

Non-Standard Allomorphy in Russian Prefixes: Corpus, Experimental, and Statistical Exploration

—
Anna Endresen

*A dissertation for the degree of Philosophiae Doctor
August 2014*

Non-Standard Allomorphy in Russian prefixes:

Corpus, Experimental, and Statistical Exploration

Anna Endresen

A dissertation submitted for the degree of Philosophiae Doctor
University of Tromsø:
The Arctic University of Norway
Faculty of Humanities, Social Sciences and Education
Department of Language and Linguistics

CLEAR group:
Cognitive Linguistics: Empirical Approaches to Russian



August 2014

CONTENTS

Acknowledgements	ix
List of abbreviations	xi
1. Introduction	1
1.1 What is Non-Standard Allomorphy?	1
1.2 Empirical exploration: Russian prefixes	2
1.3 Statistical modeling	3
1.4 Theoretical background	4
1.5 Structure of the dissertation	7
2. Allomorphy as an overgeneralized assumption: Theoretical construct vs. empirical phenomenon	9
2.1. Introduction	9
2.2. Puzzles of data: challenges to the definition	9
2.2.1. English indefinite article	10
2.2.2. English past participle marker	11
2.2.3. English plural marker	12
2.2.4. English first-person singular pronoun	14
2.2.5. German diminutive suffixes: <i>-chen</i> vs. <i>-lein</i>	15
2.2.6. What do we learn from this data?	16
2.3. Allomorphy: Asymmetry between Form and Meaning	16
2.4. Types of allomorph conditioning	19
2.4.1. Phonological conditioning	19
2.4.2. Morphological conditioning	20
2.4.3. Lexical conditioning	21
2.4.4. Semantic conditioning	21
2.4.5. Historical conditioning	23
2.5. Two additional criteria for allomorphic status	24
2.5.1. Formal phonological similarity	24
2.5.2. Common historical source	27
2.6. Allomorphy: a term with Structuralist “baggage”	27
2.6.1. Descriptivists	28
2.6.1.1. Harris 1942	28
2.6.1.2. Hockett 1947	29
2.6.1.3. Nida 1948	30
2.6.2. After descriptivists	31
2.6.3. Before descriptivists	32
2.6.3.1. Bloomfield 1933: Abandoning issues of semantics as unscientific	32

2.6.3.2.	Saussure 1916: Linguistic sign is a word	34
2.6.3.3.	Baudouin de Courtenay 1881: The “psychological autonomy” of the morpheme	34
2.6.3.4.	Summing up: Three versions of Structuralism and their implications for allomorphy	36
2.7.	Proposal of this dissertation	37
2.7.1.	Data-driven approach: From Data to Theory	37
2.7.2.	Allomorphy as a Radial Category: Standard & Non-Standard exemplars	38
2.7.3.	The role of semantics	41
2.7.4.	Statistical modeling of Non-Standard cases	43
2.8.	Conclusions	43
3.	Standard allomorphy in Russian prefixes: The case of RAZ- ~ RAS- ~ RAZO- ‘apart’	45
3.1.	Introduction	45
3.2.	Regressive voicing assimilation: #C _{voiced-} ~ #C _{voiceless-}	46
3.2.1.	General remarks	46
3.2.2.	Case study 1: RAZ- ~ RAS-	47
3.3.	Vowel / Zero alternation in consonant-final prefixes: #C- ~ #CV-	53
3.3.1.	General remarks	53
3.3.2.	Case study 2: RAZ/S- ~ RAZO-	56
3.4.	Conclusions	61
4.	The prefixes S- and SO- ‘downward, with’	63
4.1.	Introduction	63
4.2.	Historical variants of the prefix S-	64
4.3.	Problematic data and previous accounts	65
4.4.	Data collection	73
4.4.1.	Methodology	73
4.4.2.	S- and SO- in standard and marginal verbs	76
4.5.	Semantic analysis	78
4.5.1.	CENTRIPETAL MOTION VS. DOWNWARD-ABLATIVE MOTION	79
4.5.2.	CENTRIPETAL MOTION VS. CONCOMITANT ACTION	82
4.5.3.	DOWNWARD-ABLATIVE MOTION VS. ABLATIVE MOTION	84
4.5.4.	RESULTATIVE VS. SEMELFACTIVE	85
4.5.5.	Radial Category Profiling	88
4.6.	Distribution of S- and SO- and conditioning of the prefix	92
4.6.1.	Prefix alternation inside paradigm	92
4.6.2.	Prefix variation in minimal pairs	93
4.6.3.	S- and SO- stacked in one verb	97

4.6.4. Prefix variation in poetry	98
4.6.5. Verbs that exclusively attach SO-	98
4.7. Conclusions	100
5. At the crossroads of phonology and semantics: The prefixes O- and OB- 'around'	102
5.1. Introduction: The puzzle of the prefixes O- and OB-	102
5.2. The Split Hypothesis	105
5.2.1. Proponents and their arguments	105
5.2.2. Drawbacks of the Split Hypothesis	109
5.2.3. Predictions of the Split Hypothesis	110
5.3. Corpus study	111
5.3.1. Goal and justification for a new semantic model	111
5.3.2. Data collection	113
5.3.3. The cognitive model of polysemy: Radial network of submeanings and their interconnections	113
5.3.4. Radial category profiles of the prefixes O- and OB-	122
5.3.5. Statistical analysis: Combined Classification Tree & Random Forests Model	126
5.3.6. Patterns of derivation and intermediate cases	131
5.3.7. Hypothesis testing against lexicon: Discussion of results	133
5.4. Experiment	133
5.4.1. Goal, design, subjects, administration	134
5.4.2. Results	141
5.4.2.1. Variation in subjects' individual strategies	143
5.4.3. Statistical analysis	144
5.4.3.1. Linear Regression: Mixed-Effects Model	144
5.4.3.2. Combined Classification Tree & Random Forests Model	147
5.4.4. Experimental results: Discussion	149
5.5. Alternative account: O- and OB- represent a single morpheme with Non-Standard Allomorphy	150
6. Grammatically conditioned allomorphy: The prefixes PERE- 'over, across' and PRE- 'very'	152
6.1. Introduction	152
6.1.1. Diachronic relation of PERE- and PRE-	152
6.1.2. Synchronic relation of PERE- and PRE-	154
6.1.3. On the similarity of formal shapes: PERE-, PRE-, PRED-, PRI-	155
6.1.4. Specificity of Russian among Slavic: Coexistence of PERE- and PRE-	156
6.1.5. State of the art: Previous accounts	156
6.1.6. Two possible solutions and their problems	157

6.1.7. Goal of the present study and overview	158
6.2. Data	159
6.3. Semantic Analysis: Radial network of submeanings	162
6.4. Radial Category Profiling: PERE- vs. PRE- in the domain of verbs	186
6.5. Grammatical properties of PERE- vs. PRE-: Aspect and Transitivity	190
6.5.1. Shift in aspect	191
6.5.2. Shift in transitivity	192
6.5.3. Prefix stacking	193
6.6. Beyond verbs: The role of Slavonic prefix PRE-	195
6.6.1. Two domains of PRE-	196
6.6.2. PRE- as a productive intensifier of quality	196
6.7. Wrapping up the proposal	199
6.7.1. Non-standard grammatically conditioned allomorphy	199
6.7.2. Allomorphy via borrowing	200
6.8. Morphological intensifiers among other types of intensification	201
6.9. Conclusions	203
7. The prefixes VZ- and VOZ- ‘up’	205
7.1. Introduction	205
7.2. Etymological relationship and formal similarity	206
7.3. Distributional overlap and subtle semantic difference	207
7.4. Data	209
7.4.1. Methodology	210
7.4.2. Overview: VZ- and VOZ- in standard and marginal verbs	211
7.5. Semantics of the prefixes VZ- and VOZ-	213
7.5.1. Shared prototype	214
7.5.2. Different altitude	215
7.5.3. Implications for semantics	217
7.5.4. Submeanings motivated by the prototype	217
7.6. Radial Category Profiling	234
7.7. VZ- and VOZ- in minimal pairs	236
7.8. Conclusions	237
8. The prefixes VY- and IZ- ‘out of a container’	239
8.1. Introduction: What is special about this case	239
8.2. Formal similarity and historical relation	240
8.3. Data: Insights from numbers	241
8.4. Near-identical semantics	243
8.4.1. Unwrapping polysemy	243
8.4.2. Profiles within polysemy: Radial Category Profiling	249
8.5. Distribution: Overlap and contrastive uses	252
8.5.1. Verbs that can attach both prefixes	252
8.5.2. Verbs that attach only one prefix	254
8.5.3. Possible explanation: two IZ- prefixes with different	

spatial image schemas	256
8.5.4. A note on prosody and phonology	258
8.6. Discussion	258
8.7. Conclusions	259
9. Rival prefixes in word-formation of Russian change-of-state (factitive) verbs: O- vs. U-	261
9.1. Introduction	261
9.2. Terminology	262
9.3. Factitive verbs in Russian: The world of possibilities	262
9.3.1. Productivity	263
9.3.2. Morphological construction: Affixes involved in derivation	263
9.3.3. Broad variation in prefixes	264
9.3.4. Possible non-verbal bases	265
9.3.5. The Spatial Motivation Hypothesis: What determines the choice of the prefix	267
9.4. The corpus study of O- and U-	269
9.4.1. Data collection	270
9.4.1.1. Challenges of derivational analysis: Multiple motivations	270
9.4.1.2. Distribution of O- and U- across standard and marginal factitive verbs	272
9.4.1.3. Distributional overlap in the use of O- and U- across bases	273
9.4.2. The Scalarity Hypothesis: The difference in the factitive use of O- and U-	275
9.4.2.1. Argument 1: Comparatives employ U-	279
9.4.2.2. Argument 2: Adjectival classes	280
9.4.2.3. Argument 3: Beyond adjectival bases	282
9.4.2.4. Argument 4: Degree modifiers	282
9.4.2.5. Argument 5: Phonology	284
9.4.3. Factitive use of O- and U- and their spatial meanings	285
9.4.4. Conclusions for the corpus study	287
9.5. The experimental study	288
9.5.1. Research questions	288
9.5.1.1. O- vs. U-	288
9.5.1.2. Novel marginal vs. Standard vs. Nonce verbs	289
9.5.1.3. Children vs. Adults	290
9.5.1.4. A note on perspective: Competence & Performance	290
9.5.2. Experimental design: Elicitation of acceptability judgements	291
9.5.2.1. Stimuli	292
9.5.2.2. Administration	294
9.5.2.3. Subjects	295
9.5.2.3.1. Age groups and gender	296

