speakers belonging to groups (1a-b), for whom non-adjacency does not increase the acceptability of *j*-forms, are closer to Stage III, while those for whom it does, (2a-

[51] Note that the P + *n*-form string typically projects a single prosodic word (recall that most P-heads are proclitics, cf. Gribanova 2009), so [n] may associate either with the P or with the pronoun: the spelling convention may not faithfully reflect the actual locus of [n].
[52] Cf. Section 2.3.3 where Hill's diachronic observations on the occurrence of [n] on the possessor instances of GEN/ACC 3rd person pronouns are reported.

b), are closer to Stage II. Since there is considerable variation is among speakers, we can hypothesize that Contemporary Russian still has not fully reached the third stage. This also accounts for why some individual grammars are less consistent than others.

5.*N*-forms and *j*-forms in the phrasal comparative

5.1. Corpus data

The DP introducing the standard of comparison in the Russian phrasal comparative is the only environment besides the P-complement position where *n*-forms are found. This DP always bears genitive case and may also take the *j*-form: these two features are reminiscent of Class II prepositions discussed above. Although prescriptive grammarians (e.g. Bylinskij and Rozental' 2011) give only *j*-forms as acceptable, *n*-forms are widely attested in this position and have been available for a while: the earliest instance in the Russian National Corpus goes back to the 18th century, as noted in Sichinava (2013).

N-forms and *j*-forms are found with synthetic comparative adjectives, adverbs and quantifiers, as shown in (135-137) respectively.

(135) Maša vyše {ego; nego; Miši}
Masha taller he.GEN[-n] he.GEN[+n] Misha.GEN
'Masha is taller than {him; Misha}.'

(136) Maša prygaet vyše {ego; nego; Miši}
Masha jumps higher he.GEN[-n] he.GEN[+n] Misha.GEN
'Masha jumps higher than {him; Misha}.'

(137) Maša zarabatyvaet bol'še {ego; nego; Miši}
Masha earns more he.GEN[-n] he.GEN[+n] Misha.GEN
'Masha earns more (money) than {him; Misha}.

As has been pointed out and discussed in several papers (e.g. Matushansky 2002, Pancheva 2006, Grashchenkov and Lyutikova 2017), the genitive standard is not licensed with periphrastic/analytic comparative forms. Compare (135) and (136) with (138) and (139) respectively.[53]

(138) *Maša bolee vysokaja {ego; nego; Miši}
Masha more tall he.GEN[-n]/[+n] Misha.GEN
'Masha is more tall than {him; Misha}.'

[53] Sentences like (137) with a comparative quantifier do not have a periphrastic counterpart.

(139) *Maša prygaetbolee vysoko {ego; nego; Miši}
Mashajumps more highly he.GEN[-n]/[+n] Misha.GEN
‘*Masha jumps more highly than {him; Misha}.’

N-forms have also been observed (Sichinava 2013 and references therein) to occur more frequently with synthetic comparatives bearing the monosyllabic *–e/še* comparative suffix, compared to the disyllabic *–ee* and also with adverbs more than with adjectives. Let us take a closer look at the distribution of *n*-forms and *j*-forms with adverbial and adjectival comparatives.

Our study of a sample of synthetic comparatives in the three (Main, Newspaper and Speech) subcorpora of the Russian National Corpus, restricted to instances dating from 1980 to 2015, indeed suggests that comparatives in the adverbial sense account for a significantly larger proportion of *n*-forms. In other words, whenever one comes across an instance of an *n*-form in the genitive standard of comparison, the synthetic comparative is much more likely to be one formed on an adverb, (136), rather than one formed on an adjective, (135). In contrast, when one encounters a *j*-form, the synthetic comparative next to it is more likely to be an adjective. Overall, *j*-forms are found more often in this construction, at least, as far as the frequent synthetic comparatives are concerned. All this can be seen from Table 15.

Table 15 Distribution of n-forms and j-forms in the phrasal comparative

COMPARATIVE FORM	Main, Newspaper and Speech Corpora (1980-2015)						
	n-form			**j-form**			**Total**
	Adj	Adv	Total	Adj	Adv	Total	
ran'še 'earlier'	N/A	105	105	N/A	151	151	256
pozže 'later'	N/A	7	7	N/A	8	8	15
prežde 'before'	N/A	13	13	N/A	0	0	13
čaščе 'more often'	N/A	4	4	N/A	1	1	5
bol'še 'bigger/more'	17	55	72	79	96	175	247
lučše 'better'	25	105	130	252	238	490	620
xuže 'worse'	13	15	28	66	54	120	148
bystree/j 'faster'	3	10	13	5	11	16	29
dal'še 'farther/further'			13			10	23
bliže 'closer'	2	3	5	4	2	6	11
(po)sil'nee/j 'stronger'	9	2	11	103	4	107	118
(po)slabee/j 'weaker'	1	0	1	20	0	20	21

krasivee/j 'prettier'	1	0	1	15	0	15	16
(po)umnee/j 'smarter'	1	0	1	42	0	42	43
(po)legče 'easier, lighter'	0	0	0	5	0	5	5
(po)tjaželee/j 'heavier, harder'	2	2	4	12	0	12	16
(po)molože 'younger'	17	N/A	17	245	N/A	245	262
starše 'older'	43	N/A	43	478	N/A	478	521
mladše 'younger'	13	N/A	13	97	N/A	97	110
Total	**91**	**323**	**481**	**1423**	**565**	**1998**	**2479**

Note also that those comparatives that can be both adjectives and adverbs, like *bol'še* 'bigger/more', *lučše* 'better', *xuže* 'worse', occur significantly more frequently with *j*-forms. This, however, does not seem to be due to the predominance of adjectival instances (see Table 15): most likely, it reflects the general predominance of *j*-forms in this environment. This generalization is schematically presented in (140).

(140) Adv ≈ Adj → *j*-form > *n*-form

For those comparatives that can only be adverbs, the frequency contrast is less pronounced (see *ran'še* 'earlier', *pozže* 'later') and may in fact be reversed (in case there are few instances of the adverb + 3rd person pronoun strings in the corpus), see *prežde* 'before' and *čaščе* 'more often'.

Similarly, forms that are predominantly adjectival (with shaded cells in the lower part of Table 15) will be almost exclusively found with *j*-forms, as schematized in (141).

(141) Adj >> Adv → *j*-form >> *n*-form

The generalizations made in this section may prove useful for research studying the patterns of usage of the two pronominal forms in adjectival, adverbial and amount comparatives. It is clear, however, that qualitatively speaking, both pronominal forms are in principle licensed with synthetic comparatives of any form and of either function. Our aim here is thus to understand the morphosyntactic mechanisms underlying the licensing of the two forms.

5.2. The comparative puzzle: problems of previous analyses

In Chapter 2.3 we considered three treatments of n-forms – Yadroff and Franks (1999), Hill (1977) and Daniel (2015). We shall now see that neither of them can provide a principled explanation for why *n*-forms are licensed with synthetic comparatives.

For Yadroff and Franks' approach, the licensing of *n*-forms in the phrasal comparative is as problematic as their licensing in the complement of most 'lexical' (secondary, morphologically complex) prepositions. This is because in their analysis only functional prepositions that are a product of morphology license n-forms, by virtue of realizing the F0 node associated with the noun phrase, which "forces" the pronoun to be realized under F0. Since there is no functional preposition in phrasal comparatives, there is no way to derive sentences with *n*-forms in the standard of comparison.

In turn, the major weakness of Daniel's (2015) proposal that *n*-forms are adpositional (non-standard) variants of 3rd person pronominal forms triggered by certain lexical items is that a vast number of prepositions and practically all comparative adjectives and adverbs (or, probably, the synthetic comparative suffix *-e/še/že/če/ee* '-er') should be specified as *weak triggers*, since they allow but do not require n-forms. Such an approach captures the data but does not explain what determines how strong of a trigger a particular lexeme is. Furthermore, it cannot explain why synthetic comparatives are the only type of lexeme beyond prepositions that trigger *n*-forms.

Comparatives are a problem for Hill (1977) as well, since their taking of *n*-forms defines them as "actual prepositions". Why isn't it a welcome result?

Although certain comparative adverbs have semantically similar prepositional counterparts (compare *ran'še* 'earlier', *prežde* 'earlier, before' (Comp. Adv) with *pered* 'in front of, before' (P); *pozže* 'later' (Comp. Adv) and *posle* 'after' (P)) and might be considered to have undergone prepositionalization, they still select comparative clauses introduced by *čem* – a property that no preposition has, cf. (142) and (143).

(142) Ivan prišël {✓posle; ✓pozže} menja
Ivan came after later me$_{GEN}$
'Ivan came after/later than me.'

(143) Ivan prišël {*posle; ✓pozže} čem ja
Ivan came after later WH I$_{NOM}$
'Ivan came {*after; later} than I (did).'

In addition, even for those comparative forms whose morphological/phonological composition is idiosyncratic and not fully predictable from the corresponding positive form of an adjective/adverb speakers can quite easily establish a connection to the adjectival/adverbial source (*ran'še* 'earlier' – *rano/rannij* 'early.ADJ/ADV'*; pozže* 'later' – *pozdno/pozdnij* 'late.ADJ/ADV'*),* whereas there is obviously no such connection for the semantically similar prepositions. If one wants to maintain the connection between positive and comparative forms as well as to avoid postulating that words like *ran'še* 'earlier' also exist as independent, prepositional lexemes in the lexicon, one faces the problem of how to explain n-form taking on the part of comparatives.

Things are further complicated by the fact that comparative adjectives that are the predicate of a clause are also capable of licensing *n*-forms (e.g. *Miša molože* ***n****ego* 'Misha is younger than him'). Since these instances of comparatives are hardly classifiable as prepositions, Hill's definition of an actual preposition becomes problematic.

To be sure, Hill is in fact not eager to include the comparatives in the class of prepositions and does entertain the possibility that in this environment the use of *n*-forms might proliferate to disambiguate adjectival [predicative] and adverbial instances of the synthetic comparative (Hill 1977:213). Whereas this sounds reasonable and is in accordance with the corpus frequency facts (see Section 5.1), it does not enable a uniform formal treatment of *n*-forms' occurrence in the P-object and the comparative-object positions. Under this approach the two environments can only be united by means of a functionalist account, positing that *n*-forms' occurrence is motivated by the need to resolve structural ambiguity. Thus, *n*-forms in the P-complement position would signal that they are actual (immediate) complements of a preposition, rather than modifiers of an actual complement, while *n*-forms in the comparative would signal that the comparative is an adverb, i.e. a modifier semantically and an adjunct syntactically, rather than an adjective, fulfilling the

predicative function. Note, however, that even this account would not be quite uniform: in prepositional environments *n*-forms would be used to disambiguate the structural position of the 3rd person pronoun with respect to a preposition, while in the phrasal comparative they would disambiguate the function and structural position of the comparative form they depend on. Furthermore, this type of account runs into the following problem.

As presented in the previous section, in Contemporary Russian n-forms are licensed with both comparative adjectives and comparative adverbs, although they do tend to occur more frequently with adverbs. Although one could say that these tendencies reflect language change and eventually *n*-forms will only become acceptable with comparative adverbs, the change actually seems to work in the opposite direction: having first expanded to the comparative adverb environments, *n*-forms are now used more and more frequently with comparative adjectives as well, especially by younger speakers. These data do not allow us to maintain Hill's alternative proposal: we either have to posit that all comparatives taking *n*-forms are actually prepositions – a claim that would disagree with their morphological, semantic and syntactic properties – or to say that in these environments *n*-forms actually disambiguate the syntactic status of the third person pronoun itself, just like in the prepositional complement domain.

The first option would not be appealing to Hill and probably any other linguist, so we are left with the second one. Unfortunately, it does not work either, since it does not explain why the possessive/personal pronoun ambiguity should only be resolved in PPs and phrasal comparatives but not in complements of other heads, for example, verbs. Compare the sentences in (144) and (145). They involve morphologically and semantically related verbs *vljubit'sja* 'to fall in love' and *ljubit'* 'to love', which require a PP and a DP-complement respectively. We can see that the potentially ambiguous *ego* (*j*-form) is impossible in the PP-complement of *vljubit'sja* but is *required* in the DP-complement of *ljubit'.* At the same time, the potentially ambiguous *ego* can be used on a par with the non-unambiguous *nego* in the phrasal comparative, as shown in (146).

(144) Maša vljubilas' v {✓ego brata; *ego; ✓**n**ego}
Maša fell-in-love in his$_{ACC=GEN}$ brother$_{ACC=GEN}$ he$_{ACC=GEN}$ he$_{ACC=GEN}$
'Masha fell in love with his brother/him.'

(145) Maša ljubit {✓ego brata; ✓ego; *nego}
Maša loves his$_{ACC=GEN}$ brother$_{ACC=GEN}$ he$_{ACC=GEN}$ he$_{ACC=GEN}$
'Masha loves his brother/him.'

(146) Maša umnee {✓ego brata; ✓ego; ✓**n**ego}
Maša smarter his $_{ACC=GEN}$ brother$_{ACC=GEN}$ he$_{ACC=GEN}$ he$_{ACC=GEN}$
'Masha is smarter than his brother/him.'

If ambiguity resolution were the guiding principle behind the distribution of *n*-forms in Russian, we would expect that in all environments where 3.ACC/GEN did not stand for a possessor but was a DP in its own right, it would take an *n*-form. Instead, we find *n*-forms only in the P-complement and the standard of comparison positions.

5.3. Licensing of *n*-forms and *j*-forms in the phrasal comparative

Why are *n*-forms licensed in the complement of the phrasal comparative? Should those comparative adjectives and adverbs be analyzed as P-heads after all?

I argue that the phrasal comparative may or may not involve a null P-head, immediately dominating the genitive standard of comparison DP. In other words, there exist two separate constructions across and possibly within individual grammars). The structure involving a P-head requires *n*-forms; the other structure only admits *j*-forms.

We have seen in Section 5.1 that the genitive standard of comparison (SOC), in particular, *n*-forms or *j*-forms in the genitive are only licensed by the *synthetic* comparative forms. In fact, this is parallel to some other languages as well:

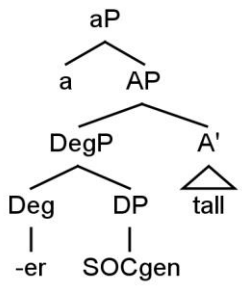
(147) Introduction of a standard of comparison DP with fixed case-marking is possible iff the adjective/adverb bears a synthetic comparative form (holds at least for Russian, Greek and Hungarian)[54]

A natural proposal then is to link the genitive case-assignment to the synthetic comparative form. This can be done in different ways. In particular, Merchant (2012) proposes for the Greek genitive of comparison, shown in (148), that the comparative –*ter*- '-er' suffix of the adjective is a Deg-head taking the standard DP as its complement and assigning genitive case to it.

(148) O Giannis ine psilo**ter**os su
The Giannis is tall.er.MSG you.**GEN**
'Giannis is taller than you.'

The full aP structure that Merchant argues for is as follows:

Figure 7. Part of Merchant's (2012) structure for the Greek genitive of comparison



Bailyn (2004, 2012) attempts to provide a uniform treatment for a number of instances of genitive case assignment in Russian and his structure suggests a different kind of link. In his analysis, the comparative adjective/adverb bears a [+q] (quantificational) feature and thus selects a QP, rather than an NP/DP complement.

[54] Cf. Merchant (2012), Wunderlich (2001) for Greek and Hungarian respectively.

This null Q-head is what assigns genitive case to the standard of comparison NP,[55] just like overt quantifiers like *mnogo* 'many', *neskol'ko* 'several' would do.

(149) $[_{AP}\ A_{[+q]}\ [_{QP}\ Q\ NP]]$ (adapted from Bailyn 2004: ex (23))

In effect, both Merchant and Bailyn connect GEN-assignment to a quantificational head. Adopting this idea allows us to account for the licensing of *j*-forms in the standard of comparison: there is no P-head in such a structure, so the *n*-form must be blocked. Also, this analysis predicts that other quantifiers assigning GEN case will license only *j*-forms. This is indeed borne out:

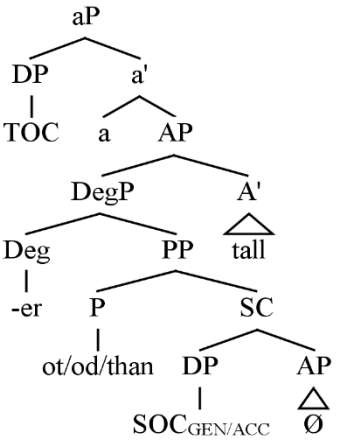
(150) Mnogo ix [zaicev] begalo vokrug
Many they.GEN hares GEN ran.N around
'*Many them/they [hares] ran around.'[56]

A different approach to the case licensing on the standard of comparison is taken in Pancheva (2006, 2010). She defends a small clause structure (PredP/FP) behind the standard of comparison DP. The small clause is a complement to preposition *ot/od* 'from' in several Slavic languages and to a null P in Russian, cf. Figure 8 below. Since the genitive/accusative case-marking on the standard DP corresponds to the case-government properties of this preposition in the respective languages, the P is assumed to exceptionally case mark the standard DP.

[55] A problem with this analysis is that the standard of comparison can be pronominal and thus definite. If we adopt the view that DPs are projected in Russian, we would run into a problem. I thus go for Merchant's analysis.

[56] There is no true grammatical equivalent of such sentences in English, since the complement of *many* does not allow pronominalization of the Russian kind ('many of them' has a different, partitive interpretation).

Figure 8. Part of the syntactic tree for phrasal comparatives under Pancheva's (2006) analysis



Pancheva (2006) considers the phrasal comparative to be similar to the pseudo-partitive construction. In particular, she proposes that the overt or null P in this structure takes as its argument a set of degrees and returns a part of this set of degrees and thus has the type <dt, dt>.

Apart from Pancheva's semantic considerations, the null P proposal for Russian is supported by the fact that certain synthetic comparatives admit an overt *ot* 'from, of' in the relevant position:

(151) Cena byla na 10% niže (ot) isxodnoj
Price was on 10% lower of original$_{GEN}$
'The (current) price was 10% lower than the original (price).'[57]

The null P analysis allows us to capture why *n*-forms are licensed in the phrasal comparative construction: this is so because they receive case from a P-head.

I thus account for the availability of both *n*-forms and *j*-forms in the phrasal comparative by proposing that both case-assignment options entertained in the literature are available. In particular, whenever the PP is projected in the structure the

[57] A brief corpus study suggests that the overt P mostly occurs with spatial comparatives, modified by a measure phrase, see (i). This is an interesting fact that warrants an independent study.

(i) Puli ušli vyše ot celi na 5 santimetrov
Bullets went-off higher from aim on 5 centimeters
'Bullets missed the aim by 5 centimeters.'
[P. Evdokimov. 2003. *Specnaz dlja Rossii*]

standard DP will receive case from the P-head and thus take the *n*-form; otherwise it will get its genitive case from the Deg-head projected by the quantificational *–er* morpheme (under Merchant's analysis) or from a Q-head (under Bailyn's approach) and take the *j*-form. Since both Deg and Q are quantificational heads and quantifiers in Russian take complements in the genitive case, we expect Deg to only license *j*-forms, as is typical for Q.

Given that the availability of *n*-forms in the standard of comparison is a relatively recent development and given that the morphologically less regular *–e/še* comparatives occur with them more often than the regular *–ee* comparatives, it is reasonable to assume that a structure with a PP-layer has developed as a reaction to Deg's losing its case-licensing ability (probably due to a fusion of A and Deg). Thus, we may posit that Deg oscillates between taking a DP (if it is capable of licensing case) and a PP-complement (if it is no longer capable of licensing case).

As a final remark, this historical development makes sense cross-linguistically. For example, the Greek genitive of comparison, shown in (148) originated in Classical Greek and is very restricted in Modern Greek (Merchant 2012), having been superseded by the 'younger' prepositional *apo*-comparative (cf. Merchant 2012). Similarly, Latin had ablative of comparison, where Italian has prepositional *di*-comparative, compare (152) and (153). This seems connected to the gradual loss of case distinctions.

(152) Marta è più alta ***di*** Giacomo *Italian*
Marta is more tall from Giacomo
'Marta is taller than Giacomo.'

Napoli & Nespor (1986: 625)

(153) Cato Cicero-ne eloquentior est
Cato.NOM Cicero-ABL more-eloquent is
'Cato is more eloquent than Cicero.'

Stassen (1985: 27)

6. The phrasal comparative: Structure behind the standard DP

This chapter continues the study of the Russian phrasal comparative, focusing now on the internal syntax behind the genitive standard DP. The construction has been studied from various angles in such works as Matushansky (2002), Pancheva (2006), Ionin & Matushansky (2012), Berezovskaya (2013), Berezovskaya & Hohaus (2015), Graschchenkov and Lyutikova (2017). Here I look at it from yet another angle and offer some novel data.

The chapter is structured as follows. In Section 6.2 the nominal genitive of comparison is compared against another phrasal comparative, the *čem*-comparative, based on several tests that probe the structure underlying the phrasal standard. Section 6.3 reviews existing analyses that the genitive of comparison is compatible with and discusses whether evidence motivating these analyses can be found in Russian as well. Section 6.4 presents a constraint on the correlates to the nominal genitive standard and develops a version of the Reduced Clause Analysis that derives the observed facts. Section 6.5 discusses data from other languages with phrasal comparatives that only allow single DP-standards and shows that some of them appear to obey correlate constraints similar to the Oblique Correlate Constraint found in Russian, whereas others do not seem to be constrained in this respect. Section 6.6 introduces additional data from Russian that suggest that the proposed constraint may be both too weak and too strong and shows how the analysis can deal with them. Section 6.7 concludes and outlines directions for further research.

6.1. Overview of the literature: the debate

This chapter continues the study of the Russian genitive of comparison and as such contributes to the growing body of research on the syntax of phrasal comparatives. *Phrasal comparatives* refer to constructions in which the degree taken as the standard of comparison is expressed as a single phrase, typically, an NP/DP or a measure phrase, (154-155). They contrast with *clausal comparatives*, wherein the standard is instantiated by a full or reduced clause, (156-157).

(154) John is taller than Bill.

(155) John is taller than six feet.

(156) John runs faster than Bill swims.

(157) John is taller than Bill is.

I refer to DPs like *Bill* as the *standard of comparison (SOC)* and to the DPs they are contrasted with, *John* in (154-157), as the *target of comparison (TOC)* or *correlate*. *Than* is the standard marker and *-er/more* is the *comparative morpheme* and *quantifier* respectively.

The fact that in English the same standard marker *than* is used in both phrasal and clausal comparatives, sparked a debate on whether phrasal comparatives like (154) are derived from clausal comparatives like (157) via ellipsis or whether *Bill* in (154) should be analyzed "directly" as a DP/NP, complement to *than*, which in that case is analyzed as a preposition. Both approaches have been advocated in the literature since the 1970s (cf. an excellent review in Lechner to appear). Arguments for the CP analysis can be found in e.g. Bresnan (1973); Lechner (2001, 2004, to appear); Bhatt and Takahashi (2011) and arguments for the direct analysis – in Hankamer (1973); Hoeksema (1983); Pinkal (1990); Kennedy (1999); Pancheva (2006; 2010).

There is still no complete consensus as far as English is concerned (Lechner to appear vs. Pancheva 2010). However, the debate became much less principled as linguists began investigating phrasal comparatives in other languages. It has been argued that both structures are needed to capture cross-linguistic facts: direct analyses have been proposed for single DP-standards in Hindi and Japanese (Bhatt & Takahashi 2011), Malayalam (Menon 2012) and Malagasy (Potsdam 2012) and clausal analyses – for Slavic (Pancheva 2006), Modern Greek (Merchant 2009) and Dari (O'Connor 2013).

Moreover, both types of structure may co-exist in a single language, typically, one that has different standard-marking strategies. For instance, unlike English, a number of languages of the European sprachbund have two types of comparative construction: a locational comparative and a particle comparative, observation and terms due to Stassen (1985, 2013). A locational comparative is essentially a construction in which the standard of comparison DP is marked by an adposition or case morpheme, typically with a locative meaning. Thus, the *di*-comparative of Italian, the adessive comparative of Hungarian and the genitive of comparison as well as the *apo*-

comparative of Greek, illustrated in (158), (160), (162-163) respectively, belong to this type.

(158) Marta è più alta ***di*** Giacomo. *Italian*
Marta is more tall from Giacomo
'Marta is taller than Giacomo.'

(159) È più alta Marta ***che*** Giacomo.
Is more tall Marta WH Giacomo
'Marta is taller than Giacomo.'

(Napoli & Nespor 1986: 640)

(160) Anna érdekes-ebb volt Péter-***nél***. *Hungarian*
Anna interesting-er was Peter-**ADESS**
'Anna was more interesting than Peter.'

(161) Anna érdekes-ebb volt, ***mint*** Péter.
Anna interesting-er was **CREL** Peter
'Anna was more interesting than Peter.'

(Wunderlich 2001: ex. 3)

(162) O Giannis ine psilo.ter.os ***mu***. *Greek*
The Giannis is tall.er.MSG 1SG.**GEN**

(163) O Giannis ine {pjo psilos; psiloteros} ***apo*** mena.
The Giannis is more tall.MSG tall.er.MSG from 1SG.**ACC**

(164) O Giannis ine {pjo psilos; psiloteros} ***ap'oti*** eimai.
The Giannis is more tall.MSG tall.er.MSG from-WH 1SG.NOM
'Giannis is taller than me.'

(based on data from Merchant 2009; 2012)

A distinctive feature of locational comparatives is that the morphological (case) form of the SOC DP is fixed. In particle comparatives, on the other hand, the form of the standard is typically parallel to that of the correlate, as can be seen from (159), (161) and (164). Particle comparatives with a single DP standard tend to reveive a *Reduced Clause Analysis* (Napoli and Nespor 1986 for the Italian *che*-comparatives; Merchant 2009 for the Greek *ap'oti*-comparatives), whereas locational comparatives

prima facie suggest a simple, *Direct Analysis*, but may be treated as reduced clauses if there is compelling evidence (cf. two analytical options that Merchant 2009 proposes for the Greek *apo*-comparative).

In Russian, five different phrasal comparative constructions can be distinguished. In one broad type of phrasal comparative, corresponding to the particle comparative of Stassen's, the standard of comparison is introduced by *čem* 'lit. what.INS' (165) or *neželi* (166), both traditionally considered (subordinating) conjunctions and so labelled in the Russian National Corpus (RNC). It should be noted though that *čem* and *neželi* are too specialized for conjunctions since they only function as standard of comparison markers. Moreover, *čem* derives from the inanimate interrogative pronoun *čto* 'what', itself also introducing finite subordinate clauses and relative clauses in Russian, suggesting it is a *wh*-operator or a *wh*-complementizer occupying Spec, CP (cf. Mezhevich 2006). Pancheva (2006) assumes this for *čem* as well, placing it in the Spec, CP of the comparative clause. Acknowledging its origin and morphology, I gloss it as WH. Determining the category of *neželi* is even more difficult: it comprises the proclitic particle *ne* (originally, negation), the emphatic enclitic particle *že* and the interrogative enclitic particle *li*, hence the gloss I adopt. I leave the question of which syntactic position it occupies for further research. Although it would be natural to assume that *neželi* is a C-head, given that the two enclitic particles are second-position (Wackernagel) clitics and that the question particle *li* might be considered a clause-typing element (Shlomina 2014 puts it in C as well), this assumption does not fit well with the following fact. In colloquial language *neželi* and *čem* may co-occur in which case *neželi* precedes *čem*, (167).[58, 59] Such sentences are licensed in the same environments as *neželi*-comparatives. In all these comparatives the standard DP bears the same morphological case as its correlate.

[58] Sentences like (167) that employ the two markers simultaneously are found much less frequently than either (165) or (166) involving only one comparative 'particle'. Furthermore, in Contemporary Russian *čem*-comparatives are more frequent than *neželi*-comparatives – the latter have a slightly archaic flavor.

[59] In fact, the reverse order *čem neželi* is also found in the corpus, but there are only five such documents in the RNC, all in the Speech Subcorpus. By contrast, instances of *neželi čem* are found in all three subcorpora (i.e. in both spoken and written registers, in fiction and newspapers), amounting to 58 documents. This frequency contrast aligns with intuitions of native speakers, who marginally accept *neželi čem* but reject *čem neželi.* I thus suggest that the former order is grammatical albeit stylistically marked, while the latter order is ungrammatical, i.e. its real occurrences are speech errors (but cf. Zevakhina & Dzhakupova 2015: fn. 8 speculating that combinations of the two markers signal the development of a new, complex comparative clause complementizer).

(165) Ivan (prygaet) vyše ***čem*** ja.
Ivan.**NOM** (jumps) higher WH I.**NOM**
'Ivan {is taller; jumps higher} than I.'

(166) Ivan (prygaet) vyše ***neželi*** ja.
Ivan.**NOM** (jumps) higher NEG-EMPH-Q I.**NOM**
'Ivan {is taller; jumps higher} than I.'

(167) Ivan (prygaet) vyše ***neželi*** ***čem*** ja.
Ivan.**NOM** (jumps) higher NEG-EMPH-Q WH I.**NOM**
'Ivan {is taller; jumps higher} than I.'

In the other type of phrasal comparative, belonging to Stassen's locational type, the standard of comparison is introduced via genitive case marking on the standard of comparison DP. I distinguish two subtypes of this genitive of comparison, which I refer to as the *nominal* genitive of comparison, (168), and the *adjectival* genitive of comparison, (169), respectively, based on the morphological form of the standard argument.

(168) Ivan (prygaet) vyše ***menja.***
Ivan.NOM (jumps) higher I.**GEN**
'Ivan {is taller; jumps higher} than me.'

(169) Ivan (prygaet) vyše ***moego.***
Ivan.NOM (jumps) higher I.**POSS.GEN**
'Ivan {is taller; jumps higher} than me.'

This chapter is devoted to the nominal genitive of comparison, while a detailed study of the adjectival variety is reserved for future research.

6.2. Probing the structure behind the standard DP

In this section, I will perform several tests adopted and adapted from the existing literature (Merchant 2009 among others) that probe the structure behind phrasal standards by diagnosing whether the standard DP and its correlate DP in the

antecedent are contained in the same clause. The (nominal) genitive of comparison will be the focus of this examination and will be contrasted with the *čem*-comparative.

We shall see that single DP-standards in *čem*-comparatives are likely to have a clausal source, while in the genitive of comparison the standard and the correlate are clause-mates, and there is no direct evidence for a clausal structure behind the genitive DP.

First, let us look at what kind of standards are allowed in *čem*-comparatives and the genitive of comparison respectively.

The most obvious argument for a reduced clause analysis of phrasal *čem*-comparatives is that, unlike genitive comparatives, they admit full clauses in the standard position: compare (170) and (171).

(170) Ivan prygaet vyše ***čem*** prygaju ja.
Ivan.NOM jumps higher WH jump.1SG. I.NOM
'Ivan jumps higher than I jump.'

(171) *Ivan prygaet vyše (prygaju) **menja** (prygaju)
Ivan.NOM jumps higher jump.1SG I.**GEN** jump.1SG
'*Ivan jumps higher than me jump.'

Note the two superficial indications of the clausal nature of *čem*-standards: 1) the standard marker is morphologically a *wh*-item, similar to the one introducing embedded indicative and relative clauses, *čto* 'what.NOM/ACC'; 2) the morphological case of the standard corresponds to that of the correlate, nominative in (7), suggesting that there is a case-assigning element underlyingly.

Next, standards in *čem*-comparatives may be instantiated by multiple XPs that do not form a constituent, as shown in (172). This is, again, not possible in the genitive of comparison, (173).

(172) Ivan segodnja prygaet vyše
Ivan.NOM today jumps higher
čem Petja v ponedel'nik.
WH Petya.**NOM** in Monday

‘Ivan jumps higher today than Peter on Monday.’

(173) *Ivan segodnja prygaet vyše
Ivan.NOM today jumps higher
Peti v ponedel’nik.
Peter.**GEN** in Monday
‘Ivan jumps higher today than Peter on Monday.’

Sentence (172) with a DP PP string after the standard marker *čem* resembles gapping constructions, cf. (174). Gapping, an ellipsis-like construction where the verb is omitted, is known to occur in coordinated clauses under identity of the predicates; the remnant XPs in the verb-less clause contrast with correlate XPs in the antecedent clause.

(174) Ivan prygal segodnja a Petja v ponedel’nik.
Ivan.NOM jumped today CONJ Petya.**NOM** in Monday
‘Ivan jumped today and Peter on Monday.’

This standard-correlate contrast is also present in (172), so it is reasonable to analyze it as involving a similar kind of ellipsis.[60] If this is on the right track, we have evidence for the presence of two clauses in (172). This is in fact one of the arguments for the reduced clause analysis of phrasal comparatives, reviewed in Lechner (to appear).

In addition, *čem*-comparatives also admit non-DP standards, e.g. adverbial phrases or PPs, while the nominal genitive of comparison disallows them.

(175) Segodnja Ivan prygaet vyše
Today Ivan.NOM jumps higher
čem {obyčno; v subbotu}.
WH {usually; in Saturday}
‘Today Ivan jumps higher than {?usually; on Saturday}.’

[60] Note that the ellipsis mechanism involved might not be identical for (172) and (174): the missing verb in the gapping construction must match its antecedent in tense, while no such requirement is imposed on the *čem*-clause undergoing reduction.

(176) *Segodnja Ivan prygaet vyše {obyčno; v subbotu}.
Today Ivan.NOM jumps higher {usually; in Saturday}
'Today Ivan jumps higher than {?usually; on Saturday}.'

Let us now turn to locality tests, probing whether there is a clausal boundary between the standard and the correlate DPs in the two types of phrasal comparative.

The first diagnostic comes from n-word distribution. Russian is a negative concord language, in which clausal negation only licenses n-words in the same minimal CP. The availability of an n-word in the standard DP under matrix negation may thus serve as a diagnostic for absence of a clausal boundary. Negative concord is observed in the nominal genitive of comparison but not in the *čem*-comparative, as shown in (177) and (178) respectively, suggesting that the standard DP and its correlate are in the same CP in the former case, but in two different CPs in the latter.

(177) Maša *(ne) vyše nikogo.
Masha.NOM NEG taller nobody.GEN
'Masha isn't taller than anyone.'

(178) *Maša ne vyše, čem nikto.
Masha.NOM NEG taller WH nobody.NOM
'Masha isn't taller than anyone.'

To be sure, something that looks exactly like an n-word may sometimes occur as the nominative DP standard in the *čem*-comparative*:*

(179) Maša vyše čem nikto.
Masha.NOM taller WH nobody.NOM
'Masha is higher than (just) nobody.'

However, in (178) *nikto* 'nobody' is interpreted as a person of no importance and as such does not require matrix negation to be licensed, as shown in (179).

(180) Maša v ètoj strane nikto.
Masha.NOM in this country nobody.NOM
'Masha is 'nobody' in this country.'

This suggests that *nikto* 'nobody' is not a true n-word in sentences (179-180), hence its grammaticality in (179) does not entail that the correlate and the standard DP are clause-mates in *čem*-comparatives.[61]

Secondly, the nominal genitive standard DP can undergo *wh*-movement, obligatorily stranding the comparative adjective (181-182) and pied-piping the comparative adverb (183-184).[62]

(181) Kogo ty vyše (v klasse)?
Who.GEN you.NOM taller in class.LOC
'Who(m) are you taller than (in your class)?'

(182) *Vyše kogo ty (v klasse)?
Taller who.GEN you.NOM in class.LOC
'*Taller than whom are you (in your class)?

(183) Vyše kogo ty segodnja prygnul?
Higher who.GEN you.NOM today jumped
'?Higher than whom did you jump today?'

(184) *Kogo ty prygnul vyše (segodnja)?
Who.GEN you.NOM jumped higher today
'Who did you jump higher than today?'