9.5.2.3.2. Place of residence	297
9.5.2.3.3. Profession	297
9.5.3. Experimental results	298
9.5.3.1. A note on the level of measurement appropriate for collected data	298
9.5.3.2. Overview: Central tendencies in data distribution	300
9.5.3.3. Age	302
9.5.3.4. Prefix	305
9.5.3.5. Stimulus type	307
9.5.3.6. All factors in a single model: Advanced statistical modeling	311
9.5.3.6.1. Ordinal Logistic Regression	313
9.5.3.6.2. Regression Mixed-Effects Model for Ordinal Data	315
9.5.3.6.3. Classification and Regression Trees (CART) and Random Forests	316
9.5.3.7. Discussion of experimental results	320
9.6. Conclusions	321
10. Conclusions	322
Appendix 1: Overview of data explored in this dissertation	328
Appendix 2: Expansion of the prefix S- in Modern Russian	329
Appendix 3: Experiment on O- and OB-: Stimuli	333
Appendix 4: Experiment on O- and OB-: Nonce words	345
Appendix 5: Minimal pairs of verbs in PERE- and PRE-	349
Appendix 6: Experiment on O- and U- in factitive verbs: Stimuli	350
Appendix 7: Experiment on O- and U- in factitive verbs: Results	357
References	359 - 376

Acknowledgements

Writing this dissertation has been a long and exciting learning journey. I can compare it with a trip to a mountain top. A journey full of breathtaking views and unforgettable moments. And when I finally make it to the top and can see the entire landscape laid out in front of me, when all separate puzzle pieces at last come together in one picture, I know that my long way was worth it. So many dreams have come true and so many great memories I share today with a great number of people that I would like to thank from all my heart.

First of all, I want to express gratitude to my supervisors – Laura A. Janda and Tore Nessel. I was extremely lucky to be guided by them, always full of ideas, energy, and passion about linguistics. Their work inspired me along all my way, their example helped me to go on. Thank you for your generous guidance, encouragement, care, time, insights, diligence, sharing your knowledge with me and giving me constructive feedback. It was a privilege to work together with you during these years, under your close supervision. Each meeting with you was extremely stimulating and inspiring. I am especially grateful for introducing me to Cognitive Linguistics through your own work as well as through reading seminars, university courses, and numerous conferences.

While working on this project, I have been fortunate to be a member of the CLEAR¹ research group at the University of Tromsø created and led by Laura A. Janda and Tore Nessel. CLEAR has been an incredible intellectual environment where I could share my ideas and doubts, receive both tough criticism and generous encouragement. I want to thank all my remarkable colleagues at CLEAR, especially Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, and Svetlana Sokolova for helpful discussions, advice, pilot experiments, and our regular meetings. I am very much indebted to you, Julia, for teaching me how to work with MySQL software and extract data. I am very grateful to you, Sveta, for our joint projects on the prefixes VZ-/VOZ- and ZA-. Working with you has always been a highly enriching, eventful and lively experience! I owe special thanks to you, Olya, for our brain-storming session in Moscow. It is because of your insights from a broader perspective I decided to inspect a case of Standard Allomorphy in greater detail in a separate chapter. I am most of all grateful to you, Nastya, for helping me to polish the details of experimental designs, comparing our linguistic intuitions, updates about everything, going together to conferences, courses, and summer schools, and also for being so wonderful in all respects.

This work would not have been possible without the funding provided by the Research Council of Norway. This doctoral dissertation is part of the grant received for the research project *“Neat Theories, Messy Realities: How to apply absolute definitions to gradient phenomena”* (2011-2014). I appreciate the privilege of being able to participate and contribute to this research enterprise.

My deepest gratitude goes to the University of Tromsø for employing me in this challenging project and for providing a perfect working environment. I am especially grateful to Eystein Dahl for being so helpful at the final stages of my writing, to Jorun

¹ CLEAR stands for Cognitive Linguistics: Empirical Approaches to Russian.

Nordmo for her assistance with all administrative matters during the entire programme, and of course to Jan Bergheim Helge and Frank Ruben Efraimsen for their advanced IT support.

Working on this dissertation has been an adventure. It took me to different countries, cities, and institutions. I am indebted to the Centre for Advanced Study (CAS) at the Norwegian Academy of Science and Letters, where I spent nine months of my PhD program, and to the Center for Slavic, Eurasian and East European Studies (CSEES) at the University of North Carolina at Chapel Hill (USA) that hosted me during eleven months of my research visit.

Many thanks also to Arlanda Moreno and Alexander Pfaff at the Center for Advanced Study in Theoretical Linguistics (CASTL) at UiT for sharing with me their intuitions about Spanish and German data, being so friendly and positive.

I am grateful to 240 participants of my linguistic experiments for their time and care about language. I highly appreciate their contribution to my work. They made it possible to test my theoretical ideas against linguistic intuitions and individual grammars of real speakers.

Last but not least, I am blessed with having awesome friends and a wonderful family. I would like to thank them for encouragement and love.

List of abbreviations

ACC	– accusative
ADJ	– adjective
INTR	– intransitive verb
IPF	– imperfective aspect
MASC	– masculine
PF	– perfective aspect
PL	– plural
RNC	– the Russian National Corpus
SG	– singular
TR	– transitive verb
V	– verb

Chapter 1

Introduction

In this dissertation I focus on one of the most fundamental notions of modern linguistic theory – the notion of allomorphy. I examine a number of non-trivial cases that do not fully satisfy traditional criteria for allomorphy. Driven by data, my approach suggests that the phenomenon of allomorphy is broader than its traditional understanding and goes beyond the axioms of complementary distribution and identical meaning. In particular, I address cases of semantic dissimilation of allomorphs and cases of conflicting factors that determine their distribution. In the light of my findings, I argue that allomorphy is a gradient phenomenon which can be best captured in terms of a radial category structure, with a central prototype, standard exemplars, and non-standard deviations.

The major objective of this work is to optimize the traditional criteria for establishing allomorphic status and morpheme identity by applying the advantages of modern linguistics: extensive and diverse linguistic data available via electronic corpora, experimental designs, and statistical modeling.

All materials produced in this work, i.e. coded databases, experimental results, and R scripts for statistical analyses are publicly available at the Tromsø Repository of Language and Linguistics (TROLLing), an international archive of data housed at the library of the University of Tromsø (<http://opendata.uit.no/>). The direct link is <http://hdl.handle.net/10037.1/10078> (all files are located at *Data & Analysis*). The full citation is:

Endresen, Anna, 2014, "Non-Standard Allomorphy in Russian Prefixes: Corpus, Experimental, and Statistical Exploration", <http://hdl.handle.net/10037.1/10078>
UiT Open Research Data [Distributor] V1 [Version]

This chapter is organized as follows. In 1.1, I present the amendments that this dissertation offers to the theory of allomorphy. In 1.2, I describe the corpus and experimental data I explore; and in 1.3, I outline the methodology of quantitative analyses that I propose. Section 1.4 provides the theoretical background of this work, followed by 1.5, where I give an overview of the chapters that follow this Introduction.

1.1 What is Non-Standard Allomorphy?

Allomorphy is created by asymmetry between form and meaning. Ideally, allomorphy is a mere variation of form where the meaning remains constant. The traditional definition of allomorphy is simple and short: it is a relationship between morpheme variants which have identical content and which complement one another in their distribution (Matthews 1974: 107; Haspelmath 2002: 27; Booij 2005: 172; Bauer 2001: 14).

For example, the variants of the English article *a* and *an* are allomorphs conditioned by the initial phoneme of the adjacent word (*a book* vs. *an apple*). The first allomorph is restricted to preconsonantal positions, whereas the second allomorph appears in prevocalic contexts, and they never overlap. Except for some cases and some

English dialects where they do (Bauer 1988: 14; Britain 2007; Gabrielatos et al. 2010). A natural question arises in this regard: is there something wrong with English or is there something wrong with our model?

The current model of allomorphy relies on absolute criteria and cannot account for such a distributional overlap. Instead, the model forces us to choose between two possible options for an analysis: either *a* and *an* are allomorphs and the overlap should be ignored, or *a* and *an* are not allomorphs but something else. Neither can this model explain other phenomena observed in natural languages like allomorphs that exhibit submorphemic differences in register or semantics (Booij 1995: 88). We can see that both criteria of distributional complementarity and identical semantics can be violated. The phenomenon of allomorphy is thus more complex and controversial, and does not fit into the narrow and rigid definition cited above. Moreover, the two criteria of allomorphy understood in absolute terms simplify and idealize the ontological nature of this phenomenon. As a result, this approach creates a gap between the idealized object of our study and the raw data, some portions of which have to be ignored.

In this dissertation, I elaborate an alternative usage-based model of allomorphy. According to this model, allomorphy is a scalar relationship between morpheme variants – a relationship that can vary in terms of closeness and regularity. The cases that perfectly satisfy both criteria for allomorphy are considered prototypical and standard, but not the only possible ones. In addition, I recognize *Non-Standard Allomorphy*, which deviates from Standard Allomorphy in that it allows violations of both semantic “sameness” and complementary distribution. However, along with this “imperfection”, Non-Standard allomorphs present compelling evidence that justifies their interpretation in terms of a single perceptible morpheme. In particular, such allomorphs have robust patterns of distribution and strong semantic similarity that can be measured quantitatively.

1.2 Empirical exploration: Russian prefixes

The approach that I pursue in this dissertation is data-driven. I conduct seven empirical studies in order to explore various facets of allomorphic relations. I show that morpheme identity is a phenomenon that often does not fit into the narrow definition of allomorphy, and that it instead can be evaluated via quantitative techniques.

This dissertation is an empirical study of modern Russian derivational morphology and the semantics of aspectual prefixes. The goal is to provide a more accurate account of relevant phenomena of the Russian language.