[61] Maayan Abenina-Adar (p.c.) also points out that clausal comparatives (and hence their reduced versions) do not admit matrix negation in the comparative clause. For that reason, too, an n-word standard will not be licensed in the *čem*-comparative.
[62] This difference may have to do with the predicative vs. adjunct nature of adjectives and adverbs respectively: typically, extraction from adjuncts is blocked, hence the whole adjunct phrase should move.

By contrast, a standard DP, introduced by *čem,* may not undergo *wh*-movement, irrespectively of whether it moves alone, as in (185-186), or pied-pipes *čem* or *čem* and the comparative adjective/adverb, as in (187-190).

(185) *Kto ty vyše čem (v klasse)?
Who.NOM you.NOM taller WH in class.LOC
'Who are you taller than in your class?'

(186) *Kto ty prygnul vyše čem (segodnja)?
Who.NOM you.NOM jumped higher WH today
'Who did you jump higher than today?'

(187) *Čem kto ty vyše (v klasse)?
WH who.NOM you.NOM taller in class.LOC
'Who are you taller than in your class?'
lit. '*Than who are you taller in your class?'

(188) *Čem kto ty prygnul vyše segodnja?
WH who.NOM you.NOM jumped higher today
'Who did you jump higher than today?'
lit. '*Than who did you jump higher today?'

(189) *Vyše čem kto ty (v klasse)?
Higher WH who.NOM you.NOM in class.LOC
'Who are you taller than in your class?
lit. '?Taller than who are you in your class?'

(190) *Vyše čem kto ty prygnul segodnja?
Higher WH who.NOM you.NOM jumped today
'Who did you jump higher than today?'
lit. '?Higher than who did you jump today?'

A natural explanation for this difference between the genitive and the *čem*-comparative is that the genitive standard, but not the standard introduced by *čem*,

belongs in the same clause as the correlate and the comparative adjective/adverb. This is supported by the fact that extraction from a *clausal* standard introduced by *čem*, such as (191), is ungrammatical as well, as shown in (192).

(191) Ty prygnula vyše čem prygnul Miša.
you.NOM jumped.F higher WH jumped.M Misha.NOM
'You jumped higher than Misha jumped.'

(192) *Kto ty prygnula vyše čem prygnul?
Who.NOM you.NOM jumped.F higher WH jumped.M
'*Who did you jump higher than jumped?

In other words, both (185-186) and (192) are ungrammatical because finite clauses are islands for extraction in Russian.[63] Pied-piping of the standard marker *čem*, alone, (187-188), or together with the comparative adjective/adverb *vyše* 'taller/higher' as in (189-190), is then ungrammatical because wh-items typically do not pied-pipe complementizer-like elements, let alone elements from a higher clause.

The final diagnostic is built on the distribution of reflexive pronouns in the standard of comparison. The nominal genitive of comparison licenses reflexive pronouns in the standard DP, while *čem*-comparatives do not always do so.

(193) Maša ljubit sebja bol'še {✓svoego; *eë} otca.
Maša loves self.ACC more self's.GEN her father.GEN
'Masha$_j$ loves herself$_j$ more than her$_j$ father.'

(194) Maša ljubit sebja bol'še čem {*svoj; ✓eë} otec.
Maša.NOM loves self.ACC more WH self's.NOM her father.NOM
'Masha$_j$ loves herself$_j$ more than her$_j$ father (does).'

(195) Maša ljubit sebja bol'še čem {✓svoego; *eë} otca.
Maša.NOM loves self.ACC more WH self's.ACC her father.ACC

[63] This, in turn, might be due to the fact that the elements introducing indicative and comparative clauses are *wh*-words, filling Spec, CP and thus precluding successive-cyclic movement of a clause-internal phrase (cf. Mezhevich 2006).

'Masha$_j$ loves herself$_j$ more than (she loves) her$_j$ father.

Sentence (193) is ambiguous between a subject and an object-correlate reading and can be paraphrased as either (194) or (195). In contrast, *čem*-comparatives in (194) and (195) are not ambiguous: (194) with a nominative case-marked standard DP has the subject-correlate reading and (195) with an accusative case-marked standard DP has the object-correlate interpretation. Sentences (194) and (195) show that the reflexive pronoun is impossible in the nominative case-marked standard DP and obligatory in the accusative standard DP in *čem*-comparatives. The facts fall out if the standard DPs in (194-195) are analyzed as remnants of an elided clause. In that case, a reflexive possessive pronoun in a NOM standard is out because it can not be bound. Reflexive pronouns in Russian must be bound by a clause-mate subject DP, but this condition is not fulfilled in (194) since the reflexive itself is part of the subject of the relevant clause, as suggested by its NOM-marking. The ACC-marked reflexive in (195), on the other hand, is a part of the object of the underlying clause, hence it can be bound by an unpronounced subject, co-referent with *Masha*. This analysis receives support from the full clausal versions of (194) and (195), which pattern exactly like the reduced versions with respect to reflexive licensing, as shown in (196) and (197).

(196) Maša ljubit sebja bol'še
Maša.NOM loves self.ACC more
čem sebja/eë ljubit {*svoj; ✓eë} otec.
WH self/she.ACC loves self's.NOM her father.NOM
'Masha$_j$ loves herself$_j$ more than her$_j$ father loves himself/her$_j$.'

(197) Maša ljubit sebja bol'še
Maša.NOM loves self.ACC more
čem ona ljubit {✓svoego; *eë} otca.
WH she.NOM loves self's.ACC her father.ACC
'Masha$_j$ loves herself$_j$ more than she$_j$ loves her$_j$ father.'

As for the reflexive in the genitive standard (193), it must be licensed by the subject DP *Masha* and hence belong in the same clause as its correlate.

As summarized in Table 16, we have quite a few pieces of evidence for the reduced clause analysis of 'phrasal' comparatives introduced by *čem*, but no such evidence for the genitive of comparison. Instead, the locality data suggest that the standard and the correlate DPs are clause-mates in the nominal genitive of comparison.

Table 16. Properties of čem- and genitive comparatives vs. their Greek counterparts

	***Čem*-comparative**	**Greek *ap'oti*-comparative**	**Genitive comparative**	**Greek *apo*-comparative**
Only DP-standards?	✗	✗	✓	✓
Multiple XPs in SOC?	✓	✓	✗	✗
Clausal standards?	✓	✓	✗	✗
Fixed case on SOC?	✗	✗	✓	✓
Refl. pronouns in SOC?	✗/✓	✗/?	✓	✓
N-words in SOC?	✗	✗	✓	✓
Wh-movement of SOC?[64]	✗	✗	✓	✓
Verdict: CP-boundary?	✓	✓	✗	✗

Table 16 also contains data for two varieties of Greek comparatives (adapted from Merchant 2009), exemplified in (198) and (199).

(198) O Giannis ine psiloteros ***ap'oti*** i Maria.
The Giannis is tall.er.MSG from-WH the.NOM Maria
'Giannis is taller than Maria (is).'

(199) O Giannis ine psiloteros ***apo*** tin Maria.

[64] Used slightly differently in Merchant (2009).

The Giannis is tall.er.MSG from the.ACC Maria
'Giannis is taller than Maria.'

It can be seen from the table that the Russian *čem*-comparative patterns with the Greek *ap'oti*-comparative, while the Russian (nominal) genitive patterns with the *apo*-comparative.

Interestingly, while the *ap'oti*-comparative receives only a reduced CP-analysis in Merchant (2009), two analytical options are entertained for the *apo*-comparative: the Direct Analysis, whereby *apo* is treated as a P with the standard as its DP-complement, and a more abstract, reduced CP-analysis, involving movement of the standard into the *apo*-headed PP of the matrix clause, which we will consider in the next section.

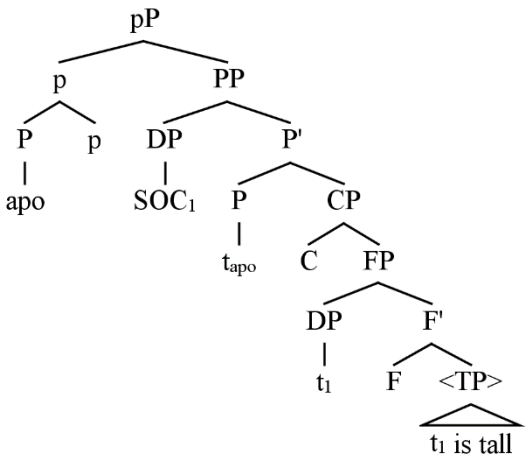
(200) O Giannis ine psiloteros [$_{PP}$ ***apo*** [$_{DP}$ tin Maria]].

In addition, the Russian (nominal) genitive of comparison, along with its Slavic counterparts (which bear more affinity to the Greek *apo*-comparatives, since their standards are marked by an equivalent preposition *ot/od* 'from'), receive a *small* clause analysis in Pancheva (2006), updated in Pancheva (2010). In other words, Merchant (2009) and Pancheva (2006, 2010) show that even when a phrasal comparative does not allow any overt material besides the standard DP and the locality tests tell us that the standard and its correlate are clause-mates, one cannot exclude the possibility that the standard is underlyingly clausal. Although the Direct Analysis is the most natural option for the Russian (nominal) genitive of comparison, it is also compatible with at least three versions of a clausal analysis, which I discuss in the next section.

6.3. Existing clausal analyses and evidence for abstract structure

In the reduced CP analysis put forward in Merchant (2009), locality facts are accounted for by postulating that the standard of comparison DP (SOC) raises out of the comparative CP into the PP-domain of the antecedent clause before the TP it originates in is elided. This is illustrated in Figure 9, with angle brackets marking the elided constituent.

Figure 9. Internal structure of standards in the Greek apo-comparative (based on Merchant 2009)



Pancheva (2006, 2010) proposes a small clause structure, where there is also no CP-boundary between the standard and the correlate. In Pancheva (2006) the small clause predicate is copied from the matrix clause at LF (after the degree quantifier *-er/more* undergoes QR and the *than*-PP undergoes late merge with it, cf. Figure 11, based on Pancheva 2006: ex. 51). That is, there is no syntactic content in the AP, as indicated by Ø in Figure 10.

Figure 10. Part of the syntactic tree for phrasal comparatives under Pancheva's (2006) analysis

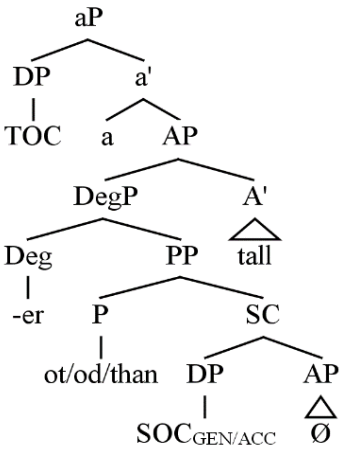
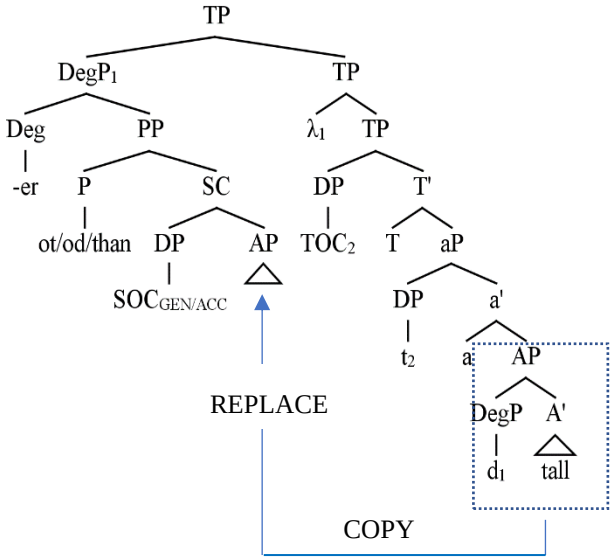


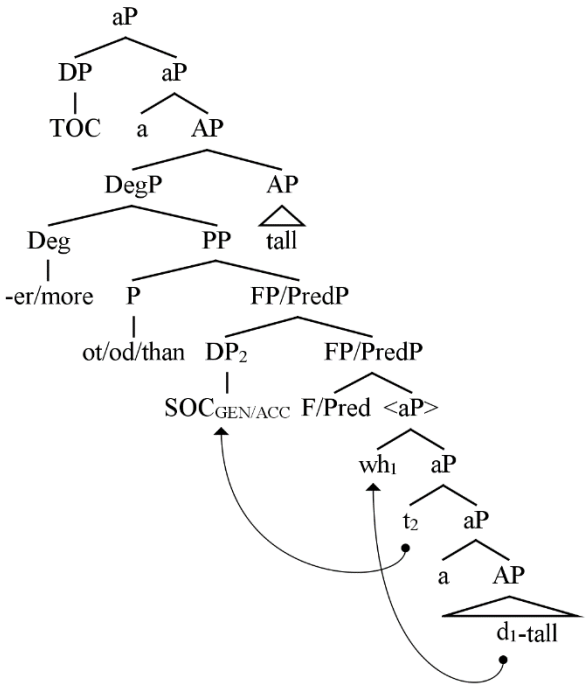
Figure 11. Pancheva (2006) LF structure for Slavic phrasal comparatives



Pancheva (2010) proposes a modification to the small clause analysis of Pancheva (2006), which consists in positing actual syntactic structure in the small clause predicate. Material in this predicate phrase undergoes PF-deletion under identity with the matrix/antecedent predicate, as shown by angle brackets in Figure 12.

The small clause is proposed to be a PredP in Pancheva (2010) and an FP (functional projection, probably corresponding to a focus phrase) in Pancheva and Tomaszewicz (2011). Pancheva proposes that a predicate of degrees is created via *wh*-operator movement in the small clause, which proceeds from the position parallel to that of the degree quantifier in the matrix clause to the most peripheral position of type <t> available, namely, the edge of the aP/vP. The standard of comparison DP originates within the aP/vP and moves to Spec, PredP/FP to be predicated of the newly formed degree predicate. In this position it also receives case from the prepositional standard-marker *ot/od* 'from', that is, is exceptionally case-marked by the preposition.

Figure 12. Part of the syntactic tree for Slavic phrasal comparatives under Pancheva (2010)



The small clause analysis looks appealing in view of the fact that the hypothetical small clause subjects, e.g. subjects of a secondary predicate, typically ECM-ed by the matrix verb, see (201), behave as if they belonged in the matrix clause. Examples (202-204) show that these accusative DPs behave similarly to the nominal genitive standard DPs as far as reflexive licensing, negative concord and A'-movement are concerned.

(201) On sčitaet Vasju umnym.
He.NOM considers Vasja.ACC smart.INS
'He considers Vasja smart.'

(202) On sčitaet sebja umnym.
He.NOM considers self.ACC smart.INS
'He considers himself smart.'

(203) On nikogo ne sčitaet umnym. [65]
He.NOM nobody.ACC NEG considers smart.INS
'He doesn't consider anyone smart.'

(204) Kogo on sčitaet umnym?
Who.ACC he.NOM considers smart.INS
'Who does he consider smart?'

Let me point out that non-direct analyses like those of Merchant and Pancheva must be supported by some syntactic evidence since there are no obvious signs of clausal structure behind the standard DP of the phrasal comparatives they consider (cf. Table 16). Note, however, that any such evidence is necessarily indirect and thus open to other interpretations. Evidence appealed to in Merchant (2009) and Pancheva (2010), Pancheva and Tomaszewicz (2011) comes from contrasts observed between two types of comparatives in Greek and Slavic respectively, which can be readily captured under a Clausal Analysis but are unexpected under the Direct Analysis.

Merchant (2009) proposes the structure in Figure 9 to capture the fact that the standard DP introduced by *apo* 'from' may not correlate with a DP that is located within an island, in contrast to the *ap'oti*-comparative, cf. (205).[66]

[65] Incidentally, the ACC-marked small clause subject cannot become or at least strongly resists becoming genitive under negation:

(i) On ne sčital {ni odnu igru; *ni odnoj igry}
He NEG considered no one.ACC game.ACC no one.GEN game.GEN
dostojnoj svoego vnimanija.
worthy.INS self's attention.
'He didn't consider any game to be worthy of his attention.'

I suggest that this is related to the strong tendency of ECM small clause subjects to be construed as specific (cf. Irimia 2016 who points out that SC subjects tend to be bear differential object marking in languages that morphologically distinguish between non-specific/indefinite/inanimate and specific/definite/animate objects). In Russian negative clauses, objects typically may either be marked accusative or genitive. It is known that the choice between the two options is conditioned by specifity and definiteness, with the generalization that the more specific the referent of the direct object DP is the more resistant it will be to receive genitive case marking under negation.

(205) Perisoteri anθropi menun sto kratos pu kivernai o Putin
More people live in.the state that governs the.NOM Putin
a. ap'oti o Bush.
from-WH the.NOM Bush
b. *apo ton Bush.
from the.ACC Bush.
'More people live in the country that Putin governs than live in the country that Bush governs.' (Merchant 2009: 142, ex. 31).

The crucial difference between the structures proposed for *ap'oti* and *apo* comparatives respectively is that the standard in the former undergoes additional movement that leaves a trace in the part of the structure that is not elided (Spec, FP in Figure 9). This trace is uninterpretable at PF since the tail of its chain is located within an island. However, Merchant (2009) also suggests an alternative, direct DP-analysis, deriving island sensitivity from constraints on covert movement of the correlate (*Putin* in (205)) rather than from overt movement of the standard. Thus, island-sensitivity may not necessarily speak for the elaborate CP-structure behind the standard.[67]

The small clause analysis of Pancheva (2010), Pancheva & Tomaszewicz (2011) is motivated by the observation that in *wh*-fronting languages the phrasal, but not the reduced clausal, comparative exhibits an asymmetry between object and (base-generated) subject amount comparatives, illustrated below for Polish.[68] (206) shows that a *more*-NP in the object position is fine with both the phrasal *od*-comparative and the reduced clausal *niż*-comparative, whereas (207) demonstrates that a *more*-NP in

[66] A brief study of Russian *čem*- and genitive comparatives in structures similar to (205) reveals a complex picture, not parallel to the one reported for Greek. I reserve an in-depth study of these for future research.