Russian is well-known for its rich morphological system, which has a broad supply of morphemes with a high degree of variation. A study of Russian data can offer useful implications for the theoretical understanding of allomorphy, because most Russian prefixes are polysemous and many have loan Church Slavonic counterparts that coexist with native prefixes and create a unique situation in the Slavic domain. Therefore, a thorough examination of Russian data promises far-reaching implications for our theoretical understanding of allomorphy. Moreover, Russian is very well-documented via the Russian National Corpus – a large electronic collection² of texts (over 500 million words, as of July 2014) with advanced search possibilities.

I examined in detail 15 Russian prefixes – RAZ-, RAS-, RAZO-, S-, SO-, PERE-, PRE-, VZ-, VOZ-, O-, OB-, OBO-, U-, VY-, and IZ- and assigned them to eight distinct morphemes.

² Freely available at <http://www.ruscorpora.ru>

All prefixes I examined have been studied before, so for each pair of prefixes there exist contradicting claims that the prefixes are either distinct morphemes or allomorphs. In each case study I take an agnostic view and propose the analysis that can best capture the complexity of the available data. Each analysis is data-driven and usage-based and goes from data to generalizations rather than the other way around.

I collected data from the Russian National Corpus, dictionaries, and psycholinguistic experiments with native speakers of Russian. The total number of verbs analyzed in this doctoral dissertation is 4,718 individual lexemes. In addition to verbs, I examined adjectives in PRE-. The experimental data includes 11,138 datapoints (responses of subjects) from two studies³. For a detailed overview of amounts and types of data analyzed for each prefix see Appendix 1.

The results offer a new perspective on the old debate of native vs. loan morphemes (the opposition of East Slavic vs. Church Slavonic strata) and the architecture of affixal semantics. Four of seven case studies are concerned with pairs of prefixes that arose from coexisting native East Slavic and loan South Slavic (Slavonic) sub-lexicons in Modern Russian. These pairs of prefixes include S- and SO-, PERE- and PRE-, VZ- and VOZ-, and VY- and IZ-. They represent different degrees of closeness between the native and the loan rival forms.

In addition, I present two experimental studies with nonce words. In the first experiment I explored the relationship of the prefixes O- and OB- and tested the hypothesis which argues that they have split in two distinct morphemes. The second experiment tested whether the acceptability of novel factitive verbs depends on the relative productivity and semantic default vs. markedness of the prefix (O- vs. U-).

1.3 Statistical modeling

The usage-based approach obliges me to account for large amounts of data (Janda 2013). For this purpose I use various quantitative methods. I suggest that statistical modeling provides additional criteria for establishing allomorphic status and makes it possible to capture Non-Standard allomorphic relations. Statistical modeling resolves many issues.

First, in case there are several factors that possibly affect the distribution of rival forms – how do we find out which factor is most important? Statistical models like Logistic Regression and Classification and Regression Trees (CART) can capture not only multifactorial data dependencies, but also complex interactions of factors. Moreover, the Random Forests analysis can rank the factorial predictors with respect to their relative strength in affecting the distribution of data.

Second, there are cases where the contrast between competing variants is present but not pure or crisp. Rather, the distribution tolerates some amount of overlap. How do we measure the zone of contrast and the zone of overlap? How do we find out what predominates, overlap or contrast? Statistical tests like Pearson's Chi-square test, Cramer's V test, and Fisher's Exact Test can be used in order to assess whether the difference in the distributions of rival variants is statistically significant, robust, and near-complementary.

³ This total number includes 3,878 datapoints collected from 60 subjects in the experimental study of the prefixes O- and OB-, and 7,260 responses from 120 subjects involved in the experiment on prefixes O- and U- in Russian change-of-state (factitive) verbs.

Third, many candidates for allomorphy are polysemous. Inconveniently enough, they might share some meanings and differ in others. How do we apply the criterion of semantic “sameness” to such candidates and objectively estimate their semantic similarity and divergence? I elaborate the methodology called *Radial Category Profiling* (Nesset et al. 2011), which identifies statistically distinct profiles of rival candidates within a radial network of their submeanings. Radial Category Profiles reveal prominent and productive submeanings of polysemous affixes – their “centers of gravity” identified on the basis of verifiable frequency information that speakers are typically sensitive to. Radial Category Profiles can be used as hypotheses about mental representations which can be further tested experimentally. The *Radial Category Profiling* methodology belongs to a family of quantitative methods employed in Cognitive Linguistics that includes inter alia behavioral profiles (Divjak & Gries 2006; Divjak 2010), constructional profiles (Janda & Solovyev 2009, Kuznetsova 2013), and grammatical profiles (Janda & Lyashevskaya 2011).

In this dissertation, I offer detailed applications of several advanced statistical models documented in R scripts. I argue that statistical modeling is a powerful and verifiable technique that extends the boundaries of the criteria of both semantic “sameness” and distributional complementation. If we can assess statistically robust patterns of data distribution, we are freed from the extreme restrictions of the absolute exceptionless criteria for allomorphy. In the following chapters I compare results from multiple statistical models. This practice makes it possible to shield the findings from unwanted biases and gain additional insights about the organization of data.

1.4 Theoretical background

In this study I approach the phenomenon of allomorphy from the perspective of Cognitive Linguistics. In this section I briefly discuss what Cognitive Linguistics is and what theoretical concepts of this framework are relevant for this dissertation.

Cognitive Linguistics is a modern approach⁴ to the study of language that emerged in the early 1970s and is originally rooted in experimental research in cognitive psychology (Rosch 1973). Cognitive Linguistics explores *language as a mental phenomenon*, which organizes, processes and conveys information about the human experience of the world. Therefore, language systems and language use can inform us about the conceptual structures employed in human mind. In this light, the Cognitive Linguistics enterprise is driven by the goal to provide psychologically realistic analyses of authentic linguistic data which can be used as testable hypotheses about mental grammars (Nesset 2008: 9).

In the framework of Cognitive Linguistics, the *language faculty* is not seen as principally different from other cognitive abilities. Quite the opposite, this approach stands for the idea that language is governed by general cognitive principles (for discussion cf. Dąbrowska 2004: 50-75).

Moreover, instead of making an assumption about an innate set of formal rules, Cognitive Linguistics elaborates the *emergentist model*, which views grammar as a system emerging from the user’s repeated exposure to language and to human experience (Tomasello 1992; Smiskova-Gustafsson 2013). Recent experimental studies on language comprehension and production show that language processing is sensitive to the frequency of units larger than individual words. As argued by **Janssen & Barber**

⁴ For detailed overviews see Evans et. al 2007; Nesset 2008: 9-28; Janda 2010.

(2012), these findings contradict the generative computation model of “words and rules” which assumes that the lexicon stores individual words, while the grammar subjects them to formal rules.

Cognitive Linguistics advocates a *non-modular approach* to language. This sets it apart from those modern linguistic schools that postulate discrete and independent modules of phonology, syntax, and lexicon (Chomsky 1965; 1981; cf. also Fodor 1983). There is a growing body of evidence suggesting that there is no clear-cut boundary between grammar and lexicon (e.g. Burzio 2006; Hilpert 2008; Bye 2014). Rather, “lexicon and grammar form a gradation” (Langacker 2008: 5, 21; 1987: 17). As Janda (2010: 6) puts it, “[m]eaning is not tidily contained in the lexicon, but ranges all through the linguistic spectrum <...> Grammar is an abstract meaning structure that interacts with the more concrete meanings of lexicon.” Taking meaning as a “driving force of language” (Janda 1993: 310), Cognitive Linguistics pursues the primacy of semantics in all linguistic phenomena. Therefore, Cognitive Linguistics advances the study of meaning on all levels of linguistic structure.

Proponents of Cognitive Linguistics do not postulate underlying representations and do not speculate about highly abstract symbolic structures. Instead, they make surface-based generalizations termed *schemas*. This practice is facilitated by interdisciplinary studies that point towards a *non-hierarchical model of language use* suggesting that linguistic structures are linear rather than hierarchical (Frank et al. 2012).

Whereas formal approaches postulate the notion of an idealized language speaker (Stokhof & van Lambalgen 2011: 6), Cognitive Linguistics takes a different position regarding this issue. Recall that formal tradition is guided by the idea that “[a]ny serious study will <...> abstract away from variation tentatively regarded as insignificant and from external interference dismissed as irrelevant. <...> a significant notion of ‘language’ as an object of rational inquiry can be developed only on the basis of rather far-reaching abstraction.” (Chomsky 1980: 219). By contrast, Cognitive Linguistics holds the view that language is shaped by its communicative function, and therefore this framework adopts a *usage-based approach* to linguistic data. In particular, this approach draws attention to variation and gradience as relevant properties of observed phenomena. It is a common practice in cognitive studies to explore extensive sets of data extracted from *electronic corpora* and distinguish between robust meaningful patterns and random effects by means of *quantitative methods* (cf. Janda 2013).

Note that Cognitive Linguistics is not a single doctrine. Rather, it is a *family of compatible approaches*, a “building with many rooms” (Geeraerts & Cuyckens 2007: 10). The scope of Cognitive Linguistics is not limited to studies of figurative language. Cognitive Linguistics has grown into a powerful innovative school of linguistic thought with elaborated terminology and methods that pertain to linguistic phenomena of different levels: syntax (Goldberg 1995; Croft 2001), phonology (Neset 2008), morphology (Janda et al. 2013), lexicon (Divjak 2010), sociolinguistics (Geeraerts et al. 2010), diachronic studies (Geeraerts 1997; Bybee 2010), typology (Evans & Levinson 2009), computation of linguistic communication in robots (Steels & Hild 2012), and other domains.

The key concept of Cognitive Linguistics relevant for this dissertation is the concept of a *radial category*. A radial category is a relationship of units hierarchically

organized around the central member or sub-category called the *prototype*.⁵ The prototype has a special status in the category because it is the most salient member of the network. On the one hand, it unifies the category into the whole, on the other hand, it motivates the variation found in less central members or peripheral members of the category. The status of other members is established on the basis of similarity with the prototype. Similarity is a scalar characteristic but can be measured quantitatively. Therefore, members of the category may share some of their properties with the prototype and each other and differ in other properties. Radial category is usually defined by the principles of *family resemblance* rather than via binary features. In this dissertation, I apply the notion of a *radial category* to model the phenomenon of allomorphy on the one hand (see Chapter 2) and to model the semantics of polysemous affixes analyzed as candidates for allomorphic relations (case studies in Chapters 3-9).

This work is inspired by a long and well-established tradition of analysis which makes subtle distinctions between different senses, or submeanings, of affixes (Janda 1986; Nessel 2009; Janda & Nessel 2010), particles, and prepositions (Brugman & Lakoff 1988). A crucial premise of this analysis is that linguistic meaning is *embodied*, that is “grounded in the shared human experience of bodily existence” (Janda 2010: 10). Therefore, the primary and central meaning that organizes the rich polysemy of affixes usually refers to concrete basic spatial relations based on orientational notions like UP, OUT, APART, WITH, ACROSS, FRONT, CONTAINER, etc. Each prefix analyzed in this dissertation implies a certain scenario, where a foregrounded object (*Trajector*) moves along a certain *Trajectory* (or Path) with regard to a backgrounded object (*Landmark*). Such a scenario is concrete and at the same time schematic. Therefore, a visual representation of this scenario is referred to as a spatial *image schema* (Johnson 1987, Evans 2007: 106; parallel to *configuration* in Janda 1986). Furthermore, one can focus on certain parts of image schemas, and thus obtain different *construals*. The human capacity to *construe* situations in various ways plays an important role in Cognitive Linguistics (Verhagen 2010).

Typically, spatial notions have rich semantic potential and serve to motivate more abstract meanings in a radial category. This becomes possible by virtue of the cognitive mechanism of *metaphor* which is a mapping from a source domain (e.g. spatial relations) to a target domain (often, a more abstract domain like temporal relations or emotional states; Lakoff & Johnson 1980). We can observe this in combinations of prefixes with different bases, where verbs that refer to concrete motion make use of the spatial image schema of the prefix (*vybežat* ‘OUT-run’ < *bežat* ‘run’), whereas non-spatial verbs interpret the same schema metaphorically (*vylečit* ‘OUT-treat’=‘cure from illness’ < *lečit* ‘treat’). Another cognitive mechanism that relates different submeanings is *metonymy*, e.g. referring to the whole by naming its part (Janda 2010: 21). In prefix semantics we often observe *metonymy* in the reduction of movement along the trajectory to the end-point of the trajectory (compare the use of the prefix SO- in *soedinit* ‘conjoin’ and *sosuščestvovat* ‘co-exist’). These are the key notions of Cognitive Linguistics that I make use of in this dissertation.

⁵ This point can be traced back to the Jakobsonian tradition, because *radial category* and *prototype* were employed in works of Roman Jakobson in terms of a hierarchy of specific meanings and the notion of “relative invariant” (cf. Janda 1993: 311; Sangster 1982: 78 for discussion).

1.5 Structure of the dissertation

The dissertation consists of ten chapters. In addition to the Introduction (Ch.1) and Conclusion (Ch.10), there is a theoretical chapter (Ch.2) and seven chapters that present individual case studies (Ch.3-9). In the first of them (Ch.3), I provide an account of a Standard allomorphic relationship, whereas the remaining six studies are devoted to Non-Standard allomorphies (Ch.4-8) and examples of Non-Allomorphy (Ch.9). I outline each chapter below.

Chapter 2 argues that allomorphy is a theoretical construct that fails to accurately capture the nuances of empirical data. I start with a number of well-known textbook examples of allomorphy that turn out to be more controversial than typically assumed. I further turn to problematic aspects of the traditional definition and criteria for allomorphy, look at the origins of this notion, and outline my theoretical proposal.

Chapter 3 reports on a study of Standard Allomorphy conditioned by phonological and morphophonological factors. I look at two phenomena – voicing assimilation across a prefix-root boundary (prefixes RAZ- ~ RAS- ‘apart’) and vocalization of consonant-final Russian prefixes (RAZ- ~ RAZO- ‘apart’). I show how a statistical analysis models the distribution of polysemous but standard allomorphs and evaluates the relative impact of each factor.

Chapter 4 examines the Non-Standard Allomorphy of the prefixes S- and SO- ‘downward, with’ conditioned by a mixture of phonological, morphophonological, semantic, and stylistical factors.

Chapter 5 addresses a long-standing debate about the status of the prefixes O- and OB- ‘around’ conditioned by conflicting phonological and semantic patterns. I present novel corpus and experimental data that speaks for Non-Standard Allomorphy and against the Hypothesis of Morphological Split. I also show how statistical analysis can shed more light on the hierarchical ranking of factors involved in the conditioning.

Chapter 6 tells the story of the native Russian prefix PERE- ‘across’ and its loan cognate Slavonic counterpart PRE- ‘very’. Contrary to most accounts that view them as different morphemes, I propose an account in terms of grammatically conditioned Non-Standard Allomorphy suggesting that allomorphy via borrowing is possible. I propose that both prefixes represent a single morpheme-intensifier, where PERE- functions as a productive perfectivizer and intensifier of activity, specialized for the verbal domain, whereas PRE- is an intensifier of property and weak perfectivizer specialized beyond verbs.

Chapter 7 explores the correlation of the Russian and Slavonic prefixes VZ- and VOZ- ‘up’. This study is primarily devoted to their semantics which is largely motivated by the central spatial image schema of upward movement. I argue that the two prefixes differ in terms of scale of the path: the native prefix refers to a shorter trajectory of upward motion than the loan prefix. The difference in prototype makes this a borderline case between Non-Allomorphy and Non-Standard Allomorphy with historical and semantic conditioning.

Chapter 8 presents the relation between the prefixes VY- and IZ- ‘out of’ – the pair that has the largest number of verbs that combine with both prefixes among all case studies of this dissertation. I propose that the Modern Russian IZ- conflates uses of two origins: the East Slavic IZ-, native to Russian, and the loan Slavonic IZ-. The loan IZ- correlates with VY- in terms of Non-Standard Allomorphy conditioned by register. The native IZ- is a distinct morpheme that differs from VY- and semantically implies a

different spatial schema. This makes a borderline case between Non-Allomorphy and Non-Standard Allomorphy.

Chapter 9 examines prefixes productively employed in Russian change-of-state (factitive) verbs. They share the same function of verbalization, differ in formal shape, and thus appear to be candidates for suppletive allomorphy. I address the overall picture and focus on the two most prominent prefixes O- and U-. I present and analyze novel corpus and experimental data and explore the factors that condition the choice of the prefix.

Chapter 10, the Conclusion, brings together my findings and summarizes the contribution of this dissertation. I propose that the pairs of prefixes I have examined constitute a scale of different degrees of “closeness” and discuss additional criteria that are relevant in making subtle distinctions between these cases.

Chapter 2

Allomorphy as an overgeneralized assumption: Theoretical construct vs. empirical phenomenon

“It is, rather, a common-sense statement of situations with which we deal constantly, and which we must describe in such a way as not to violate the very evident relationship to which we intuitively react.”

E. Nida (1948: 433)

2.1 Introduction

This chapter aims to provide a critical overview of the scholarly literature, presenting both theoretical discussion and empirical investigations of the phenomenon of allomorphy. I focus on the problematic aspects of the notion as it has been shaped in modern linguistics and argue that the current theoretical model of allomorphy should be revisited. I suggest instead that the model of allomorphy should be optimized with respect to the complex and gradient nature of authentic language data. This chapter discusses the origins of the notion *allomorphy* in structuralist linguistics and the theoretical assumptions that constitute the “baggage” of this term.

I start in 2.2 with a number of textbook examples that are more controversial than usually assumed. In 2.3-2.5 I proceed to the current theoretical model of allomorphy which fails to account for these examples. Next, I turn to the structuralist implications of the term *allomorphy* in 2.6 and discuss the original theory of *morphemes*, *morphs* and *allomorphs* as it was advanced by American descriptivists. The chapter concludes in 2.7. with my proposal of an extended and modernized model of allomorphy that will be further tested and elaborated in this dissertation. I wrap up the discussion in 2.8.

2.2 Puzzles of data: challenges to the definition

Allomorphy is traditionally defined as a structural relation of two or more variants of a single morpheme that satisfy two criteria: 1) identical meaning (or function) and 2) complementary distribution, so that their phonological, grammatical, or lexical environments never overlap (Matthews 1974: 107; Haspelmath 2002: 27; Booij 2005: 172; Bauer 1988: 13; Bauer 2001: 14). Crucially, these criteria are traditionally understood in absolute terms. This means that if our data satisfies these criteria, then we attribute it to allomorphy, and if these criteria are not satisfied, then we assume that it is clearly not allomorphy but something else.

There is a handful of examples (mostly from English morphology) that are often cited when one attempts to illustrate what allomorphy is. Let us look at some of them, which are favored by dictionaries and textbooks as “simple”, “clear”, and “representative” cases of allomorphy, but which in fact reveal a more complicated nature when examined in depth.

2.2.1 English indefinite article

It is well-known that the English indefinite article has two shapes: *a* in front of consonants and *an* in front of vowels (e.g. *a man* vs. *an apple*). The relationship of the two articles in English is typically analyzed as phonologically conditioned allomorphy of a single morpheme {indefinite article} (Nida 1948: 420). This analysis is based on the assumption that the two variants of the indefinite article never occur in the same phonological environment and thus exhibit complementary distribution (Bauer 1988: 13). However, while making this assumption, Laurie Bauer is aware of that in reality the distribution of *a* and *an* cannot always be described in terms of perfect complementation:

“In some rather conservative varieties of English this rule is not quite true, since it is possible to say *an hotel* and *an historical novel*. Not all such speakers pronounce these words without an /h/, which would make them conform to the general rule. **We shall provisionally ignore these varieties.**” (Bauer 1988: 14; highlighted by me – AE)

Indeed, such controversial data does not fit the strict definition of allomorphy. These instances might be just unimportant peripheral examples that deviate from the “normal” neat complementary distribution. Indeed, deviations like this are often “swept under the carpet”, because such examples might seem minor, marginal and irregular and therefore can easily be ignored for the sake of clear-cut definitions. They seem to “spoil” an overall general picture.