[67] In his later work (Merchant 2012) Merchant argues for the latter option, that is the Direct Analysis with covert movement of the correlate.

[68] *Wh*-fronting languages that are cited in Pancheva (2010), Pancheva and Tomaszewicz (2011) are Polish, Bulgarian, Serbian/Croatian, Slovenian, Greek and Hungarian; the non-*wh*-fronting languages are represented by Turkish, Hindi, Japanese and Korean. However, no actual examples are provided for languages other than Polish and Bulgarian. Moreover, data from Merchant (2009: 9b), constituting a phrasal comparative with a *more NP* subject might suggest that the generalization does not hold for Greek.

(ii) Perisoteri anθropi θelun na maθun anglika apo germanika
More people want.3PL SUBJ learn.3PL English from German
'More people want to learn English than German.'

the subject position is severely degraded in the *od*-comparative, but not in the *niż*-comparative.

(206) Marek zwiedził *więcej miejsc* {✓**od** Anny**;** ✓**niż** Anna}.
Marek visited more places from Anna.GEN; than Anna.NOM
'Marek visited more places than Anna (did).'

(207) *Więcej uczniów* zwiedziło Czechy{??/*__od__ Słowacji**;** ✓**niż** Słowację}
More students visited Czech R. from Slovakia.GEN; than Slovakia.ACC
'More students visited the Czech Republic than Slovakia.'
(adapted from Pancheva and Tomaszewicz 2011: ex. 2-3)

Pancheva (2010) proposes to account for the gradient acceptability of subject amount *od*-comparatives by appealing to constraints on *wh*-operator movement inside the unpronounced small clause behind the standard DP (see Figure 12). The presence of a subject/object *more*-NP asymmetry in the phrasal *od*-comparative and its absence in the reduced clausal *niż*-comparative is explained as follows. Both comparative structures involve a *wh*-operator in the standard clause, occupying a position parallel to that of the degree quantifier/comparative morpheme in the antecedent clause. This operator pied-pipes its quantified complement and A'-moves to the periphery of the clause, Spec, CP (or adjoins to CP[69]) in the *niż*-comparative, shown in (208-209), and to Spec, vP in the *od*-comparative, shown in (210-211), as this is the most peripheral position of type <t> available.

(208) …*niż* [$_{CP}$ ***wh-many places***$_1$ [$_{FP}$ Anna$_2$ [$_{TP}$ t$_2$ [$_{vP}$ t$_2$ visit **t**$_1$]]]]]^{70}
(209) …*niż* [$_{CP}$ ***wh-many students***$_1$ [$_{FP}$ Slovakia$_2$ [$_{TP}$ **t**$_1$ [$_{vP}$ **t**$_1$ visit t$_2$]]]]
(210) …*od* [$_{FP}$ Anna$_2$ [$_{vP}$ ***wh-many places***$_1$ [$_{vP}$ t$_2$ visit **t**$_1$]]]
(211) *…*od* [$_{FP}$ Slovakia$_2$ [$_{vP}$ ***wh-many places***$_1$ [$_{vP}$ **t**$_1$ visit t$_2$]]]

This movement is blocked in just one case, namely, when the operator is part of the underived subject NP, i.e. is generated in Spec, vP and moves to another Spec, vP. This is precisely the case of *more*-NP subjects in the *od*-comparative, (207), (211).

[69] Pancheva (2010) adopts Bare Phrase Structure (Chomsky 1995), which does not distinguish between specifiers and adjuncts.
[70] Shading marks the elided constituent (TP and larger vP respectively).

Adopting the Bare Phrase Structure approach which does not make a distinction between specifiers and adjuncts, Pancheva claims that such movement would be vacuous for being too local and hence must be blocked by an Anti-Locality principle, formulated in (212).

(212) The Specifier of a head H cannot move to a Specifier of H (Pancheva 2010: ex. 22)

To account for the fact that *more*-NP subjects are degraded but not completely unacceptable in the *od*-comparative Pancheva suggests that an alternative derivation involving subextraction of the *wh*-operator is available. This derivation, schematized in (213), does not violate the universal Anti-Locality constraint but does violate the violable Subject Island Constraint.

(213) **??/*** …*od* [**FP** Slovakia$_2$ [**vP** ***wh-many***$_1$ [**vP** **d**$_1$-students visit t$_2$]]]

In sum, the degraded status of *od*-comparatives with *more*-NP subjects has to do with the more general constraint on subextraction from subjects. The prediction then is that sentences like (207) with the standard DP introduced by *od* 'from' should be judged similarly to sentences with overt subextraction from subjects. This prediction was tested for Polish in the experimental study of Pancheva and Tomaszewicz (2011), where they concluded that it was confirmed. The greater degradation of *od*-comparatives with *more*-NP subjects compared to overt subject subextraction sentences that was found in Pancheva and Tomaszewicz (2011) was accounted for by two additional factors: the fact that *od*-comparatives are in general less acceptable than *niz*-comparatives[71] and that subject amount comparatives are in general less acceptable than object amount comparatives.

Note that Pancheva's account of the subject/object amount comparatives contrast in the *od*-comparative and its absence in the *niż*-comparative requires that the *wh*-operator pied-pipe the NP and move to Spec, CP in the latter. In other words, subextraction of the *wh*-operator should be ruled out in this construction to derive the absence of subject/object-NP asymmetry in the *niż*-comparatives. This analysis yields the following (empirical) problem. Note that unlike in the *od*-comparative, in the *niż*-comparative the moved *wh*-NP is not elided at PF, cf. (208-209) vs. (210-211). It is,

[71] I suggest that the reason might be the *od*-comparative being an older and less productive construction.

however, never pronounced, to the best of my knowledge. Notably, though, a *wh*-word (operator) *jak* 'how' does marginally occur in *niż*-comparatives, following *niż* and preceding the standard of comparison (as noted in Pancheva and Tomaszewicz 2010: p.190). Now, judging by Pancheva (2006), *jak* in comparatives is a *wh*-operator that ends in the Spec, CP position. If so, how would one explain why the quantified NP must always remain silent in these cases? Recall that postulating movement of the *wh*-operator alone is not an option, since it would predict a subject/object asymmetry in the *niż*-comparative as well, which is not observed.[72] The only solution seems to be to treat *jak* as something other than the moving *wh*-operator. This problem casts doubt on the correctness of the asymmetry account.

Abstracting away from these issues for a moment, let us see whether argumentation à la Pancheva could in principle be adopted for the Russian genitive of comparison.

Unfortunately, Pancheva's argument for a clausal structure does not seem to work for Russian, but not because "in Russian no nominal [that is, amount – *T.P.*] comparatives can be phrasal" (Pancheva 2010: fn.1). Although Russian equivalents to the Polish and Bulgarian sentences examined in Pancheva (2010) turn out to be quite bad, irrespectively of the locus of comparative, a minimal pair of sentences with a synonymous verb *navestit'* 'visit', requiring both subject and object to be animate, was much more readily accepted by the native speakers that I consulted. On a 1-5 discrete scale, where 1 corresponds to completely unacceptable and 5 – to fully acceptable, the object amount comparative in (214) received a mean judgement of 3.62 (SD=1.56) and the minimally different subject amount comparative, (215), was rated 3.31 (SD=1.38) on average, with the two means not significantly different (paired t-test, p=0.15, one-tail). 8/13 speakers judged them as equally acceptable, 4/13 speakers judged the object amount comparative more acceptable than the subject amount comparative and 1 speaker accepted (215) but not (214).

(214) Maša bol'še Miši navestila druzej.

[72] This analysis is nevertheless entertained in Pancheva and Tomaszewicz (2011), where the empirical problem it yields is not addressed. Note also that under this analysis it is not only that the subject amount comparative should be degraded compared to the object amount comparative in the *niż* case, but if anything, the subject island violation in the *niż* case should yield a worse result than its counterpart in the *od* case, since the head of the island-violating movement chain (the *wh-many*) is elided in the latter but not in the former (this argument is inspired by Merchant's idea that traces of PF-island-violating movements cause ungrammaticality unless elided, cf. Merchant 2004, 2009).

Masha.NOM more Misha.GEN visited.F friends.GEN
'Masha visited more friends than Misha.'

(215) Mašu bol'še Miši navestilo druzej.
Masha.ACC more Misha.GEN visited.N friends.GEN
'More friends visited Masha than (visited) Misha.'

The acceptability difference between (214) and (215) is thus unlikely to be categorical.

Now let us look at minimally different sentences involving object and subject subextraction.

(216) Skol'ko/skol'kix Maša navestila druzej?
How-many.NOM/ACC Masha.NOM visited.F friends.ACC
'How many friends did Masha visit?'

(217) Skol'ko Mašu navestilo druzej?
How-many.NOM Masha.ACC visited.N friends.GEN
'How many friends visited Masha?'

(216) and (217) received an average judgment of 4.77 (SD = 0.44) and 4.62 (SD = 0.51)[73] respectively from a pool of 13 speakers (identical to the ones who judged (214-215). 11 of the informants reported no acceptability contrast between these instances of object and subject subextraction and judged both as either 4/5 or 5/5, whereas 2 speakers rated (216) as 5 and (217) as 4, i.e. perceived a slight contrast.

We have seen that there is no significant acceptability difference either between object and subject amount phrasal comparatives, (214) and (215), or between object and subject subextraction in parallel wh-interrogatives, (216) and (217). Interestingly, while subextraction from both object and subject may be almost fully acceptable in Russian,[74] object and subject amount NPs are apparently somewhat degraded in the genitive of comparison.[75] This fact is rather irrelevant to the issue at hand, but the

[73] $p > 0.05$, one-tail on t-test (paired two sample for means/assuming equal variances)
[74] This finding calls for an in-depth study of the grammaticality status of subject subextraction.
[75] $p < 0.05$ one-tail t-tests, comparing (214) and (216), (215) and (217) respecrively.

degradation might be due to the availability of the more productive *čem*-comparative construction.

We conclude that argumentation à la Pancheva cannot be adopted for Russian. Therefore, an argument for an underlying clausal structure of a phrasal comparative (including the Russian genitive of comparison) should be built on a different type of evidence.

In the next section I present one such piece of evidence. It comes from a morphological constraint on correlates to the genitive standard and falls out from a clausal analysis under certain assumptions concerning case assignment. In contrast, the data are unexpected under the direct DP-analysis – some independent principle must be postulated to derive them.

6.4. Constraints on the correlate: an ellipsis account

In the (nominal) genitive of comparison construction, nominative and accusative DPs (subjects and direct objects) can be freely construed as targets of comparison, that is, correlates that the genitive SOC DP is contrasted with. For this reason, comparative sentences with transitive predicates may be ambiguous between the subject and the object reading. This is exemplified in (218).

(218) Ja uvažaju Petrovu bol'še
I.NOM respect Petrova.**ACC** more
{Ivanova; Ivanovoj}.
Ivanov.GEN/ACC Ivanova.GEN/DAT/INS/LOC
'I respect Petrova more than Ivanov/Ivanova.'

NOM-reading: 'I respect Petrova more than {Ivanov; Ivanova} does.'
ACC-reading: 'I respect Petrova more than I respect {Ivanov; Ivanova}.'

By contrast, DPs that are neither nominative nor accusative do not always qualify as correlates. Consider sentence (219) with a (morphologically reflexive) transitive verb *gordit'sja* 'to be proud', whose internal argument DP bears instrumental case. We have just seen in (218) that the NOM/ACC ambiguity is observed with both

morphologically masculine and feminine surnames in the standard of comparison position (*Ivanov* and *Ivanova*). However, in (219), which involves an instrumental-marked rather than an accusative-marked complement DP, only the variant with a feminine surname is ambiguous between the nominantive and the instrumental reading, while the masculine surname variant only has the nominative (subject) reading, i.e. the oblique instrumental reading is blocked.

(219) Ja gordžus' Petrovym bol'še
I.NOM proud Petrov.**INS** more
{Ivanova; Ivanovoj}.
Ivanov.GEN/ACC Ivanova.GEN/DAT/INS/LOC
'I am proud of Petrov more than Ivanov/Ivanova.'

NOM-reading: 'I am proud of Petrov more than {Ivanov; Ivanova} is.'
INS-reading: 'I am proud of Petrov more than of {*Ivanov; Ivanova}.'

Thus, only some bare oblique DPs qualify as correlates to the genitive standard of comparison. I suggest that the choice of the masculine/feminine surname in the standard DP makes a difference here just because the genitive case form of the feminine variant coincides with its form in the instrumental case (case that appears on the internal argument DP), while the genitive form of the masculine variant does not display genitive/instrumental syncretism.

DPs that are introduced by prepositions (P-complements), however, can never be correlates to the nominal genitive standard DP, irrespectively of the morphological (case) properties of the standard DP. Thus, sentences like (220) with a verb taking a PP-complement only have subject correlate readings.

(220) Ja verju v Petrova bol'še
I.NOM believe in Petrov.**ACC** more
{Ivanova Ivanovoj}.
Ivanov.GEN/ACC Ivanova.GEN/INS/DAT/LOC
'I believe in Petrov more than Ivanov/Ivanova.'

NOM-reading: 'I believe in Petrov more than {Ivanov; Ivanova} does.

PP-reading: 'I believe in Petrov more than I believe in {*Ivanov; *Ivanova}.'

Based on these data we can formulate a constraint that is imposed on the Russian nominal genitive of comparison, which I dub the Oblique Correlate Constraint (OCC).

(221) *The Oblique Correlate Constraint (OCC)*

a. A nominal genitive standard DP may have an oblique correlate iff its morphological form is syncretic for genitive and the relevant oblique case.

b. A nominal genitive standard DP may not have a P-object or a PP correlate.

The OCC can be rather straightforwardly derived from either the reduced CP-structure proposed in Merchant (2009) or from the elliptical small clause structure developed in Pancheva (2010), Pancheva & Tomaszewicz (2011). Specifically, the OCC facts fall out if one makes the following two assumptions.

The first assumption is that a DP may receive multiple case values in the course of the derivation. Such an approach to case has been independently advocated in Béjar and Massam (1999), Merchant (2006) and Richards (2013) and is also adopted in Merchant's (2009) reduced CP analysis of the Greek *apo*-comparative.

The second assumption is that structural, but not inherent/lexical case values may be "overwritten", that is, not realized morphologically. The tendency for preservation of inherent/lexical case morphology has been observed at least for Norwegian in Béjar and Massam (1999) and for Lardil in Richards (2013). The former authors capture the inherent/structural case contrast by proposing that inherent case is more marked, while Richards holds the view that inherent case is semantically *interpretable*, and structural case is semantically *uninterpretable*. I will assume that inherent case must be morphologically realized because, unlike structural case, it encodes information that cannot be retrieved otherwise (e.g. read off the structure).

I will further assume that there are two mechanisms responsible for multiple case value resolution in Russian: Overwrite and Match.[76] Overwrite instructs the morphology to realize the last assigned case and Match resolves the conflict via insertion of a syncretic morpheme. Both can freely apply to all case value

[76]Equivalent to Attraction and Matching of Assmann et al (2014).

combinations, but the former will yield an ungrammatical result if the case value to be overwritten is inherent/lexical.

I adopt and adapt the small clause analysis à la Pancheva (2010), Pancheva and Tomaszewicz (2011), rather than Merchant's (2009) reduced CP alternative because the latter requires postulating more silent structure and presupposes the availability of DP-raising to the matrix clause Spec, PP out of an embedded CP, cf. Figure 5. Although there is some independent evidence for this kind of raising in Modern Greek (see Merchant 2009 and references there), nothing of the kind has been proposed for Russian, to the best of my knowledge.

Let me now show how the facts captured by the Oblique Correlate Constraint are derived from the small clause structure à la Pancheva and the case assignment assumptions just introduced. The crucial part of the structure, deriving the object (accusative) reading of sentence (218) is presented in (222), with glosses used for ease of exposition.

(222) $[_{DegP}$ -er $[_{PP}$ $Ø_{ot}$ $[_{FP}$ Iv.$_1$ $[_{vP}$ $[_{AdvP}$ *wh* much$]$ $[_{vP}$ I $[_{v'}$ respect $[_{VP}$ $[_{VP}$ tV $t_1]$ $t_{AdvP}]]]]]]]$

GEN ACC

The structure behind the genitive standard of comparison, i.e. *Ivanovoj*, is an FP (as in Pancheva and Tomaszewicz 2011).[77] F takes a *v*P as its complement; there is no TP or CP layer in the structure. The *v*P behind the standard is identical to the *v*P of the matrix clause except for the standard-to-be DP. There is a *wh*-operator in the small clause, originating in the position parallel to that of the degree head in the matrix and moving to the most peripheral position of type <t> available, i.e. the edge of the *v*P, thereby creating a degree predicate.[78] The SOC-to-be moves from inside the *v*P to Spec, FP. At PF, the maximal *v*P undergoes ellipsis. The standard DP in (218)/(222) originating as the direct object of the verb *uvažat'* 'to respect', gets a [ACC] value from the *v*+V combination. After moving to Spec, FP it receives [GEN] from the null variant of preposition *ot* 'from' via Exceptional Case Marking. Under standard assumptions both these cases are structural, rather than inherent (since they are not

[77] Pancheva and Tomaszewicz (2011) suggest that FP is equivalent to a focus phrase, but do not commit to it.

[78] The *wh*-operator may move on its own if the phrase it is contained in allows subextraction (cf. discussion in the previous section). In our case, however, the *wh* is within an AdvP, commonly considered an adjunct. Since adjuncts are islands for extraction, I assume that the entire AdvP has to move.

assigned in conjunction with a theta-role). Thus, the standard-of-comparison DP gets a structural genitive on top of a structural accusative. The morphological case that we see on the SOC must be genitive, meaning that the structural genitive 'overwrites' the structural accusative. In more general terms, the structural case that is assigned last in the derivation is the one that gets morphologically realized. Overwriting the 'original' structural accusative by structural genitive incurs no penalty, since structural case values do not have to be morphologically realized. The proposed derivation thus accounts for the fact that accusative direct objects are typically allowed as correlates to the genitive standards.

Now let us see what happens if the first case value that the SOC DP receives is inherent/lexical. The instrumental case that the verb *gordit'sja* 'to be proud of' requires its internal argument to bear is likely an instance of such a case. Following Bailyn (2012), I assume that the assignment of instrumental case is a lexical property of such verbs and they assign it upon Merge with their internal argument.[79]

(223) $[_{DegP}$ -er $[_{PP}$ Ø$_{ot}$ $[_{FP}$ Iv.$_1$ $[_{vP}$ $[_{AdvP}$ *wh* much$]$ $[_{vP}$ I $[_{v'}$ proud $[_{VP}$ $[_{VP}$ t$_V$ t$_1$$]$ t$_{AdvP}$$]]]]]]]$

GEN INS

The SOC DP bearing lexical [INS] and structural [GEN] case feature values must be both instrumental and genitive from the morphological point of view. This is so because on the one hand, morphology must be faithful to lexical case and on the other hand, the genitive of comparison cannot possibly be formed when SOC is not morphologically genitive. Given the morphological resources of Russian, which lacks overt case-stacking, these conflicting requirements can only be satisfied if the morphological form of the SOC displays genitive-oblique syncretism, i.e. under Case Matching. This derives Part (a) of the Oblique Correlate Constraint.

What about Ppart (b), namely, the impossibility of PP-/P-complement correlates? Let us look at how sentences like (220) are derived under the small clause analysis.

(224) $[_{DegP}$ -er $[_{PP}$ Ø$_{ot}$ $[_{FP}$ Iv.$_1$ $[_{vP}$ $[_{AdvP}$ *wh* much$]$ $[_{vP}$ I $[_{v'}$ believe $[_{VP}$ $[_{VP}$ t$_V$ $[_{PP}$ in t$_1$$]]$ t$_{AdvP}$$]]]]]]]$

✗

[79] For a slightly different approach to the Russian instrumental case see Pereltsvaig (2007).

I assume that the whole PP *v Ivanova* 'in Ivanov' may not move from inside the VP to Spec, FP because PPs are not licit subjects of a small clause predicate (perhaps, their semantic type is not right). Thus, the only way to derive the desired interpretation is to move the DP *Ivanov* out of the PP. This, however, must be blocked in a non-P-stranding language like Russian: unlike certain island violations, preposition-stranding is not ameliorated by ellipsis (cf. Abels 2003), so deletion of the vP does not repair the violation. Thus, there is no licit derivation in Russian that would lead to the PP/P-complement correlate interpretation of the nominal genitive comparative. We have thus derived OCC (b) as well.

The most salient, nominative readings of sentences (65-67) are derived via movement of the small clause subject from Spec, *v*P to Spec, FP. Since there is no NOM-assigning T-head in the small clause, these standard DPs receive only one case value, namely, [GEN] from the null preposition *ot* 'from'. NOM-readings are then always available because there is no multiple case value conflict.

6.5. Correlate constraints beyond Russian

If correlate constraints like OCC provide evidence for multiple case assignment and thus clausal structure, their presence in a phrasal comparative that admits only single DP standards may be used as a diagnostic for clausal structure in various languages.

Languages indeed differ in whether their phrasal comparatives impose correlate constraints like the OCC. Data for the Hindi-Urdu *se*-comparative from Bhatt and Takahashi (2011), who argue for the Direct DP analysis on independent grounds, show that, in addition to nominative (zero-marked) subjects, Hindi allows ergative, dative and even P-complement DPs as correlates, despite lack of syncretism in the case system.

(225) Atif-ne Mina-ko Tina-se zyaadaa tohfe diye
Atif-ERG Mina-DAT Tina-from more presents.M give.PFV
'Atif gave more presents to Mina than Dina.'
ERG-reading: 'Atif gave more presents to Mina than Tina did.'
DAT-reading: 'Atif gave more presents to Mina than to Tina.'

(226) Amrika-me Rus-se zyaadaa log rah-te haiN
America-in Russia-from more people stay-HAB.MSG be.PRS.PL
PP-reading: 'More people live in America than in Russia.'
(Bhatt and Takahashi 2011: 592-3, ex. 21)

This is quite different from Russian. If we take OCC-like constraints as evidence for multiple case assignment and hence an underlying complex structure, absence of OCC-effects may be taken as an additional argument *against* clausal analysis.

The English phrasal comparative does not impose correlate constraints either, as shown in (227-229), but this fact is less notable, since the language only has rudiments of the morphological case system. The fact that OCC (b) ruling out PP-correlates does not apply in English is not in itself an argument for the direct DP analysis. This falls out from the standard reduced CP-analysis, in which the standard DP moves out of the TP to be elided, since P-complements may move out stranding the P in English. Note that in Hindi-Urdu P-stranding is blocked, so the grammaticality of (226) may not be accounted for in this way.

(227) Atif gave more presents to me than Tina.
Subject-reading: Atif gave more presents to me than Tina did.
PP-reading: Atif gave more presents to me than to Tina
(Bhatt and Takahashi 2007: 25, fn. 8)

(228) You've got more money to make **out of** me than them.
[Jack London. *The Burning Daylight*. 1910]

(229) I thought that the burden of directing and warning would be more efficiently borne **by** him than me. [Emily Brontë. *Wuthering Heights*. 1847]

By contrast, OCC appears to hold in the Hungarian phrasal comparative in which the standard DP is marked by adessive case (which also makes it typologically similar to the Russian genitive of comparison). Wunderlich (2001) examines a variety of such sentences, concluding that only nominative, accusative and dative DPs can be contrasted with the adessive standard, that is, only such DPs are acceptable correlates.

(230) (Én) Péter-t jo-bb-an szeret-em nál-ad.
I Peter-ACC good-COMP-ADV love-1SG ADESS-2SG
'I love Peter more than you.'
NOM-reading: 'I love Peter more than you do.'
ACC-reading: 'I love Peter more than I love you.'
(adapted from Wunderlich 2001: ex. 19)

(231) Anná-nak Péter-nél nehez-ebb franciául olvas-ni.
Anna-DAT Peter-ADESS difficult-COMP French read-INF
DAT-reading: 'For Anna it is more difficult to read French than for Peter'
ACC-reading: #'For Anna it is more difficult to read French than Peter'
(adapted from Wunderlich 2001: ex. 23)

Especially telling is the minimal pair in (232-233). The verb *telefonál* 'to call up' can in principle take either a dative or an instrumental DP referring to the person being called up. However, if one wants to say that *he calls X more often than Y*, only the dative option is available. In other words, an instrumental DP may not be a correlate to the adessive standard: for that reason, in (233) the adessive DP is necessarily interpreted as a locative phrase.

(232) Gyakra-bb-an telefonál-t-am nek-i Péter-nél.
often-COMP-ADV telephone-PAST-1SG DAT-3SG Peter-ADESS
'I called him up more often {than; at} Peter.'
DAT-reading: 'I called him up more often than I called up Peter.'
ADESS=LOC reading: ??'I called him up more often at Peter.'

(233) Gyakra-bb-an telefonál-t-am vel-e Péter-nél.
often-COMP-ADV telephon-PAST-1SG INSTR-3SG Peter-ADESS
'I called him up more often {than; at} Peter.'
INS-reading: *'I called him up more often than I called up Peter'
ADESS=LOC reading: 'I called him up more often at Peter's.'
(adapted from Wunderlich 2001: ex. 26)

The fact that DPs bearing semantic case are not possible correlates to the adessive standard is actually expected under the analysis proposed in Section 6.4 for Russian. Since Hungarian has neither case syncretism, nor case-stacking (of the Lardil type addressed in Richards 2013), there is no way for the standard DP to morphologically realize both the 'downstairs' semantic case and the 'upstairs' adessive case. A structural case, in turn, can be overwritten by adessive, so structural case-marked DPs are acceptable correlates.

It is interesting that dative case in Hungarian behaves on a par with structural cases. In fact, Wunderlich (2001: fn. 5) mentions independent arguments for analyzing the Hungarian dative as structural. This suggests that the ability of a case-marked DP to serve as a correlate to a phrasal standard (if OCC effects are present) may also be potentially used to diagnose structural case.

6.6. Correlate Constraints: back to Russian

6.6.1. Is OCC too weak?

It turns out that Russian speakers vary as to which readings of the sentences in (218-220) they find available. The speakers I initially asked (4 including myself) for whom the range of readings is constrained by the OCC can be considered the most liberal ones. Another group of speakers did not accept the oblique readings irrespectively of syncretism, allowing only NOM- and ACC-correlates. Still others accepted only the NOM-correlate interpretations of (218-220). These correspond respectively to Group 1, 2 and 3 in Table 17.

How can we model this variation under the small clause analysis proposed in Section 6.4? I have suggested above that there exist two mechanisms for multiple case value resolution in the morphological component of Russian – Overwrite and Match. Since morphology must be faithful to lexical/inherent case values, only Match produces a grammatical result for multiple case derivations involving case values of this type, whereas derivations involving multiple case values of the structural type can be resolved by either Match or Overwrite. The speakers then differ in which mechanisms for case value resolution they have in their grammars, as illustrated in Table 17.

Since the derivation of NOM-readings does not involve multiple case checking, they would be available for any speaker and would be the only possible ones for speakers lacking both Overwrite and Match (Group 3). Group 2 speakers only have Overwrite and hence will only allow structurally case-marked correlates. Group 1 speakers have both Overwrite and Match and thus admit all structurally case-marked correlates and inherent case-marked correlates, in case the OCC is obeyed. Finally, the system provides an additional logically possible grammar (Group 4), which only has Match. Such speakers (if they exist) would require all non-NOM-construals to have SOCs that are syncretic between the genitive and the relevant non-NOM case.

Table 17. Modelling inter-speaker variation

	Group 1	**Group 2**	**Group 3**	**Group 4**
OVERWRITE	+	+	-	-
MATCH	+	-	-	+
Readings allowed	NOM, ACC, syncretic obliques	NOM & ACC	NOM	NOM, syncretic non-NOM

An important question that arises in the context of speaker variation is how categorical the judgements of Group 2 and Group 3 speakers are. Since I only consulted a small set of informants about isolated sentences, it may be premature to conclude that the readings they did not accept are necessarily ruled out by the morpho-syntactic properties of their grammar. It may be the salience of nominative readings that has led to the rejection of the oblique readings on the part of some speakers. The assessment of the grammatical status of genitive comparative sentences with non-nominative readings require an independent controlled study, collecting judgments on a wider range of such sentences presented in and out of context. This study will likely have to involve several psycholinguistic experiments, which is beyond the scope of research presented in this chapter; I plan to to engage in this study in future research.

6.6.2. Is OCC too strong?

While the previous subsection shows that the OCC may be too weak, here I present cases that violate the OCC and thus suggest that the proposed analysis might undergenerate.

Russian personal pronouns only display GEN/ACC syncretism, so oblique correlates are expected to be disallowed with such standards. However, corpus searches of the adverbial *bol'še* 'more' with 3rd person pronouns revealed 3 such examples,[80] out of 70 relevant instances (62 involved a nominative correlate and 5 – an accusative correlate); they are presented in (234-236). Each exceptional instance has a different type of correlate: a dative indirect object, a dative subject-like DP and the possessor-introducing PP, headed by *u* 'at'.

(234) Bol'še nego krasnuju kartočku pokazyvali tol'ko Ju. Kovtunu.
More he.GEN red.ACC card.ACC showed.PL only Y.DAT Kovtun.DAT
'It was only Yury Kovtun that they gave the red card to more than to him.'
[Premjer-liga. 27 tur. *Sovetskij Sport*, 2006.11.09]

(235) Bol'še nix udalos' zarabotat' tol'ko PIFam.
More they.GEN managed.N.REFL earn only open-end funds.DAT
'It was only open-end funds that managed to earn more than them [bond funds].'
[I. Šlygin. *RBC Daily*, 2013.01.11]

(236) Bol'še nego iz igrokov sbornoj tol'ko u…Malkina.
More he.GEN from players.GEN team.GEN only at Malkin.GEN
'Of all national team players, it is only Malkin who has [scored] more [goals] than him.'
[S. Galajdo. *Sovetskij Sport*, 2012.05.18]

Another example of an apparent OCC-violation is the judgment pattern for a sentence with a topicalized oblique (dative) DP:

[80] All of them come from the Newspaper Subcorpus of the Russian National Corpus, available at http://ruscorpora.ru/search-paper.html.

(237) Maše$_1$ ja plaču t_1 bol'še Miši.
Masha.DAT I.NOM pay more Misha.GEN≠DAT
'I pay Masha more than Misha.'
DAT-reading: 'I pay Masha more than I pay Misha.'
NOM-reading: 'I pay Masha more than Misha pays Masha.'

Only one of my 6 informants (a linguist) accepted just the NOM-reading (with some hesitation), as predicted by the analysis. The remaining 5 speakers got the OCC-violating DAT-reading first. Only three of them also got the NOM-reading afterwards. Two of these three furthermore said that NOM was in fact the only acceptable one, while one speaker still accepted both. Thus, for 1 speaker the sentence only allowed a NOM-correlate reading, for 2 speakers the sentence only had a DAT-correlate interpretation and 3 speakers found it ambiguous, with questionable acceptability of the DAT-reading. Two things are notable here: the considerable pre-judgment hesitation of the OCC-abiding speaker and the rejection of the initially accepted OCC-violating DAT-reading on the part of two speakers. These raise the question of whether the unexpected acceptability of the DAT-reading of (237) translates into grammaticality or whether topicalization of the dative DP biases the speakers into the ungrammatical DAT-reading, i.e. creates a grammatical illusion.

The same question is relevant for sentences in (234-236) as we know that not all real-occurring utterances correspond to grammatical sentences. To assess the grammaticality of sentences like (234-237) under the oblique reading, one would need to quantitatively compare the respective acceptability of nominative and oblique readings in disambiguating contexts. If they turn out to be equally acceptable, the proposed small clause analysis with multiple case-checking will only be able to capture this if the relevant oblique Cases are shown to be structural (e.g. all instances of DAT case in Russian are structural for Pereltsvaig 2007). If the latter is impossible (as it is for *u*-PP correlates), the unexpected acceptability of the oblique reading will have to be accounted for by different means. The fact that all the exceptions to OCC involve a non-basic word order, with topicalization of the comparative adverb + standard phrase in (234-236) and of the dative DP in (237), as well as obligatory (sic!) modification of the correlate by the focus sensitive *tol'ko* 'only' in (234-236) suggest that the interpretation of phrasal comparatives is sensitive to the information structure

of the sentence. Thus, we cannot exclude the possibility that correlate constraints are of an information structural nature.

6.7. Conclusion and questions for further research

This chapter focused on the internal structure of standard DPs in the Russian genitive of comparison. I ran several locality tests, which suggested that the genitive standard DP and its correlate were clause-mates. Whereas the Direct Analysis remained the null hypothesis at this stage, the data were shown to be compatible with certain Reduced Clause Analyses (Merchant 2009, Pancheva 2010) as well. I then offered novel evidence for a clausal structure underlying the genitive of comparison. I showed that the genitive of comparison obeys the Oblique Correlate Constraint (OCC), which blocks PP-or P-complement correlates and imposes a syncretism constraint on the standard DP in the case of oblique correlates. I showed how the OCC can be derived from a small clausal structure à la Pancheva (2010), employing the assumptions that DPs may receive multiple case values in the course of the derivation and that inherent/lexical case features have to be morphologically realized, unlike structural case features.

Next, I briefly looked at correlate constraints in languages other than Russian. I demonstrated that the analysis correctly predicted the range of correlates available in the Hungarian adessive of comparison. I also suggested that the absence of correlate constraints in the Hindi-Urdu *se*-comparative provided additional evidence for the Direct Analysis of Bhatt and Takahashi (2011).

Finally, I discussed additional Russian data suggesting that the OCC might be both too strong and too weak and showed that the proposed analysis can explain the former if speakers differ in which multiple case resolution strategies are available in their idiolect and can capture the latter if the relevant oblique cases can be shown to be structural. I then noted, however, that the grammatical status of the non-complying data should be clarified: speakers who rejected OCC-*abiding* oblique readings might have done so for information-structural reasons, while real occurrence or acceptability of OCC-*violating* oblique readings may not translate into grammaticality.

To fully understand the nature of correlate constraints, one would need to systematically compare the acceptability of NOM and non-NOM readings (especially, oblique readings) in disambiguating contexts. If acceptability ratings are significantly

different, with oblique readings approaching the acceptability of ungrammatical controls, the small clause analysis proposed here will be supported. If they turn out to be of comparable acceptability, correlate constraints observed in isolated sentences would best be explained by appealing to information-structural or processing constraints. In the latter scenario, all things being equal, one would opt for the simpler, DP-analysis of the genitive standard. I leave this issue for future research.

7. Conclusions and implications

In this dissertation I proposed a morphosyntactic account of licensing of the so-called *n*-forms, a special series of 3rd person pronouns that have traditionally been called *pripredložnye* "adpositional", based on their tight connection to prepositional environments.

I showed that we can indeed state that *n*-forms are only licensed on the immediate complement of a P, thus maintaining the traditional intuition that *n*-forms appear in the prepositional object position.

However, keeping that simple and appealing idea required a close examination of the behavior of individual prepositions with respect to the N-factor, since the distribution of *n*-forms in the language is quite intricate and the traditional informal description cannot capture it.

Based on corpus data, I showed that some prepositions require *n*-forms (Class I), some prepositions allow both *n*-forms and the unmarked *j*-forms (Class II) and some block *n*-forms, only allowing *j*-forms (Class III). Secondly, both *n*-forms and *j*-forms are licensed in the standard of comparison DP of the Russian phrasal comparison, where no preposition appears.

The second fact had long been a puzzle to linguists and precluded a uniform explanation. Nevertheless, since *n*-forms do not appear in any environment beyond P-complements and genitive standards in the phrasal comparative, I strived for a uniform solution.