However, a number of large-scale corpus studies (Gabrielatos et al. 2010; Britain & Fox 2009; Britain 2007) have shown that in this particular case we are dealing with a much bigger phenomenon than just a couple of counterexamples. Rather, the overlap in the distribution of *a* and *an* in English dialects turns out to be a broad and frequent phenomenon attested both in old traditional British dialects and in modern youth vernacular slang in London, as well as in a number of sociolects of the American, Australian and New Zealand Englishes. Gabrielatos et al. (2010: 298) report that the article form *a* in both prevocalic and preconsonantal positions is a common feature of many varieties of English today:

“There is evidence from British English dialects of *a* as a universal indefinite article. The absence of alternation (*a* vs. *an*) was already identified by Write (1905) and has been, for example found in traditional dialects in the southwest of England (Wagner 2008).” (Gabrielatos et al. 2010: 298)⁶

“Prevocalic indefinite article *a* is also found in Bolton in the North-West (Shorrocks 1999: 45), in Peasmarch in Sussex in the South-East (Lodge 1984), and across the South-West of England (Wagner 2004: 155).” (Britain 2007: 182)

Moreover, Gabrielatos et al. (2010) report on a significant increase of both density and spread of the use of the *a* article in vowel-initial contexts in London in the period between 1993 and 2005: “We argue that *a* + vowel does not primarily index social class but rather that it is a feature of the young speakers in inner London. <...> It indexes inner-city youth language.” (Gabrielatos et al. 2010: 323).

⁶ References of works cited in this and other quotations should be found in the original works.

Further pieces of evidence on the distributional overlap of the indefinite articles *a* and *an* come from a large number of studies cited in Britain 2007:

“Beyond the UK, Taylor (2003) finds *a* + vowel in written court transcripts of defendants in 1850s Sydney, **Australia**; Wolfram and Fasold (1974), Mufwene (2001), and Craig, Thompson, Washington, and Potter (2003) mentioned it as a characteristic of **African American Vernacular English**, Labov (1972) finds it in **New York**. Watermeyer (1996: 118) notes that for speakers of **Afrikaans English in South Africa**, “there is usually no distinction between *a* and *an* (*a energetic class; a absolute miracle*).” (Britain 2007: 182) (boldfaced by me – A.E.)

To conclude, while the complementary distribution of *a* and *an* still holds for Standard English, within a number of dialectal and sociolectal varieties of English this observation is not correct. Overlap in phonological environments of the two articles does exist and should be accounted for. A question arises of whether one can still call *a* and *an* standard (or regular) allomorphs, given that their distribution might be far from what we call complementary. If so, one has to admit that *a* and *an* represent distinct synonymous morphemes, but this contradicts native speakers’ intuition.

If we want to maintain a more natural observation of their submorphemic mutual status, one should instead adjust the definition of allomorphy so that it can be applied to empirical data. We find that complementary distribution is not an absolute and necessary criterion for an allomorphic relation even in case of clear phonological conditioning of distribution. However, if we dismiss the complementary distribution criterion from the definition of allomorphy, we risk losing a crucial part of the asymmetry between form and meaning.

If we agree that there can be a number of exceptional counterexamples of distributional overlap that we can simply ignore, what is the maximum number of exceptions that we are allowed to dismiss? On the other hand, since exceptions to complementary distribution exist, why should we continue to state the distributional criterion of allomorphy in absolute terms?

2.2.2 English past participle marker

The past participle formation in English is another commonly used illustration of allomorphy. The suffixes *-en* and *-ed* are not related phonologically and are usually analyzed as suppletive allomorphs of a single morpheme. Most verbs attach the suffix *-ed* (e.g. *pave* > *paved*) but some verbs make use of the suffix *-en* (e.g. *give* > *given*) (Haspelmath 2002: 28; Matthews 1991: 116).

However, there are some verbs that can take either of the two suffixes, like the verb *show*: both forms (*have*) *shown* and (*have*) *showed* are attested. Admitting this fact, Nida (1948: 432) states that *-en* and *-ed* no longer exhibit complementary distribution. Moreover, Nida argues that the forms *shown* and *showed* also demonstrate a difference in terms of register: “the difference of meaning is precisely that which depends on the distribution of the two allomorphs. The allomorph /-d/ is productive: its potential distribution is not arbitrarily fixed like that of /-n/. Both allomorphs may occur in the same person’s speech, but /-d/ occurs in more colloquial socio-linguistic environments.” Nida classifies the cases like *shown* – *showed* as “instances of ‘overlap’, i.e. forms which are in complementary distribution except at certain points where there is a contrast resulting from fluctuation of forms” (Nida 1948: 431-2).

Bernard Bloch, who assigns *-en* and *-ed* to a single morpheme, in the face of the overlap in *shown – showed* decides to recognize in *show* two separate morphemes: *show*₁ in *shown* and *show*₂ in *showed*, so that the allomorphic definitional requirement of complementary distribution will be satisfied (Bloch 1947: 406). Nida comments on Bloch’s account that it is not wise to go to such extremes “within the rigid confines of the method” (ibid: 432).

Another example of such an overlap comes in the verbs *wake* and *awake*, which can form a participle with either of the two suffixes – *(have) woken* vs. *(have) waked* and *(have) awoken* vs. *(have) awaked*. The suffix *-en*, used by “weak” (irregular) verbs, is considered standard for *wake*, whereas suffix *-ed*, typical for “strong”, or regular, verbs, is also possible: *wake – woke – (have) woken* and *wake – waked – (have) waked*; *awake – awoke – (have) awoken* and *awake – awaked – (have) awaked*.

One more example of distributional overlap involves the two rival forms of the past participle for the verb *hang*: *hung* vs. *hanged*. Instead of two competing participial suffixes, this pair of competing forms features *umlaut* vs. the suffix *-ed*, but the principle is the same. For centuries these forms have been used in English interchangeably. In Modern English, the two competing forms show some semantic distinctions: *hanged* refers to official executions, whereas *hung* is possible elsewhere: *killers are hanged* vs. *pictures are hung*.

Summing up, the examples of overlap in the distribution of the English participial formants are not numerous, but they do exist and violate the criterion of distributional complementarity. Strictly speaking, the suffixes *-en* and *-ed* cannot be called allomorphs since one of the criteria is not satisfied. This means that we have to lose an important generalization about the two affixes that perform the same function in English grammar. Instead, one has to go for a very unnatural and uneconomic analysis and divide the base verbs into homonymic but distinct morphemes. Again, should not we rather revise the definition of allomorphy so that it could account for such examples?

2.2.3 English plural marker

The English plural suffix *-s* is commonly analyzed as a morpheme that exhibits a clear case of phonologically conditioned allomorphy (Matthews 1991: 118; Trask 1996: 272; Haspelmath 2002: 29, 30; Spencer 1991: 6; Lyons 1968: 184; Bloomfield 1933 [1961]: 211). One of the three allomorphs [s], [z], or [əz] / [ɪz]⁷ (*cats; dogs; buses, dishes, edges*) is chosen depending on the phonological environment.⁸ However, it was noticed long ago that the distribution in this case is not precisely complementary:

“Sometimes we are confronted with a set of alternants with apparently identical meaning which are almost, but not quite, in complementary distribution. So with the two alternants meaning ‘noun plural’ in *hoofs* and *hooves*, or *laths* with /θ/ and with /ð/. These would forbid the tactically desirable conclusion that there is but one noun-plural morpheme in English.” (Hockett 1947: 323)

⁷ The vowel represented here by ə and ɪ can be pronounced in different ways depending on the variety of English (Spencer 1991: 461).

⁸ By contrast, Trask (1996: 162) argues that this alternation is triggered by the “presence of a particular grammatical morpheme”, and therefore is grammatically, or morphologically, conditioned: *knife ~ knives* as opposed to *knife ~ knife’s*, where a “phonologically identical suffix fails to trigger the alternation”.

Other nouns that allow both pronunciation variants of the plural marker are those that end with *f* or *fe* like *belief*, *grief*, *handkerchief*, *chief*, as well as *proof*, *poof*, *wharf*, and *dwarf*. Interestingly, the latter noun makes a slight semantic distinction for the two rival plural forms: *dwarfs* and *dwarves*. The form *dwarves* was popularized by J.R.R. Tolkien and specifically refers to the legendary dwarf race, whereas the only accepted plural form at that time was *dwarfs*. Today, both forms are used in English. Similarly, Tolkien insisted on distinguishing between his legendary *elves* and “other” *elfs*.

Regarding the forms *roofs* ~ *rooves* and *hoofs* ~ *hooves*, Nida (1948: 432) points out that “some speakers would agree that they are in completely free variation, and they may be right”. Yet, the difference between the two variants in each pair can be attributed to productivity and sociolinguistic connotations: “The former is the productive type. The latter is non-productive, with a strictly limited distribution; within that range, it has greater socio-linguistic acceptability in some social groups but is considered pedantic in others.” (ibid:432)

We can again observe that the strict model of allomorphy which requires complementary distribution forces us either to ignore the examples of distributional overlap or to reject the allomorphic analysis of this data and thus distort the morphology of English.

Moreover, as for the plural suffixes, the allomorphy of English plural markers can be extended to such nouns as *oxen*, *teeth*, *formulae*, *cherubim*, *criteria*, *memoranda*, *mafiosi*, *schemata*, *indices*, and *crises*. Spencer (1991: 40) points out that “[i]n *oxen* we have a rare vestigial *-en* affix; in *formulae*, *criteria* and *memoranda* we have a Greek or Latin plural ending replacing what might be thought of as a singular ending *-a*, *-on*, or *-um*. *Schemata* shows a more complex example of a Greek plural.” In addition, the formation of the English plural permits optional doublets like *formulae* – *formulas* (Spencer 1991: 430).

Spencer raises a natural question with regard to this variety of English plural markers: “Are all these allomorphs of a single morpheme? If so, what reasonable theory of allomorphy will allow us to say that the vowel ablaut of *teeth* and the *-im* in *cherubim* bear the same relation (of allomorphy) to each other as the different pronunciations of the *-z* plural morpheme bear to each other?” (ibid: 40)

There are different solutions in the literature. A symptomatic suggestion is proposed by Lyons (1968: 186), who argues that irregular plurals should be viewed as exceptions and ignored in the analysis: “Since the formation of the word *oxen* is an irregular fact of English, which, despite the segmentability of the word into two constituent morphs, can only be handled by an *ad hoc* ‘rule’ applying to this one instance, there is little point in recognizing /ən/ as an allomorph of {s} in the description of Contemporary English.”