The proposal put forward in this thesis is that *n*-forms are licensed only by P-heads and the variation among prepositions stems from non-homogeneity of their categorial status. Thus, I argued that Class I prepositions are P-heads, Class II prepositions are not fully lexicalized P-heads, and oscillate between P-heads and N-heads or P-N combinations: *n*-forms are obligatory in the former case and *j*-forms – in the latter. Class III prepositions are not P-heads: some of them preserve the categorial status of their source lexeme (adjective/adverb or a verb), whereas in those of the P-N composition, either the nominal element licenses and assigns theta-related inherent case to the pronominal object or the case of the pronoun is licensed externally to the 'preposition'. This accounts for the impossibility of *n*-forms with Class III lexemes. In phrasal comparatives, *n*-forms were argued to be licensed by a null P-head, whose existence was independently motivated, whereas *j*-forms were proposed to be licensed by a quantificational head (the comparative quantifier) in a different syntactic configuration.

Based on historical data, contemporary corpus data and acceptability judgments of a large number of native speakers on n-forms in the single P-complement and n-forms in the coordinated P-complement positions, I proposed that in the history of Russian, the initial 'n' of 3rd person pronouns evolved from a (3rd) person object agreement marker on the P into a morphological reflex of a [P-case] superfeature. The current status of 'n' as a case-marker is evident from clear-cut acceptability judgments of certain younger speakers. For other informants the licensing mechanism for n-forms is not as obvious, as their acceptability judgments vary, reflecting the incomplete change of the morphosyntactic status of the initial 'n'.

In addition, I pointed out that the fate of *n*-form/*j*-form contrast was not the same in different Slavic languages. Rather they fall into three types, with respect to this factor:

(1) Languages where *n*-forms distinguish P-complements from complements of other heads (or P-case from other types of case). These are represented by Russian, Ukrainian and West Slavic;
(2) Languages where *n*-forms have evolved into prosodically strong pronouns as opposed to second position pronominal clitics; they are also obligatory in P-complement position, apparently since most P-heads are phonologically weak. This is true for South Slavic languages.
(3) Languages where *n*-form/*j*-form contrast had been lost and the 'j' stem generalized across case forms, including nominative. This property characterizes Belorussian and certain Russian dialects.

The research presented in this dissertation has implications for the typology of case, since it introduces a distinction between cases assigned by prepositions and cases assigned otherwise. As far as I can see the prepositional/non-prepositional case distinction is orthogonal to the inherent (lexical)/structural case distinction. If the approach taken here is on the right track, it has interesting implications for Case theory as well.

In addition, the present study contributes to the classification and typology of pronouns. I showed in this thesis that a formal distinction between pronominal forms may be due to factors other than differences in their internal structure, which is what the prominent generative typologies of Cardinaletti and Starke (1999) and Déchaine and Wiltschko (2002) are built on.

The approach taken here may also prove useful for the investigation and classification of adpositions in other languages, especially other Slavic languages.

Finally, the classification of prepositions proposed here may lead to a better understanding of the variable behavior of adpositions in various constructions, such as P-omission under sluicing and P-stranding.

In this dissertation I also dealt with the internal structure of the Russian phrasal comparative – the genitive of comparison. I offered novel indirect evidence for the Reduced Clause Analysis, involving silent structure behind the genitive standard DP. I suggested that the new morphosyntactic observation, dubbed the Oblique Correlate Constraint, might be relevant for probing the structure of phrasal comparatives in other languages as well.

In order to derive the facts, I assumed the availability of Multiple Case Checking even in languages that do not show overt stacking of case morphemes. This adds to the growing corpus of studies, arguing for that mechanism on independent grounds, which definitely has implications for Case theory.

Finally, I discussed that the issue of what kind of DPs may serve as correlates to the genitive standard of comparison is intricate, since the judgments are quite variable. I showed that the availability of a certain DP as a correlate may also depend on word order and prosody, which suggests that the choice of the correlate might be constrained by information-structural factors as well. I came to the conclusion that a series of controlled acceptability judgment studies are necessary for a better understanding of the nature of correlate constraints and decided to leave it for future research.

8. Appendix

A1. *Distribution of n-forms and n-less forms with different prepositions*[81]

Class I: Requiring n-forms (j-forms </= 1%)					
Preposition	***Frequency in ipm (rank among all lexemes)***[82]	***3rd person pronouns***	***Total no. of docs***	***Of all (%)***	***Case on the PRN***
Simple & double					
k 'towards/to'	5389 (15)	j-forms	20	0.04	DAT
		n-forms	51375	99.96	
po 'along'	5786.7 (11)	j-forms	5	0.05	
		n-forms	10873	99.95	
v 'in'	31374.2 (2)	j-forms	24	0.05	LOC
		n-forms	49846	99.95	
na 'on'	15867.3 (4)	j-forms	7	0.08	
		n-forms	9088	99.92	
pri 'at'	1550.8 (63)	j-forms	4	0.22	
		n-forms	1776	99.78	
o/ob 'about'	3407.1 (31)	j-forms	5	0.03	
	714.5 (129)	n-forms	14769	99.97	
pod 'under' (locative)	1126.0 (81)	j-forms	0	0	INSTRUMENTAL
		n-forms	1208	100	
nad 'above'	537.0 (170)	j-forms	0	0	
		n-forms	2319	100	
pered 'in front of'	626.7 (143)	j-forms	0	0	
		n-forms	2601	100	
za 'behind'	3904.1 (24)	j-forms	2	0.03	
		n-forms	6973	99.97	
s 'with'	11311.9 (8)	j-forms	12	0.04	
		n-forms	29561	99.96	
o(b) 'against'	3407.1 (31)	j-forms (sent-f)[83]	0	0	ACC
	714.5 (129)	n-forms (sent-f)	17	100	ACC
pod 'under' (directional)	1126.0 (81)	j-forms (sent-f)	0	0	
		n-forms (sent-f)	165	100	

[81] Based on the Main, Newspaper and Speech Subcorpora of the Russian National Corpus and dated 1980-2015.

[82] Based on the Lyashevskaya and Sharoff (2009) frequency dictionary (ipm = instances per million words).

[83] 'Sent-f' marks cases, where only sentence-final occurrences of pronouns were counted (due to the complicating ambiguity of j-forms in the genitive case).

pro 'about' (1316-actual)	392.1 (248)	j-forms (sent-f)	0	0	ACC
		n-forms (sent-f)	278	100	
za ‘behind (DIR)’, 'for', 'instead'	3904.1 (24)	j-forms (sent-f)	1	0.07	
		n-forms (sent-f)	1458	99.93	
po 'for'	5786.7 (11)	j-forms (sent-f)	0	0	
		n-forms (sent-f)	0	0	
s ‘off/of’	11311.9 (8)	j-forms (sent-f)	0	0	GENITIVE
		n-forms (sent-f)	**3359**	100	
u 'at'	4306.1 (21)	j-forms (sent-f)	46	1	
		n-forms (sent-f)	**4550**	99	
iz 'from, of'	4314.1 (20)	j-forms (sent-f)	1	0.07	
		n-forms (sent-f)	**1391**	99.93	
ot 'from, of'	3672.5 (29)	j-forms (sent-f)	5	0.15	
		n-forms (sent-f)	**3420**	99.85	
do 'till, up to, before'	2061.1 (51)	j-forms (sent-f)	2	0.21	
		n-forms (sent-f)	**934**	99.79	
bez 'without'	1018.8 (88)	j-forms (sent-f)	3	0.19	
		n-forms (sent-f)	**1578**	99.81	
iz-za 'from behind; due to'	289.6 (353)	j-forms (sent-f)	0	0	
		n-forms (sent-f)	235	100	
iz-pod 'from under'	86.8 (1424)	j-forms	0	0	
		n-forms	123	100	
po-nad ‘above (PATH)’ (Ukrainiasm)	0.6 (41085)	j-forms	0	0	INS
		n-forms	0	0	
po-za ‘behind (PATH)’[84]	0.4 (48208)	j-forms	0	0	
		n-forms	3	100	

[84] All three instances come from the work of the same author (the writer V. Astafiev).

Root prepositions					
skvoz' ‘through’	110.5 (1095)	j-forms	0	0	ACC
		n-forms	134	100	
čerez 'across, through'	805.4 (115)	j-forms (sent-final)	0	0	
		n-forms (sent-final)	365	100	
protiv 'against, opposite'	236.5 (451)	j-forms (sent-final)	0	0	GEN
		n-forms (sent-final)	**581**	100	
bliz ‘near, beside’	10.6 (7846)	j-forms	0	0	
		n-forms	31	100	
sred' 'among' [archaic/dialectal]	4.8 (13606)	j-forms	1		
		n-forms	13		
mež 'between'	22.2 (4598)	j-forms (plural)	1	2.56[85]	
		n-forms (plural)	38	97.44	
		j-forms (plural)	0	0	INS
		n-forms (plural)	112	100	
Suffixed prepositions					
meždu ‘between’	607.5 (149)	j-forms (plural)	5	0.1	INS
		n-forms (plural)	5235	99.9	
		j-forms (plural)	0	0	GEN
		n-forms (plural)	11	100	
dlja 'for'	3229.3 (32)	j-forms (sent-final)	3	0.12	
		n-forms (sent-final)	**2492**	99.88	
radi ‘for the sake of’	109.2 (1109)	j-forms	3	0.96	
		n-forms	308	99.04	
krome ‘except for’	343.6 (286)	j-forms	3	0.39	
		n-forms	763	99.61	

[85] The slightly higher percentage of *j*-forms for this preposition might be due to its archaic status and stylistic considerations.

mimo ‘past’	75.9 (1610)	j-forms	0	0	
		n-forms	289	100	
sredi ‘among’	314.9 (319)	j-forms (plural)	2	0.02	
		n-forms (plural)	**11361**	99.98	
Prefixed prepositions					
vozle (Adv)’beside’	111 (1088)	j-forms	0	0	GENITIVE
		n-forms	310	100	
podle (Adv) ‘beside’	5.8 (11864)	j-forms	0	0	
		n-forms	72	100	
posle 'after'	1080.1 (85)	j-forms (sent-f)	2	0.59	
		n-forms (sent-f)	336	99.41	
okolo (Adv) ‘near, by, next to’	228.1 (473)	j-forms	0	0	
		n-forms	186	100	
vokrug ‘around’	154.4 (755)	j-forms	1	0.07	
		n-forms	**1376**	99.93	
*vkrug ‘around’ [archaic]	0.9 (33607)	j-forms	0	0	
		n-forms	18	100	
vmesto ‘instead of’	155.7 (742)	j-forms	2	0.19	
		n-forms	**1071**	99.81	
vrode ‘like’	78.5 (1579)	j-forms	0	0	
		n-forms	22	100	
		n-forms	13	100	
Class II: Allowing n-forms (j-forms > 1%)					
Strongly favoring n-forms (1% < j-forms </= 10%)					
Prefixed prepositions					
vdol' ‘along’	66.7 (1804)	j-forms	2	1.16	GEN
		n-forms	170	98.84	
pomimo “besides”	64.6 (1866)	j-forms	3	1.89	
		n-forms	156	98.11	
nasčët “concerning”	59.9 (1996)	j-forms	1	2.44	
		n-forms	40	97.56	
naprotiv “against, facing”	23.2 (4416)	j-forms	7	3.61	
		n-forms	187	96.39	
szadi “behind”	3.0 (18436)	j-forms	3	4.23	
		n-forms	68	95.77	
pozadi “behind”	12.7 (6937)	j-forms	10	5.62	
		n-forms	168	94.38	
poverx “over”	20.1 (4934)	j-forms	8	5.71	
		n-forms	132	94.29	
posredi “amidst, in the middle of”	32.7 (3341)	j-forms	3	6.38	
		n-forms	44	93.62	

Suffixed prepositions					
krugom “around”	1.6 (25832)	j-forms	1	7.14	GEN
		n-forms	13	92.86	
No strong preference for either form (10% < j-forms < 90%)					
Adverbial prepositions					
otnositelno ’in relation to, regarding’	48.6 (2409)	j-forms	10	18.18	GEN
		n-forms	45	81.82	
Suffixed prepositions					
posredstvom ‘by means of’	21.3 (4729)	j-forms	8	72.73	GEN
		n-forms	3	27.27	
Prefixed prepositions					
vblizi “in proximity to”	12.9 (6838)	j-forms	7	10.94	GEN
		n-forms	57	89.06	
vperedi “in front of”	13.3 (6703)	j-forms	19	15.2	
		n-forms	106	84.8	
sverx “above, over”	9.7 (8371)	j-forms[86]	2	16.67	
		n-forms	10	83.33	
promež "between" (including INSTR)	1.9 (23907)	j-forms	2	18.18	
		n-forms	9	81.82	
iznutri "from inside"	1.1 (31045) - prep[87]	j-forms	2	18.18	
	23.8 (4329) - ADV	n-forms	9	81.82	
vnutri “inside”	63.4 (1902)	j-forms	56	19.24	
		n-forms	235	80.76	
vnutr' ‘into, inside’	9.8 (8284)	j-forms	11	29.73	
		n-forms	26	70.27	
suprotiv "against" [archaic]	0.8 (36801)	j-forms	2	25	
		n-forms	6	75	
poperëk ‘across’	9.7 (8363)	j-forms	5	31.25	
		n-forms	11	68.75	
vsledstvie 'as a result of'	24.0 (4303)	j-forms	2	40	
		n-forms	3	60	
vperëd 'to the front of'	1.6 (25642) - prep	j-forms	3	42.86	
	128.1 (936) - ADV	n-forms	4	57.14	
vzamen ‘in exchange for’	5.6 (12079)	j-forms	31	65.96	
		n-forms	16	34.04	
naverxu 'on top of'	21.2 (4742) - ADV	j-forms	2	66.67	
		n-forms	1	33.33	

[86] *Promež* "between" licenses both GEN and INS complements; both of them are counted here, since the preposition is very rare (because of its archaic status, apparently).

[87] Lyashevskaya and Sharoff (2009) treat certain lexemes as both prepositions (when transitive) and adverbs (when intransitive), hence two frequency values appear for some of the prepositions in this table.

vglub' ‘deep into’	0.6 (39579)	j-forms	7	70	
		n-forms	3	30	
posredi**ne** ‘in the middle/midst of’	2.3 (21494)	j-forms	5	83.33	
		n-forms	1	16.67	
poseredi**ne** ‘in the middle of’	5.0 (13186)	j-forms	13	86.67	
		n-forms	2	13.33	
nakanu**ne** ‘on the eve of’	29.4 (3672)	j-forms	8	88.89	
		n-forms	1	11.11	
Strongly disfavoring n-forms (95% < j-forms < 100%)					
Suffixed prepositions					
v**ne** (Adv) ‘out of’	59.4 (2008)	j-forms	176	96.17	GEN
		n-forms	7	3.83	
Adverbial prepositions					
soglasno ‘in accordance with’	91.9 (1335)	j-forms	746	97.26	DAT
		n-forms	21	2.74	
Gerund prepositions					
blagodarja ‘thanks to’	94.5 (1293)	j-forms	2000	98.72	DAT
		n-forms	26	1.28	
Class III: Blocking n-forms (j-forms = 100%)					
Gerund prepositions					
(ne) projdja '(not) having passed'	No data	j-forms	16	100	ACC/GEN
		n-forms	0	0	
ne schitaja "excluding [not counting]"	No data	j-forms	2	100	GEN
		n-forms	0	0	
ne doezzhaja "not reaching"	No data	j-forms	1	100	
		n-forms	0	0	
vkljuchaja ‘including’	58.6 (2033)	j-forms	26	100	ACC
		n-forms	0	0	
iskljuchaja ‘excluding’	3.0 (18209)	j-forms	1	100	ACC
		n-forms	0	0	
Adverbial prepositions					
podobno ‘like, similarly to’	29.9 (3626)	j-forms	35	100	DAT
		n-forms	0	0	
soobrazno ‘in conformity with’	1.5 (27139)	j-forms	6	100	
		n-forms	0	0	
sorazmerno 'in proportion to'	No data	j-forms	2	100	
		n-forms	0	0	
sootvetstvenno ‘according to’	1.2 (30477)	j-forms	13	100	
		n-forms	0	0	

Prefixed prepositions[88]					
v(o)sled ‘following, after’	5.7 (11922) – prep.	j-forms	106	100	DAT
	1.4 (27392) - ADV	n-forms	0	0	
vdogonku 'after, in pursuit of'	5.0 (13080) - ADV	j-forms	16	100	
		n-forms	0	0	
vopreki ‘despite (of)’	26.0 (4056)	j-forms	70	100	
		n-forms	0	0	
vzamen "in stead of/in replacement of"	5.6 (12079)	j-forms	3	100	
		n-forms	0	0	
navstreču ‘towards’	56.8 (2091) - ADV	j-forms	325	100	
		n-forms	0	0	
naperekor ‘athwart, in defiance’	2.5 (20318)	j-forms	9	100	
		n-forms	0	0	
napererez ‘cutting (way)’	No data	j-forms	15	100	
		n-forms	0	0	
nazlo 'to spite'	9.1 (8731) - ADV	j-forms	37	100	
		n-forms	0	0	
srodni 'akin to'	3.2 (17683)	j-forms	4	100	
		n-forms	0	0	
podelom 'rightly'	No data	j-forms	32	100	
		n-forms	0	0	
Combination (P+N) prepositions					
na radost' 'to please smb'	No data	j-forms	2	100	DAT
		n-forms	0	0	
na smenu 'to change smb/smth'	No data	j-forms	354	100	
		n-forms	0	0	
v protivoves 'as a counterbalance'	No data	j-forms	33	100	
		n-forms	0	0	
v protivopolozhnost' 'in contrast'	No data	j-forms	16	100	
		n-forms	0	0	
v piku 'to spite'	No data	j-forms	15	100	
		n-forms	0	0	
v otmestku 'in revenge'	No data	j-forms	3	100	
		n-forms	0	0	
ne v primer 'unlike'	No data	j-forms	9	100	
		n-forms	0	0	

[88] Lexemes in the shaded cells can be ‘prepositions’ as well as ‘postpositions’ and can be stranded.