Trask (1996: 203) suggests that in case of plurals like *cacti*, *men*, *sheep*, *children* we deal with the use of “exceptional allomorphs in the presence of specific lexical stems”, and the choice of the plural marker is lexically conditioned. Bauer (1988: 14) expresses the same line of argument: “[i]t is a peculiarity of the lexeme *ox* (as opposed to *box*, *cox*, *fox*) that it takes a plural marker *-en*.” Therefore, this allomorph of the morpheme {plural} is determined by a particular lexeme and is lexically conditioned.

By contrast, Spencer (1991: 40), who follows Matthews 1972, argues that in this case it is more appropriate to distinguish between grammatical categories and their exponents (“i.e. the linguistic material that expresses those categories”) rather than to interpret English plural markers in terms of morphemes and their allomorphs. In particular, Spencer (1991: 431) discusses the Separation Hypothesis advocated most

explicitly in works of Beard (1982⁹; 1987; 1995). This theory stands for the idea of separating morphological form and morphological function of affixes and denies the prior status of morpheme as a bearer of morphological meaning. In other words, this account escapes the problem of the distributional requirement by means of eliminating the very idea of morpheme as a pairing of form and meaning altogether.

Summing up, in the case of English plural formation the standard theory of allomorphy is challenged by a) the distributional overlap of allomorphs and b) a variety of remnant plural markers that are assigned the status of exceptions. Again, one can choose to pursue the theory or the data.

2.2.4 English first-person singular pronoun

The forms of English first person pronoun *I* and *me* can be analyzed as allomorphs that are distributed in terms of syntactic positions: *I* is used in preverbal subject position, whereas *me* occurs after prepositions and in postverbal object position (Nida 1948: 422). As Nida points out, this distribution is complementary except for phrases that indicate an overlap, like *It's me* vs. *It's I* and *for you and me* vs. *for you and I* (ibid: 422, 433).

Hockett (1947: 342) observes that there are also dialectal differences: in some dialects the forms *I* and *me* (as well as *we* and *us*, *he* and *him*) are in complementary distribution, whereas in others they exhibit partial overlap and non-contrastive distribution. According to Hockett, exceptions to complementary distribution are “either on the Latin pattern (*It's I*, or *Who's here? – I*, instead of *Me*), or are overcorrections (*between you and I*). For many speakers whose usage of *I* and *me* does not put them in complete complementation, there is no contrast between, for example, *It's I* and *It's me*. In other dialects and styles, on the other hand, the forms are in contrast: literary English, schoolteachers' on-duty English, and certain whimsical styles.” (ibid: 342)

Strictly speaking, this distributional overlap of *I* and *me* violates the requirement for perfect complementary use which defines allomorphs in the first place. According to this logic, the two forms of English first person pronoun should be assigned to distinct root morphemes. However, such an analysis is counter-intuitive. Both Hockett and Nida propose an allomorphic account of this data and modify the strict requirement for complementary distribution. Hockett argues for non-contrastive distribution, which allows for distributional overlap with no semantic difference. Nida suggests a more subtle distinction: “Related forms which occur in the same environment, but which could otherwise be regarded as allomorphs, can still be so regarded if (a) there is no apparent difference in meaning between them or (b) the difference of meaning is derivable from the distributions of the related forms.” (ibid: 431) The former case (a) refers to free-variation under semantic identity of allomorphs and is similar to what Hockett proposes. The second condition (b) is what Nida calls sub-morphemic semantic differences of allomorphs – a natural assumption which follows from his idea that distributional differences of variants inevitably introduce and reflect semantic contrasts. Examples of this type are given in the next subsection.

⁹ Beard 1982 argues that pluralization of English nouns is a lexical derivation. See also (Wickens 1992: 5) for critical discussion.

2.2.5 German diminutive suffixes: *-chen* vs. *-lein*

German has two diminutive suffixes *-chen* /-çən/ and *-lein* /-ləjn/ which can be analyzed as allomorphs of a single morpheme {diminutive}. Both suffixes assign a neuter gender to nominal bases (Sternefeld 2006: v.1: 31). Both suffixes trigger umlaut of the vowels *a*, *o*, and *u* in the base¹⁰ (Féry 1994: 9; Sternefeld 2006: v.2: 696). Moreover, both suffixes are able to attach to a stem after a plural marker: *die Kind+er+chen*, *die Kind+er+lein* ‘little children’, where the stem is *Kind* ‘child’ and the plural marker is *-er* (Sternefeld 2006: v.1: 106). There is a lot of dialectal variation in the use of these suffixes (e.g. the suffix *-le* instead of *-lein* in Southern varieties of German, and the suffix *-el* discussed in Beard 1995: 74, 95). The choice between *-chen* and *-lein* can be governed by dialectal preferences, phonological conditioning, and idiomatized uses. Crucially, in some phonological environments these suffixes show complementary distribution typical for standard allomorphs: e.g. *Buch* ‘book’ > *Büchlein* ‘small book’, no **Büchchen*, but *Teil* ‘component’ > *Teilchen* ‘part, particle (in chemistry)’, no **Teillein* (Luschützky 2000: 460). Beard (1987: 7; 1995: 74) points out that only *-chen* attaches to nouns ending in /l/ and only *-lein* occurs with stems ending in /x, ŋ, g/.

However, upon closer consideration, both distributional and semantic criteria for standard allomorphy turn out to be violated by a number of counterexamples. In particular, there are nouns that can attach either of the two suffixes, and the resultant derivatives are often synonymous, as in the pairs *Tischchen* ~ *Tischlein* ‘small table’ (< *Tisch* ‘table’), *Schränkchen* ~ *Schränklein* ‘small cupboard’ (< *Schrank* ‘cupboard’), *ein Türchen* ~ *Türlein* (or *Tourchen* ~ *Tourlein*) *machen* ‘take a short walk’ (< loan word *Tour* ‘tour’).

Other derivatives formed by the two suffixes can share some meanings but differ significantly in others, as in the pair *Männlein* and *Männchen* (< *Mann* ‘man’). Both of these diminutive nouns can denote ‘short man’, but they split in the senses ‘boy’, exclusive for the former *Männlein*, and the sense ‘male (as biological species)’, reserved to the latter *Männchen*.

The difference in semantics is even more distinct in idiomatic usages like *Fräulein* ‘unmarried woman, miss’ vs. *Frauchen* ‘female owner (of a pet)’ both of which are derived from *Frau* ‘married woman’.

On the one hand, this data clearly violates both semantic and distributional requirements for allomorphy: the two suffixes can attach to the same stems and in addition contribute different semantic content. In other words, this undermines the allomorphic account for *-chen* and *-lein*. On the other hand, the two suffixes share general diminutive semantics, have in common some morphological properties (like assigning a neuter gender), trigger umlaut in the base, and exhibit an idiosyncratic distributional pattern.

Luschützky (2000: 460) examines these diminutive suffixes and suggests that transitional phenomena of this kind are common among allomorphs. Luschützky

¹⁰ This property is rather special for these suffixes. Féry (1994: 9) claims that in German “[v]ery few, indeed only two, suffixes seem to trigger umlaut productively, namely the diminutive suffixes *-chen* and *-lein*. In their case only, umlaut is a morphologically-triggered phonological phenomenon. <...> Derivation with *-lein* has essentially the same properties as with *-chen* except for a few lexical and phonologically conditioned variations.” In addition, both suffixes share the condition in which they trigger umlaut – “only when a Foot is formed by the last syllable of the stem and the suffix itself.” (Féry 1994: 4).

explicitly states that in order to draw a conclusion in such cases it is common to eliminate the unwanted counter-intuitive outcomes of the analysis.¹¹

Again we can see that one has to choose between being loyal to data or to the traditional definition of allomorphy. For some reason, it is not enough to encounter a relevant correlation of linguistic forms in order to name it allomorphy. The definition puts too strong restrictions on the type of phenomena that are included in its scope, leaving all deviations, fluctuations, and transitions beyond the focus of interest. I suggest that the notion of allomorphy is in itself a theoretical construct that is rarely perfectly represented in real data and should be revisited.

2.2.6 What do we learn from this data?

Summing up this discussion, I have looked at five textbook examples of allomorphy – English indefinite article, past participle marker, plural marker, first-person singular pronoun, and German diminutive marker. Although we expect them to conform to the mainstream definition of allomorphy, all these examples are in fact problematic. Upon closer investigation, they reveal severe violations of the two principal criteria for regular allomorphy. In each example, we observe “defective” complementary distribution, because in some environments allomorphs overlap. Moreover, the variants of English past participle marker and German diminutive marker develop semantic distinctions and therefore do not satisfy the requirement for identical meaning.

According to the definition of allomorphy that modern linguistics operates with, none of these examples represents an allomorphic relationship. This suggests that the definition of allomorphy is imperfect, as it fails to account for cases of significant relationships in natural languages by excluding them from its scope of application. There is a strong contradiction, or a gap, between the definition of allomorphy and the data that represents this phenomenon. I will now turn to a discussion of the modern definition of allomorphy and will focus on its problematic aspects.

2.3 Allomorphy: Asymmetry between Form and Meaning

We can pose a question: why do we need the notion of allomorphy? The answer is: in order to describe a certain type of mismatches, or asymmetries, between two sides of a linguistic sign – the form and the meaning.

The theory of the linguistic sign introduced by Ferdinand de Saussure (1916) posits that a sign is a pairing of form and meaning – a signifier and a signified, as I show in Figure 1. The rectangles represent two sides that constitute a linguistic sign, and the vertical line visualizes an arbitrary link which connects form and meaning. However, the correlation between two sides of a sign, be it a word, or a morpheme, is often more complex than a simple *one meaning – one form* relationship.

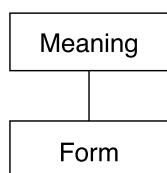


Figure 1: The structure of a linguistic sign.

¹¹ “Die Einwort-Definition von Allomorph als ‘Morphemvariante’ läßt zwar offen, inwieweit sich solche Varianten in ihrer lautlichen Erscheinungsform voneinander unterscheiden dürfen, doch hat es sich eingebürgert, weitere Kriterien zur Ausschaltung intuitiv unerwünschter Analyseergebnisse heranzuziehen.”; “Übergangsphänomene dieser Art kommen häufig vor...” (Luschützky 2000: 460).

The mismatch, or asymmetry, of form and meaning can be of different kinds. In Figure 2, I schematically depict four types of asymmetries that correspond to well-known linguistic phenomena – synonymy, allomorphy, homonymy, and polysemy. This visualization might simplify these phenomena, but it captures the key differences and at the same time places allomorphy in the context of other form – meaning relationships.

Asymmetry 1 represents synonymy – the situation when several linguistic forms are associated with the same meaning. Ideally, synonymous forms are phonologically different and are not related to each other, therefore the rectangles which represent Form 1, Form 2, and Form 3 lack connection lines. This is the only thing that makes synonymy different from allomorphy.

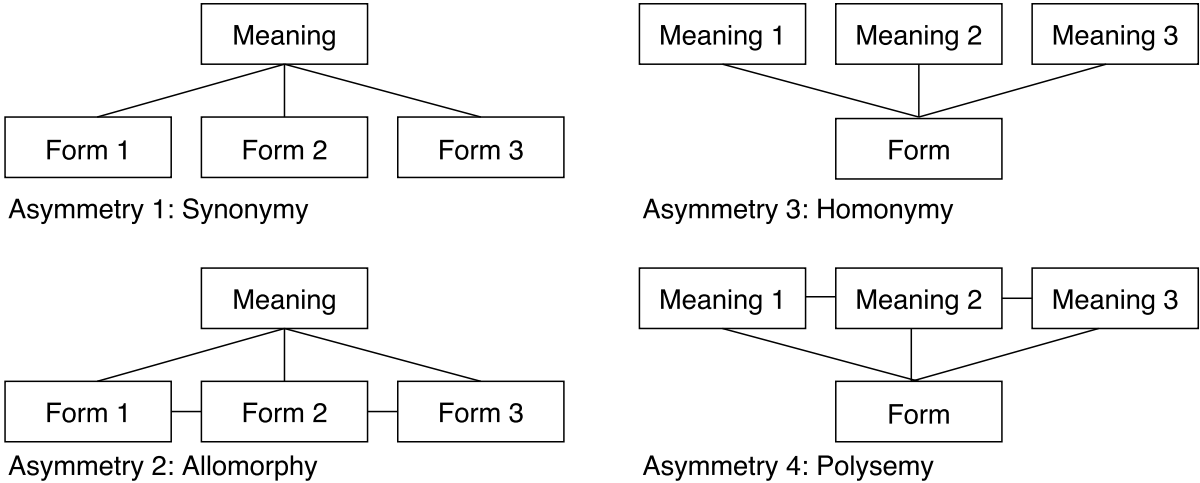


Figure 2: Types of *form – meaning* relationship.

Allomorphy describes a similar type of asymmetry between form and meaning: there are several forms that correspond to a single meaning. The term itself is a combination of Greek roots *allos* ‘other’ and *morph* ‘form’. The fundamental idea of allomorphy is that these forms are related to each other – via distributional rules and preferably also by similarities of their formal shapes. Allomorphy is visualized as Asymmetry 2 in the bottom left corner of Figure 2. Even if allomorphs are not phonologically similar and are not phonologically derivable from each other (as in case of suppletive allomorphs) they are related in terms of identical meaning and non-overlapping, i.e. complementary, distribution. This model depicts a traditional understanding of allomorphy widely presented in scholarly literature.

Asymmetries 3 and 4 correspond to homonymy and polysemy respectively. As opposed to synonymy and allomorphy which deal with the problem of multiple forms, both homonymy and polysemy are concerned with variation of meaning. In both cases, several meanings correspond to a single form, with the only difference that in case of homonymy these meanings are unrelated, while in polysemy the meanings are related. As a consequence, homonyms are analyzed as distinct linguistic signs (words or morphemes), unlike related meanings that are interpreted as semantic components of a single polysemous linguistic sign.

The four types of *form – meaning* asymmetries are represented in Figure 2 as clearly different from each other in theory. However, they do not exhaust all the possibilities that linguistic data can exemplify. First of all, we can encounter variation of both form and meaning – when several forms have several meanings with no one-to-one

correspondence. Linguistic data undermines clear-cut boundaries set up by a strict definition, and the contrasts that we search for can vary, showing gradient nature and dependence on the context. Moreover, many theories claim that there are no synonyms that are semantically identical (Nida 1948, Janda & Solovyev 2009, Divjak 2010).

The four types of asymmetries in Figure 2 are rather idealizations, or theoretical constructs, that empirical data can be assigned to. The data can appeal not only to allomorphy but also, in some degree, to synonymy, or polysemy.

As I showed in section 2.2, the model of allomorphy presented in Figure 2 fails to represent some relevant phenomena of natural languages. This model is a theoretical construct created within a particular linguistic tradition. It fails to account for various non-trivial cases. In Figure 3, I show two alternative possibilities of *form-meaning* asymmetry that I interpret in terms of Non-Standard Allomorphy.

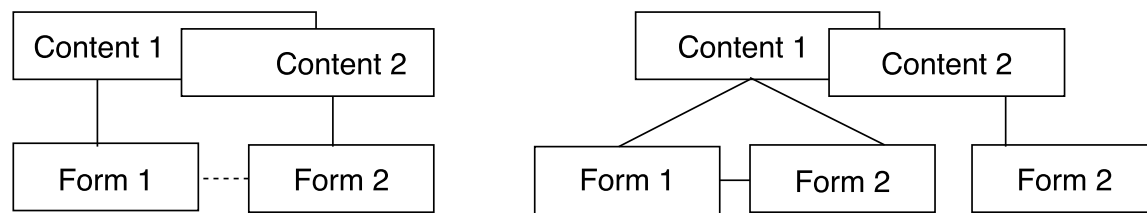


Figure 3: Non-Standard Allomorphy.

The leftmost graph presents two semantic contents with a strong overlap and an equally strong split. These contents correspond to two linguistic forms that are related to each other, but in a defective way: for example, their phonological similarity might be lost, or distributional criterion can be violated by some exceptions or instances of overlap, or the automatic rule that relates the forms can be restricted to a subpart of the lexicon. The leftmost graph can also represent a case of what was originally a single morpheme which underwent partial split. As a result, its allomorphs developed partial semantic differences. We know that such cases exist. It is common knowledge that allomorphs often develop a contrast in terms of register or meaning. For example, consider the data from Dutch discussed by Booij:

“In medieval Dutch a phonological process took place in which the sequence /də/ was deleted both word-initially and in intervocalic positions. It was a process subject to lexical diffusion, that is, it affected a number of words, one by one, and then it stopped. Consequently, there are a number of words in Dutch with two forms, one with, and one without /də/. **As may be expected, the two allomorphs often got different meanings, or at least a stylistic differentiation** (the *de*-less allomorph is more informal, or the allomorph with *de* archaic).” (Booij 1995: 88; boldfaced by myself – A.E.)

- a. Word-initially in the words:
 - sne**de*** ‘cut’ vs. *snee* ‘cut (slice of bread)’;
 - la**de*** ‘drawer’ vs. *la*;
 - wei**de*** ‘meadow’ vs. *wei*;
 - kou**de*** ‘cold’ vs. *kou*
- b. Intervocalically:
 - broe**der*** ‘brother; male nurse; form of address in church’ vs. *broer* ‘brother’;
 - moe**de**r* ‘mother’ vs. *moer* ‘female animal, female screw’.

The rightmost graph in Figure 3 presents a different case of Non-Standard Allomorphy: the two semantic contents overlap to a smaller extent and diverge to a greater degree. These contents are paired with two linguistic forms, where some instances of Form 2 are clearly phonologically related to Form 1, but other occurrences of Form 2 appear without any clear reason and seem to be associated with partially different content.

In this light, in the post-structuralist era, it is useful to explore the non-trivial asymmetries of form and meaning in order to deepen our knowledge of the nature of linguistic phenomena.

2.4 Types of allomorph conditioning

An allomorph is a conditioned variant of a morpheme (Trask 1996: 85). Each allomorph is associated with a particular environment, or context, and types of such contexts define the types of allomorph conditioning and types of allomorphic relations. In this section I am concerned with the variety of types of allomorph conditioning – from the types widely accepted in the linguistic community to those types that are marginal and underinvestigated. In particular, I will draw attention to semantic conditioning and the role of semantics in allomorphy in general.

Linguists differ in what types of allomorph conditioning they recognize (or accept). The scope of recognized types of allomorph conditioning is very theory-dependent and data-dependent, as it is often restricted by the theoretical framework and by the data that linguists need to account for. Moreover, the terminology employed in this domain is neither uniform nor consistent from one reference work to another.

Many linguists recognize only **phonological** and **morphological** (also termed **grammatical**) conditioning of allomorphs. For example, Lyons (1968: 184-186) and Fábregas & Scalise (2012: 15) discuss only allomorphy conditioned by phonological and grammatical, or morphological, factors. In addition to these two types, it is common to recognize also **lexical** conditioning of allomorphs (Bauer 1988: 240, 2004: 15-20; Trask 1996: 85; Haspelmath 2002: 29-30). Mascaró (2007: 715) goes further and states on the basis of his data that “[t]he conditions that determine the choice of one allomorph rather than another can be **semantic, syntactic**, morphological, or phonological, and in many cases they are highly idiosyncratic.” Bauer (1988: 15) considers also written language and distinguishes **orthographic** conditioning of allomorphs (e.g. English *come* and *com-* (as in *coming*). Rarely one can come across such terms as **historical** conditioning or **mixed** conditioning of allomorphs, although these also exist (Booij 2005: 170; 2010: 252). It is worth taking a closer look at the major types.