A2. Questionnaire data for individual prepositions across 52 speakers

Preposition *krome 'except'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.90 (1.07)	No data	1.89 (1.10)	1.77 (1.06)	1.42 (0.80)
& [+n]	3.12 (0.94)	3.14 (1.14)	3.35 (0.95)	1.60 (0.98)	3.48 (0.85)
	*p<<0.001	---	*p<<0.001	p=0.39, 2-tail	*p<<0.001

Preposition *pomimo 'besides'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	2.08 (1.22)	2.60 (1.27)	1.96 (1.17)	1.75 (0.99)	1.67 (0.9)
& [+n]	2.83 (1.06)	3.12 (1.15)	3.10 (0.1)	1.62 (1.03)	3.64 (0.69)
	*P=0.0006, 1-tail	*p=0.016, 1-tail	*p<<0.001	p=0.5, 2-tail	*p<<0.001

Preposition *dlja 'for'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.44 (0.75)	1.58 (0.87)	1.87 (1.05)	1.25 (0.74)	1.02 (0.14)
& [+n]	3.04 (1.01)	2.94 (1.15)	No data	1.37 (0.84)	3.77 (0.68)
	*p<<0.001	*p<<0.001	--	p=0.46, 2-tail	*p<<0.001

Preposition *radi 'for the sake of'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.89 (1.08)	1.94 (1.20)	1.85 (1.11)	1.40 (0.93)	1.12 (0.47)
& [+n]	3.04 (1.03)	3.19 (1.07)	3.25 (1.06)	1.42 (0.87)	3.89 (0.47)
	*p<<0.001	*p<<0.001	*p<<0.001	p=0.91, 2-tail	*p<<0.001

Preposition *pered 'in front of'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.92 (1.05)	1.98 (1.13)	1.89 (1.13)	1.15 (0.46)	1.17 (0.38)
& [+n]	2.35 (1.01)	2.29 (1.14)	2.35 (1.08)	1.15 (0.61)	3.69 (0.70)
	*p=0.02, 1-tail	p=0.09, 1-tail	*p=0.02, 1-tail	p=1, 2-tail	p<<0.001

Preposition *vperedi 'in front of'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.64 (0.91)	2.21 (1.23)	1.87 (1.10)	1.46 (0.67)	1.44 (0.78)
& [+n]	2.54 (1.18)	2.96 (0.99)	3.08 (1.03)	No data	3.83 (0.59)
	*p<<0.001	*p=0.0004, 1-tail	*p<<0.001	--	*p<<0.001

Preposition *za 'behind'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	2 (1.19)	2 (1.09)	1.71 (1.02)	1.10 (0.41)	1 (0)
& [+n]	2.46 (1.21)	2.5 (1.11)	2.08 (1.01)	1.23 (0.61)	3.77 (0.55)
	*p=0.03, 1-tail	*p=0.01, 1-tail	*p=0.03, 1-tail	p=0.19, 2-tail	*p<<0.001

Preposition *pozadi 'behind'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.81 (1.03)	2.15 (1.23)	1.89 (1.04)	1.64 (0.99)	1.37 (0.69)
& [+n]	2.83 (1.06)	2.81 (1.09)	2.98 (1.18)	1.54 (0.1)	3.83 (0.51)
	*p<<0.001	*p=0.002	*p<<0.001	p=0.62, 2-tail	*p<<0.001

Preposition *među 'between'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	2.17 (1.18)	2.04 (1.05)	1.71 (1.05)	1.56 (1.06)	1.31 (0.9)
& [+n]	2.77 (1.00)	2.83 (1.12)	2.35 (1.08)	1.25 (0.56)	3.96 (0.19)
	*p=0.003, 1-tail	*p=0.0002, 1-tail	*p=0.002	*p=0.03, 1-tail	*p<<0.001

Preposition *naprotiv 'opposite'*

Second conjunct	*First conjunct*				*Single*
	DP	**[-3]**	**[+n]**	**[-n]**	
& [-n]	1.77 (1.11)	2.17 (1.22)	1.98 (1.06)	No data	1.35 (0.79)
& [+n]	3 (1.05)	3.39 (0.99)	2.87 (1.09)	1.40 (0.75)	3.75 (0.59)
	*p<<0.001	*p<<0.001	*p<<0.001	--	*p<<0.001

A3. Auxiliary questionnaire probing the status of the 10 examined prepositions with a single 3rd person pronominal complement.

Since any corpus comprises data from a variety of speakers, the result may be quite messy. In particular, we do not know whether the optionality of Class II prepositions with respect to the pronominal form that they take is a result of their being such in individual grammars or rather a consequence of idiolectal variation with respect to which form a given P takes. To resolve the issue, I prepared a short questionnaire involving 10 Class I-Class II prepositions.

The questionnaire contained 20 target sentences with 2 sentences for each P. Each pair of sentences was a minimal pair – only the form of the pronoun was manipulated. The linear structure of the sentences was as follows:

(1) DP V (NP) [P 3rd pron]

(2) Vasja pišet stixi dlja {neë; eë}
Vasya writespoems for she.GEN**[+n]** she.GEN**[-n]**
'Vasya writes poems for her.'

(3) Nastja neslas' za {nim; im}
Nastja rushed behind he.INS**[+n]** he.INS**[-n]**
'Nastya rushed behind/after him.'

Methods and participants

The sentences were organized into a questionnaire using the eSurv on-line survey platform.[89] The sentences were distributed over 4 pages, containing 5 sentences each; the order of sentences on each page was automatically randomized by the platform for each participant, but the order of pages was the same for each speaker (the program does not allow for randomization of pages). The decision to present sentences on different pages was motivated by two factors: 1) it is easier for the participants to see fewer sentences on a page; 2) preventing certain sentences to appear close to each other. To elaborate on the latter factor, the aim was to have the minimal pairs maximally distant from each other and not to have sentences with the paired prepositions on the same page. Thus, odd pages of the questionnaire contained

[89] http://esurv.org/

prepositions *krome* ‘except’*; radi* ‘for the sake of’*; za* ‘behind’*; vperedi* ‘in front/ahead of’*; meždu* ‘between’ and even pages comprised sentences with *pomimo* ‘besides’; *dlja* ‘for’ *pozadi* ‘behind’ *pered* ’in front of’ *naprotiv* ‘opposite’. Finally, each page always contained two sentences with n-forms in the P-complement position and three sentences with n-less forms in this position or vice versa, for balance.

The questionnaire was prefaced with detailed instructions (full text to be given in the Appendix) as well as six personal information fields asking for the participant’s:

1) Name/nick name[90]
2) Gender/Age
3) Native language
4) Village, town or region where s/he spent the first 12 years of life
5) Professional sphere (the relevant information being whether a person is a linguist or otherwise)

In the instruction text the informants were asked to evaluate a number of Russian sentences from the point of view of their acceptability in spoken Russian by choosing one of the four answer options: 1) absolutely unacceptable; 2) possible but sounds bad; 3) OK (‘normal’) but sounds weird; 4) good. These options were presented in a line below the sentence, from left to right.

The participants were instructed to base their judgments solely on their own linguistic intuition and not to consult with anyone. In addition, it was recommended that they read the sentences out loud (as this could help to complete the task faster) and that they do not ponder over sentences, relying instead on their initial reaction and choosing the most fitting answer option. The participants were given the option to save their progress and get back to the questionnaire later – the preface instructions explained how to do this.

[90] This was done to enable comparison of speakers’ judgments in two questionnaires.

Participants

86 subjects completed the questionnaire. However, four of them indicated Tatar, Azerbaijani or a Belarussian dialect as their native language and were excluded from the analysis. Thus, only the data obtained from the remaining 82 speakers will be considered below.

The sociological profile of the group is as follows: 34 male and 48 female subjects aged 19-70 (mean = 31, SD = 14.12; median = 25); 24 linguists and 58 non-linguists.

Almost a half (n=39, 48%) of the group spent their childhood and early adolescence years in the Moscow region of Russia/USSR; 4 speakers grew up in the former Soviet Republics (Belarus, Latvia or Georgia); 36 participants were raised in other regions of Russia/Russian part of the USSR and 3 speakers in an unspecified region of Russia/USSR.

To enable a quantitative analysis of the data as well as its more convenient presentation, I converted the verbal judgments into numbers, with the lowest judgment corresponding to '1' and the highest to '4'.

Results

A major overall result of the survey is that *n-less forms* appear to be acceptable with more prepositions than expected from the initial, corpus-based classification. Table below shows that if we take the number of speakers who accepted a particular sentence with an n-form and with an n-less form to roughly correspond to the number of occurrences of these two forms with a particular preposition in the corpus and apply the same class thresholds, all ten prepositions will have to be classified as Class II. However, if we take into account the mean acceptability of n-less forms across speakers, we will see that for none of the considered prepositions does it reach the level of marginal acceptability ('2'), with *pomimo* 'besides' being the only one coming close to it. Especially notable is the difference between the mean acceptability of n-less forms and that of n-forms: it varies between 2.02 (for *pomimo* 'besides') and 2.83 (for *dlja* 'for') with n-forms judged between 3.56-3.98 (that is, clearly acceptable). Setting the issue of precise classification aside, I point out that the simple preposition *za* and the suffixed preposition *dlja* behave differently from the simple *pered*, suffixed *meždu* and the remaining prepositions down the table (the difference

between the means for *za* and *pered* is statistically significant ($p<0.05$, 2-tail t-test), with the suffixed *radi* being in between the two groups.

Acceptability vs. corpus frequency of n-less forms

Preposition	**Acceptability data**				**Corpus data**	
	Mean (SD)	No. (%) speakers	% n-less/all	Class	% n-less/all	Class
dlja	1.02 (0.16)	2 (2.44%)	2.41	Class II	0.12	Class I
za	1.04 (0.25)	2 (2.44%)	2.41	Class II	0.03	Class I
radi	1.15 (0.47)	9 (10.98%)	10.11	Class II	0.96	Class I
pered	1.24 (0.56)	16 (19.51%)	17.20	Class II	0	Class I
mezhdu	1.33 (0.90)	11 (13.41%)	11.83	Class II	0.1	Class I
naprotiv	1.33 (0.72)	17 (20.73%)	17.17	Class II	3.61	Class II
pozadi	1.40 (0.77)	22 (26.83%)	21.15	Class II	5.62	Class II
vperedi	1.40 (0.72)	24 (29.27%)	22.86	Class II	15.2	Class II
krome	1.48 (0.86)	24 (29.27%)	23.08	Class II	0.39	Class I
pomimo	1.70 (0.88)	37 (45.12%)	31.36	Class II	1.89	Class II

At the level of individual speakers, the following patterns of responses were distinguished:

0. Only n-forms allowed with all Ps (n=31, 37.8%)
1. Both forms allowed for 1 P (n=16, 19.5%), subgroups:
2. True optionality for *mežđu* ‘between’ (n=3)
3. Optionality for *pomimo* ‘besides’ (n=7)
4. Optionality for *vperedi* ‘in front of’ (n=3)
5. Optionality for *krome* ‘except’ (n=2)
6. Optionality for *pered* ‘in front of’ (n=2)
7. Both forms allowed for 2 Ps (n=8), subgroups:
8. Optionality for *pozadi* ‘behind’ & *pomimo* ‘besides’(n=3)
9. Optionality for other 2P combinations - inhomogeneous (n=6)
10. Both forms allowed for 3 Ps (n=7)

11. Both forms allowed for 4 Ps (n=2)
12. Both forms allowed for 5 Ps (n=6)
13. Both forms allowed for 6 Ps (n=7)
14. Both forms allowed for 7 Ps (n=2; 6/7 Ps overlap)
15. Both forms allowed for all 10 Ps (n=1)

Regularly, in case both the n-form and the n-less form variant were accepted by a speaker, the former received a higher or identical judgment than the latter. Only 3 speakers from groups 2, 3 and 7 respectively preferred n-less forms over n-forms for 1-2 Ps, namely *naprotiv 2:4*, *vperedi 2:3*, *krome 2:4* and *pomimo 2:3*.

Preposition	**Only n-form**	**n-form/n-less form**	**Only n-less form**	**Unacceptable**
Za ‘behind’	79 speakers	2 speakers	-	1 speaker
Dlja ‘for’	79 speakers	2 speakers	-	1 speaker
Radi ‘for the sake of’	71 speakers	9 speakers	-	2 speakers
Mezhdu ‘between’	71 speakers	11 speakers	-	-
Pered ‘in front of’	62 speakers	15 speakers	1 speaker	4 speakers
Naprotiv ‘opposite’	65 speakers	17 speakers	-	-
Krome ‘except’	58 speakers	22 speakers	2 speakers	-
Pozadi ‘behind’	60 speakers	22 speakers	-	-
Vperedi ‘in front of’	57 speakers	24 speakers	-	1 speaker
Pomimo ‘besides’	45 speakers	36 speakers	1 speaker	-

We can see from the table above that each of the 10 prepositions (belonging to Class I-II, based on the corpus) is considered to only allow n-forms by a majority of speakers (55-96%). The number of speakers admitting both forms, with the preference for n-forms, ranges from 2.5% to 44%. Finally, three prepositions are only accepted with n-less forms by 4 speakers overall: *pered* and *pomimo* – by one speaker each and *krome* – by two speakers. Additionally, 7 speakers did not accept a particular sentence

with either n- or n-less form (*za, dlja, radi, pered, vperedi*); for another speaker there were two sentences, not accepted with either form (*radi* and *pered*).

Under the assumption that acceptability judgments may serve as a model for actual performance, we may say that prepositions classified as Class II based on the corpus, that is, those occurring with both n-forms and n-less forms, might have distinct status in individual grammars: some grammars only admit n-forms, some admit both and some only admit n-less forms. Crucially though, there does seem to be optionality in individual grammars.

9. Bibliography

Abels, Klaus. (2003). *Successive cyclicity, Anti-locality, and Adposition Stranding.* Ph.D. dissertation, University of Connecticut, Storrs.

Aoun, Joseph, Hornstein, Norbert, Lightfoot, David, & Weinberg, Amy. (1987). Two types of locality. *Linguistic Inquiry, 18*, 537-577.

Assmann, Anke, Edygarova, Svetlana, Georgi, Doreen, Klein, Timo, & Weisser, Philipp. (2014). Case stacking below the surface: On the possessor case alternation in Udmurt. *The Linguistic Review, 31*(3-4), 447-485.

Babby, Leonard H. (1978). *Negation and subject case selection in existential sentences*. Bloomington: Indiana University Linguistics Club.

Babby, Leonard H. (1986). The locus of case assignment and the direction of percolation: Case theory and Russian. In R. D. Brecht & J. S. Levine (Eds.), *Case in Slavic* (pp. 170-219). Columbus, Ohio: Slavica Publishers.

Bailyn, John F. (2004). *The Case of Q*. Paper presented at the Formal Approaches to Slavic Linguistics 12: The Ottawa Meeting, Ann Arbor, MI.

Bailyn, John F. (2012). *The Syntax of Russian*. New York: Cambridge University Press.

Baker, Mark C. (2003). *Lexical Categories: Verbs, Nouns and Adjectives*. Cambridge: Cambridge University Press.

Béjar, Susana, & Massam, Diane. (1999). Multiple Case Checking. *Syntax, 2*(2), 65-79.

Berezovskaya, Polina. (2013). Acquisition of Russian comparison constructions. Semantics meets first language acquisition In M. Kohlberger, K. Bellamy & E. Dutton (Eds.), *Proceedings of ConSOLE XXI (8-10 January 2013, Potsdam)* (pp. 45-65). Leiden: Leiden University Centre for Linguistics.

Berezovskaya, Polina, & Hohaus, Vera. (2015). The Crosslinguistic Inventory of Phrasal Comparative Operators: Evidence from Russian *Proceedings of FASL 23*.

Bhatt, Rajesh, & Takahashi, Shoichi. (2007). Direct comparisons: Resurrecting the direct analysis of phrasal comparatives. In M. Gibson & T. Friedman (Eds.), *Proceedings of SALT XVII* (pp. 19-36). Ithaca: Cornell University, CLC Publications.

Bhatt, Rajesh, & Takahashi, Shoichi. (2011). Reduced and unreduced phrasal comparatives. *Natural Language & Linguistic Theory, 29*(3), 581-620.

Blažev, B. I. (1962). K voprosu ob upotreblenii ličnogo mestoimenija 3-go lica s načal'nym n-. [On the use of n-initial 3rd person pronouns]. *Russkij jazyk v škole, 2*, 13-16.

Botwinik-Rotem, Irena. (2004). *The Category P: Features, Projections, Interpretation.* Ph.D. thesis, Tel Aviv University.

Bresnan, Joan W. (1973). Syntax of the Comparative Clause Construction in English. *Linguistic Inquiry, 4*(3), 275-343.

Bulygina, T.V. (1977). *Problemy teorii morfologičeskix modelej [Issues in the theory of morphological models]* Moscow: Nauka.

Bylinskij K.I, & Rozental', D.E. (2011). *Literaturnoe redaktirovanie: Učebnoe posobie [Copy-editing: A reference book]*. Moscow: Flinta/Nauka.

Campbell, Lyle. (2013). *Historical Linguistics: An Introduction* (Third ed.). Cambridge, MA: MIT Press.

Cardinaletti, Anna, & Starke, Michal. (1999). The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns. In H. van Riemsdijk (Ed.), *Clitics in the Languages of Europe* (pp. 145-234). Berlin and New York: Mouton de Gruyter.

Chomsky, Noam. (1986). *Knowledge of language*. New York: Praeger.

Chomsky, Noam. (1995). Bare Phrase Structure. In G. Webelhuth (Ed.), *Government and Binding Theory and the Minimalist Program* (pp. 383-439). Cambridge: Blackwell.

Cinque, Guglielmo, & Rizzi, Luigi (Eds.). (2010). *Mapping spatial PPs: the cartography of syntactic structures volume 6*. Oxford: Oxford University Press.

Comrie, Bernard, & Corbett, Greville G. (Eds.). (1993). *The Slavonic Languages*. London/New York: Routledge.

Daniel, M.A. (2015). Stem initial alternation in Russian third person pronouns: variation in grammar. In V. P. Selegei. e. al. (Ed.), *Computational Linguistics and Intellectual Technologies: Papers from the 14th Annual International Conference "Dialogue" (2015)* (Vol. 1, pp. 95-104). Moscow: RSUH Press.

Déchaine, Rose-Marie, & Wiltschko, Martina. (2002). Decomposing pronouns. *Linguistic Inquiry, 33*(3), 409-442.

Grashchenkov, Pavel, & Lyutikova, Ekaterina. (2017). O sintaksise komparativa v russkom jazyke [On the syntax of comparatives in Russian]. *Russkij jazyk v naučnom osveščenii, 33*(1), 116-141.

Gribanova, Vera. (2009). Phonological Evidence For a Distinction between Russian Prepositions and Prefixes. In G. Zybatow, U. Junghanns, D. Lenertová & P. Biskup (Eds.), *Proceedings of the 7th meeting of Formal Desription of Slavic Languages (FDSL 7), Leipzig 2007* (pp. 383-396). Frankfurt am Main: Peter Lang.

Grimshaw, Jane. (2000). Locality and Extended Projection. In P. Coopmans, M. Everaert & J. Grimshaw (Eds.), *Lexical Specification and Insertion* (pp. 115-133). Amsterdam: John Benjamins.

Hagège, Claude. (2010). *Adpositions*. New York: Oxford University Press.

Halle, Morris, & Marantz, Alec. (1993). Distributed Morphology and the Pieces of Inflection. In K. Hale & S. J. Keyser (Eds.), *View from Building 20* (pp. 111-176). Cambridge, Mass.: MIT Press.

Hankamer, Jorge. (1973). Why there are two than's in English. *Papers from the Ninth Regional Meeting of the Chicago Linguistic Society* (pp. 179-188). Chicago.

Hill, Steven P. (1977). *The N-factor and Russian prepositions. Their development in 11th-20th century texts*. The Hague/Paris/New York: Mouton.

Hoeksema, Jack. (1983). Negative Polarity and the Comparative. *Natural Language & Linguistic Theory, 1,* 403-434.

Ionin, Tania, & Matushansky, Ora. (2012). *More than one comparative in more than one Slavic language: an experimental investigation*. Paper presented at the Proceedings of Formal Approaches to Slavic Linguistics (FASL) 21.

Itkin, I. B., & Tolkacheva, V. A. (2015). *Postpositivnye upotreblenija predloga "radi" v russkom literaturnom jazyke XVIII-XXI vv. [Postpositional use of preposition "radi" in the Russian literary language of the XVIII-XXI c.]*. Paper presented at the 4th Conference on Constructional & Lexical Semantic Approaches to Russian.

Irimia, Monica Alexandrina. (2016). How small are small clauses? Embedded adjectives and restructuring. In K.-m. Kim, P. Umbal, T. Block, Q. Chan, T. Cheng, K. Finney, M. Katz, S. Nickel-Thompson & L. Shorten (Eds.), *Proceedings of the 33rd West Coast Conference on Formal Linguistics* (pp. 207-216). Somerville, MA: Cascadilla Proceedings Project.

Jakobson, Roman. (1971 [1935]). Les enclitiques slaves. *Selected Writings, vol. II: Word and Language* (pp. 16-22). The Hague: Mouton &. Co. [First published in

Atti del Congresso di Linguistica tenuto in Roma il 19-26 settembre 1933, Firenze, 1935.].

Kennedy, Christopher. (1999). *Projecting the Adjective: The Syntax and Semantics of Gradability and Comparison*. New York: Garland Press.

Kholodilova, M. A. (2013). *Pozicionnye svojstva mestoimenij v russkom jazyke [Clausal positions of pronouns in Russian]* MA thesis. Saint Petersburg State University.

Koopman, Hilda. (2000). Prepositions, postpositions, circumpositions, and particles. In H. Koopman (Ed.), *The Syntax of Specifiers and Heads* (pp. 204-260). London: Routledge.

Landau, Idan. (2010). *The Locative Syntax of Experiencers*. Cambridge, MA: MIT Press.

Landau, Idan. (2013). *Control in Generative Grammar: A Research Companion*. Cambridge, UK: Cambridge University Press.

Lechner, Winfried. (2001). Reduced and Phrasal Comparatives. *Natural Language & Linguistic Theory, 19*(4), 683-735.

Lechner, Winfried. (2004). *Ellipsis in comparatives*. Berlin: Mouton de Gruyter.

Lechner, Winfried. (to appear). Clausal vs. phrasal comparatives (v. 1.0, February 2015). In L. Matthewson, C. Meier, H. Rullmann & T. E. Zimmermann (Eds.), *The Blackwell Companion to Semantics*. Malden: Blackwell.

Lyashevskaya, Olga, & Sharoff, Serge. (2009). *Častotnyj slovar' sovremennogo russkogo jazyka (na materialax Nacional'nogo korpusa russkogo jazyka) [Frequency dictionary of Contemporary Russian [based on data from the Russian National Corpus)]*. Moscow: Azbukovnik.

Matushansky, Ora. (2002). More of a good thing: Russian synthetic and analytic comparatives *Proceedings of FASL: The Second Ann Arbor Meeting* (pp. 143-162): Michigan Slavic Publications.

Mayo, Peter. (1993). Belorussian. In B. Comrie & G. G. Corbett (Eds.), *The Slavonic Languages*. London/New York: Routledge.

Menon, Mythili. (2012). Adverbial Comparatives: Evidence from Malayalam. *Proceedings of Sinn und Bedeutung 16* (pp. 461-474).

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. (2006). Polyvalent case, geometric hierarchies, and split ergativity In J. Bunting, S. Desai, R. Peachey, C. Straughn & Z. Tomkova (Eds.), *Proceedings of the 42nd annual meeting of the Chicago Linguistics Society*. Chicago, IL: Chicago Linguistics Society.

Merchant, Jason. (2009). Phrasal and clausal comparatives in Greek and the abstractness of syntax. *Journal of Greek Linguistics, 9*, 134-164.

Merchant, Jason. (2012). *Two phrasal comparatives in Greek.* Ms., University of Chicago.

Mezhevich, Ilana. (2006). *Featuring Russian Tense: A Feature-Theoretic Account of the Russian Tense System.* Ph.D. dissertation. University of Calgary.

Napoli, Donna Jo, & Nespor, Marina. (1986). Comparative Structures in Italian. *Language, 62*(3), 622-653.

Norde, Muriel. (2000). The grammaticalization of adpositions in the history of Swedish. In G. Þórhallsdóttir (Ed.), *Proceedings of the 10th Conference of Nordic and General Linguistics* (pp. 177-187). Reykjavík: Institute of Linguistics.

Nykiel, Joanna. (2013). Clefts and preposition omission under sluicing. *Lingua, 123*, 74-117.

O'Connor, Ellen. (2013). A Clausal Analysis of Dari and Tajik Phrasal Comparatives. In Y. Fainleib, N. LaCara & Y. Park (Eds.), *Proceedings of the 21st meeting of the North East Linguistic Society (NELS 41)* (Vol. 2, pp. 83-95). Amherst, MA: GLSA Publications.

Pancheva, Roumyana. (2006). Phrasal and Clausal Comparatives in Slavic. In J. Lavine, S. Franks, M. Tasseva-Kurktchieva & H. Filip (Eds.), *Formal Approaches to Slavic Linguistics 14: The Princeton Meeting* (pp. 236-257).

Pancheva, Roumyana. (2010). More students attended FASL than CONSOLE. In W. Browne, A. Cooper, A. Fisher, E. Kesici, N. Predolac & D. Zec (Eds.), *Formal Approaches to Slavic Linguistics 18: The Cornell Meeting, 2009* (pp. 383-400): Michigan Slavic Publications.

Pancheva, Roumyana, & Tomaszewicz, Barbara. (2011). Experimental Evidence for the Syntax of Phrasal Comparatives in Polish. *University of Pennsylvania Working Papers in Linguistics, 17*(1), Article 21.

Pereltsvaig, Asya. (2007). *Copular Sentences in Russian: A Theory of Intra-Clausal Relations*. Dordrecht: Springer.

Pesetsky, David. (1982). *Paths and Categories.* PhD dissertation, MIT.

Pesetsky, David. (2013). *Russian Case Morphology and the Syntactic Categories* Cambridge, MA: The MIT Press.

Philippova, Tatiana. (2014). P-omission under sluicing, [P clitic] and the nature of P-stranding. In M. Kohlberger, K. Bellamy & E. Dutton (Eds.), *ConSOLE XXII: Proceedings of the 22nd Conference of the Student Organization of Linguistics in Europe* (pp. 133-155). Leiden: Leiden University Centre for Linguistics.

Pinkal, Manfred. (1990). On the logical structure of comparatives. In R. Studer (Ed.), *Natural language and logic*. (pp. 146-167). Berlin: Springer.

Podobryaev, Alexander. (2009). "Postposition Stranding" and related phenomena in Russian. In G. Zybatow. e. al. (Ed.), *Studies in Formal Slavic Phonology, Morphology, Syntax, Semantics and Information Structure* (Vol. 21. Linguistik International, pp. 197-208). Frankfurt am Main: Peter Lang.

Potsdam, Eric. (2012). A direct analysis of Malagasy phrasal comparatives. *Proceedings of the 18th Meeting of the Austronesian Formal Linguistics Association.* (pp. 140-155).

Pšexotskaja, E. A. (2012). *Kosvennoe dopolnenie kak subkategorizovannyj i nesubkategorizovannyj aktant (na materiale russkogo jazyka) [Indirect objects as core and non-core arguments in Russian]*. Ph.D. thesis. Lomonosov Moscow State University.

Rappaport, Gilbert C. (1988). On the Relationship between Prosodic and Syntactic Properties of Pronouns in the Slavic languages. In A. M. Schenker (Ed.), *American Contributions to the Tenth International Congress of Slavists. Sofia, September 1988. Linguistics* (pp. 301-327): Slavica Publishers.

Richards, Norvin. (2013). Lardil "Case Stacking" and the Timing of Case Assignment. *Syntax-a Journal of Theoretical Experimental and Interdisciplinary Research, 16*(1), 42-76.

van Riemsdijk, Henk C. (1990). Functional Prepositions. In H. Pinkster & I. Genee (Eds.), *Unity in Diversity: Papers presented to Simon C. Dick on his 50th birthday*. Dordrecht: Foris.

van Riemsdijk, Henk C. (1998). Categorial Feature Magnetism: The Endocentricity and Distribution of Projections. *The Journal of Comparative Germanic Linguistics 2*(1), 1-48.

Rodrigues, Cilene, Nevins, Andrew, & Vicente, Luis. (2009). Cleaving the Interactions between Sluicing and Preposition Stranding. In D. Torck & W. L. Wetzels (Eds.), *Romance Languages and Linguistic Theory 2006* (pp. 175-198). Amsterdam: John Benjamins.

Rose, Sharon. (2000). Rethinking Geminates, Long-Distance Geminates and the OCP. *Linguistic Inquiry, 31*, 85-122.

Šaxmatov, A. A. (1957). *Istoričeskaja morfologija russkogo jazyka [Diachronic morphology of Russian]*. Moscow: Učpedgiz.

Seliščev, A.M. (1951). *Staroslavjanskij jazyk: Vvedinije. Fonetika [The Old Slavic language: Introduction. Phonetics]* (Vol. 1). Moscow: Gosudarstvennoje Učebno-Pedagogičeskoje Izdatel'stvo Ministerstva Prosveščenija RSFSR.

Shlomina, Anna. (2014). *Ellipsis v kosvennom obshchem voprose (na materiale russkogo jazyka). [Ellipsis in embedded polar questions in Russian]*. MA Thesis. Russian State University of Humanities, Moscow.

Sichinava, D. V. (2013). *Mestoimenie. Materialy dlja proekta korpusnogo opisanija russkoj grammatiki [The pronoun. Materials for the "Corpus-based description of Russian grammar" project]* http://rusgram.ru/%D0%9C%D0%B5%D1%81%D1%82%D0%BE%D0%B8%D0%BC%D0%B5%D0%BD%D0%B8%D0%B5#bibl.

Stassen, Leon. (1985). *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

Stassen, Leon. (2013). Comparative Constructions. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. Available online at http://wals.info/chapter/121.

Švedova, N. Ju. (Ed.). (1980). *Russkaja Grammatika* [Russian grammar] Moscow: Akademija Nauk.

Svenonius, Peter. (2003). Limits on P: filling in holes vs. falling in holes. *Nordlyd 31.2*, 431-445.

Valova, E.A. & Slioussar, N.A. (2016). Comparing Corpus and Experimental Approaches: a Study of Syntactic Properties of the Russian enclitic *že* [Sravnenie korpusnogo i eksperimental'nogo metoda na primere issledovanija

sintaksiceskix svojstv enklitiki *že*]. In S. V. P. e. al (Ed.), *Computational Linguistics and Intellectual Technologies: Proceedings of the Annual International Conference "Dialogue" (2016)* (pp. 806-816). Moscow: RSUH Press.

Vondrák, Wenzel. (1928). *Vergleichende slavische Grammatik*, 2nd ed. (Vol. 2). Göttingen: Vandenhoeck und Ruprecht.

Vostokov, Aleksandr X. (1874). *Russkaja grammatika (Izd. 12) [Russian grammar (2nd Ed.)]*. Sankt-Peterburg: Imperatorskaja Akademija Nauk.

Wunderlich, Dieter. (2001). Two comparatives. In I. Kenesei & R. M. Harnish (Eds.), *Perspectives on semantics, pragmatics and discourse. A Festschrift for Ferenc Kiefer* (pp. 75-89). Amsterdam: Benjamins.

Yadroff, Michael, & Franks, Steven. (1999). *The Origin of Prepositions.* Paper presented at the Third European Conference on Formal Desciption of Slavic Languages, Leipzig.

Zaliznjak, A. A. (2002 [1967]). *Russkoe imennoe slovoizmeneniye [Russian nominal inflection]*. Moscow: Jazyki Slavjanskix Kul'tur.

Zevakhina, Natalia A., & Dzhakupova, Svetlana S. (2015). Russian metalinguistic comparative: a functional perspective *Working Papers in Linguistics*. Moscow, Russia: National Research University Higher School of Economics.

Zwarts, J. (1997). Complex prepositions and P-stranding in Dutch. *Linguistics, 35*(6), 1091-1112.

ON-LINE RESOURCES

Russian National Corpus: http://ruscorpora.ru/

Max Vasmer's Russian Etymological Dictionary: https://vasmer.lexicography.online/

eSurv Survey Platform: https://esurv.org/

תקציר

התזה הנוכחית עוסקת בתופעות מורפוסינטקטיות ברוסיות בהן מופיעות מילות יחס. חלק גדול של העבודה הזאת מוקדש לניתוח הדיסטריבוציה של כינויי הגוף מיוחדים ברוסית שמופיעים בדרך כלל לצד מילות יחס. הדיסטריבוציה שלהם שהינה כאוטית לכאורה, לא הוסברה במחקרים הקודמים. תמוהה במיוחד היתה העובדה שמעבר למשלים של מילות יחס, כינויי גוף אלה מופיעים ב (ורק ב) מבנה ההשוואה ברוסית, למרות שבה לא מופיעה שום מילת יחס.

אני מציעה ניתוח אחיד של הדיסטריבוציה של כינויי הגוף האלה ברוסית, אשר שומרת על האינטואיציה שהם כפופות למילות יחס.עם זאת, אני מחדדת את ההכללה באומרו כי כינויי גוף אלה מורשים רק במשלים של-P . זאת אומרת שאני עושה הבחנה מכרעת בין קבוצה של מילות יחס שמוגדרת באופן לא מדויק ובין מילים המוגדרות כ-P במובן התחבירי.

הרעיון הוא שרק מילות היחס שיש להן מעמד של P-Head מרשות את כינויי הגוף המיוחדים במשלימן.

בטקסט העיקרי אני בוחנת קבוצה גדולה של ״מילות יחס״ בפירוט, ומספקת טיעונים שרק חלק מהן הן P-Heads. התעלומה של המבני ההשוואה נפתרה על ידי הצעה כי המבנה הזה כולל P-head שאינו ״ריק״ ושמסוגל להרשות את הכינויי גוף הנ״ל.

בתזה אני מציעה שמילים שקוראים להם ״מילות יחס״ במסורת הדקדוקית למעשה שייכות ל-3 סוגים. מילים בסוג א׳ הינן מילות יחס ״אמיתיות״, כלומר P-Heads. מילים שבסוג ב׳ אינן הומוגניות מבחינה מורפוסינטקטית: הן יכולות להופיע במעמד של P או להיות קומפינציה של מילית יחס ושם עצם. מילים בסוג ג׳ אינן P-Heads מבחינה התחבירית, אלה שיכות לקטגוריות של הפועל, שם התואר או מהוות קומבינציות של של מילית יחס ושם עצם.

ברצוני לציין שמבחינתי, הטעינות שלי ש״מילות יחס״ של סוג ג׳ הינן מילות יחס אמיתיות ושלמילים מסוג ב׳ יש מעמד דואלי הינם מקובלות על ידי בלשנים שעובדים בכל פרדיגמה. אם כך, התזה הזאת תוכל לעניין קהל די רחב.

עם זאת, שמירה על אותו סוג של הניתוח כדי להסביר למה כינויי הגוף הנ״ל מופיעים גם במבני ההשוואה דורשת קבלת הטענה שבתחביר יש מילים ומבנים ״ריקים״. הטענה הזאת מקובלת בפרדיגמה הגנרטיבית אבל לא בהכרח בכל פרדיגמה בלשנית. ייתכן שזה מהווה תמיכה לפרדיגמות פורמליות שכן מאמצות את האבסטרקציות כאלה.

החלק הזה של התזה תורם למחקר בטיפולוגיה של מילות יחס, כינווי גוף, יחסה ותורת היחסה.

החלק השני של התזה מתמקד על התחביר הפנימע של מבני ההשוואה ברוסית. אני מציע נתונים וממצאים חדשים שהינם רלוונטיים לחקר מבני ההשוואה גם בשפות אחרות. המחקר הזה גם תורם לדיון המפורסם לגבי התחביר הפנימי של מבני השוואה.

מילות מפתח: תחביר, מורפולוגיה, רוסית, מילות יחס, כינויי גוף, יחסה, מבני השוואה, השמטה