2.4.1 Phonological conditioning

Phonological conditioning of allomorphs is recognized by all sources, it is the most studied and most thoroughly described type of allomorphy. By definition, phonologically conditioned allomorphs are triggered by particular properties of phonological context. Here we should distinguish between phonologically conditioned distribution of allomorphs and phonological relation of their shapes. These two things often characterize the same data but do not necessarily coexist. The choice of allomorph can be phonologically conditioned and serve the ease of pronunciation, but the allomorphs themselves can at the same time be unrelated phonologically – not related by means of

an automatic phonological rule (phonologically conditioned suppletive allomorphy, cf. Paster 2009).

2.4.2 Morphological conditioning

The second type of allomorph conditioning is usually called morphological or, more broadly, grammatical conditioning. This type of conditioning takes place when the choice of allomorph is triggered by a grammatical or morphological factor – such as “gender, conjugation or declension, or the presence of a particular type of affix” (Bauer 1988: 240). I provide examples of these subtypes below.

Conditioning by a grammatical category can be illustrated with German adjectives: they change their form (the suffix) depending on the gender of the noun they modify, as we see in the examples *ein gross-er Wagen* ‘a big car_{MASCULINE}’, *ein gross-es Haus* ‘a big house_{NEUTER}’, *eine gross-e Feder* ‘a big feather_{FEMININE}’ (Bauer 1988: 14). Gender is a grammatical property of nouns and therefore the choice of the suffix on adjectives is conditioned grammatically.

Conditioning by declension class is the case of Latin ablative plural markers: first and second declension nouns attach *-i:s*, whereas nouns of other declensions take *-(i)bus* (Bauer 1988: 240).

Allomorphy conditioned by the presence of a particular grammatical morpheme takes place in the alternation of English stems in singular and plural: e.g. in *knife ~ knives* the stem allomorphy is determined by the plural marker, whereas in *knife ~ knife’s* the possessive suffix, phonologically identical to the plural marker, fails to trigger the stem alternation (Trask 1996: 162). Another example comes from different shapes of the English suffix *-able*: compare *enjoy-able* vs. *enjoy-abil-ity*, where the presence of *-ity* arguably triggers the shift of *-able* into *-ible* (Fábregas & Scalise 2012: 15).

According to Trask (1996: 162), morphological conditioning of allomorphs also includes those cases when the distribution of morpheme variants is affected by properties of “the grammatical structure of a word”, such as morpheme boundaries. For example, in Scottish English, Aitken’s Law produces a short vowel in the closed syllable in *brood*, whereas in *brew-ed* the morpheme boundary prevents shortening, and this form retains the long vowel of the base *brew* (ibid: 162).

A coherent and systematic account for morphophonological variation is proposed in terms of Cognitive Grammar in Nessel 2008. In this monograph, it is persuasively shown that a set of *softening* and *truncation* alternations in Russian verbs serves to differentiate between the present and past tense stems (ibid: 6). “Truncation” alternations refer to relations between a longer and a shorter allomorph of a stem, as in *piš-et* 3.SG.PRESENT ‘s/he writes’ vs. *pisa-l* SG.MASC.PAST ‘he wrote’ < *pisat’* INF ‘write’. In addition, “softening” alternations manipulate stem-final consonants, as in the case of *s ~ š* in the same pair of forms, and contribute to marking the contrast between the present and past stems of a verb. Nessel suggests that both phenomena carry a semiotic function, as they mark grammatical meaning. Contrary to formal accounts, Nessel analyzes stem allomorphy in the Russian verb system not in terms of abstract underlying representations and formal rules, but in terms of schemas and surface-based generalizations. Crucially, this work shows that allomorphic variation is motivated by morphological meaning on a regular basis and that this interaction of phonology and morphology can be captured in terms of a concrete cognitive model.

2.4.3 Lexical conditioning

Note that phonological and morphological conditioning are described by rules that apply regularly or even automatically in many words. If an allomorph is observed only in a single lexeme or a few lexemes, or if an allomorph cannot be explained by the phonological or morphological characteristics of its context, then it is assumed that such an allomorph is lexically conditioned (Bauer 1988: 14). This type of conditioning is often illustrated by irregular plural markers in English nouns like *men*, *sheep*, *children*, *cacti* (Trask 1996: 203; cf. more discussion in 2.2.3). Similarly, ablaut as a marker of past tense in English verbal forms is considered lexically conditioned, as we encounter it in *hang* > *hung*, but not in *bang* > *banged* (Bauer 1988: 237). The use of umlaut in German is also largely lexicalized and depends on individual stems (Féry 1994: 5; Booij 2005: 169).

Lexical conditioning implies that allomorphs are analyzed as exceptional (Trask 1996: 203) and irregular (Lyons 1968: 186). Therefore, it is assumed that lexically conditioned allomorphs have to be stored in the lexicon as a part of a particular lexeme and “must be learned individually for each case” (Haspelmath 2002: 30).

2.4.4 Semantic conditioning

A problem occurs when the number of exceptions is too large. For example, Bye (2014: 32) in the study of plural markers and forms in Burushaski, a language isolate spoken in northwestern Pakistan, discovers that the most striking feature of this data is the great number of exceptions. According to his calculations, “[o]ver a quarter (26%) of the nouns in the sample have an unpredictable choice of plural suffix”. In this case, to simply accept that all these exceptional plural forms make an arbitrary list and should be assigned to the “lawless” lexicon means to fail to find any pattern in this data. Instead of calling this distribution a lexically conditioned allomorphy, Bye (2014) provides a thorough analysis which reveals that the list of exceptions contains an internally motivated system with a robust semantic pattern.

The allomorphy of plural markers in Burushaski makes a distinction between human and non-human classes. Bye (2014: 17) reports that “[t]he plural suffixes {-daru}, {-štaru}, {-tiŋ}, {-kón}, and {-bák} may only appear with nouns in the human class.” Other distinctions concern semantic subclasses of nouns. In particular, the nouns that take the allomorph {-štaru} (or {-šteru}) all refer to “blood relatives in the parental generation or older”: e.g. *bap* ‘grandfather’, *nzu* ‘aunt’. Another group of nouns forms the plural by the suffix {-daru}, and most of these nouns “refer to close family members in the same generation as or younger than ego” (ibid: 17). Furthermore, the suffix {-tiŋ} “is used with nouns referring to individuals that in one sense or another require ‘handling with care’, either because they are deemed especially worthy of consideration, or because they spell trouble” – royal persons, individuals with special spiritual gifts, supernatural creatures like angels and fairies, people “whose jobs entail working with sharp implements” like a knife-grinder, as well as people that have “socially disruptive character traits” like a crazy person or womanizer (Bye 2014: 17).

Bye (2014: 44) concludes that allomorphy of the plural marker in Burushaski is semantically conditioned: “the selection of certain plural allomorphs is conditioned by the semantics or pragmatics rather than the phonology of the stem”. Moreover, “where semantic/pragmatic and phonological conditions conflict, the semantic/pragmatic

conditions have priority by virtue of what we have called the Core Semantic Override Principle” (ibid: 44).

Interestingly, Bye observes that these classes of words share not only semantics but also certain phonological or morphological properties. Therefore, Bye proposes a concept that he calls the *morpholexicon* (Bye 2014: 44). The idea of the *morpholexicon* and the conclusion that the exceptionality of Burushaski plural formation is patterned in terms of semantics bring this thorough empirical study to a fundamental amendment of linguistic modularity: from this data it appears that the modules of language, viewed in the generative framework and Optimality Theory as independent from one another, are in fact not that autonomous. Bye argues that “[i]n the current approach, selection is governed by the morpholexicon, where the boundaries between autonomous syntactic, semantic and phonological modules break down. Allomorph selection under the current proposal calls not just on morphosyntactic features, such as gender, but calls directly on semantic and phonological information as well” (ibid: 15).

A similar case of semantic conditioning is discussed by Haspelmath (2002: 30), although it is termed lexical conditioning. Haspelmath points out cases where the selection of an affixal allomorph “is dependent on semantic properties of the base.” Haspelmath illustrates this type of conditioning with Persian plural marking. Similarly to Burushaski, the distribution of allomorphs in Persian is determined by semantic subclasses of words: human nouns attach the allomorph *-an*, whereas non-human nouns use the allomorph *-ha* (Mahootian 1997: 190):

Human nouns: <i>-an</i>	Non-human nouns: <i>-ha</i>
<i>mærd</i> ‘man’ – <i>mærd-an</i> ‘men’	<i>gorbe</i> ‘cat’ – <i>gorbe-ha</i> ‘cats’
<i>geda</i> ‘beggar’ – <i>geday-an</i> ‘beggars’	<i>etterfaq</i> ‘incident’ – <i>etterfaq-ha</i> ‘incidents’

Interestingly, Persian and Burushaski exhibit a similar pattern – both languages semantically condition allomorphs of plural markers, and in both languages the semantic distinction concerns nouns that refer to humans vs. non-humans.

Apart from these two instances, semantic conditioning of allomorphs is largely overlooked in the literature. The first reason might be that semantic conditioning is a debatable and “unsafe” zone of argumentation because it implies that allomorphs are semantically different, and this violates the criterion of semantic identity through which allomorphs are established in the first place. Second, modern discussion of allomorphy is mostly centered around phonology and the issues of variation of formal shapes and distribution of variants are discussed at the expense of semantic issues. Underestimation of the role of semantics in allomorphic relationships is rooted in the long-term tradition initiated by Bloomfield (cf. section 2.6.3.1 on Bloomfield’s version of structuralism).

In Optimality Theory (OT) there has been a number of proposals which recognize various semantic effects in allomorph selection and aim at capturing them in the analysis. Mostly, they propose special constraints which should capture the encountered semantic effects of allomorph selection. For example, Bonet et al. (2007) discuss an additional property of the lexical representation of allomorphs – lexical ordering of allomorphs. In particular, the authors propose the constraint RESPECT. This constraint “ensures compliance with idiosyncratic lexical specifications, which further interacts with allomorph selection” (ibid: 903). Another study, Bye 2007, argues that selection among suppletive allomorphs “reflects language-specific combinatorial requirements (subcategorization frames)” and therefore introduces a constraint MORPHOLEXICAL

CONTROL which checks “whether the relevant language-specific combinatorial restrictions are satisfied or not” (ibid: 63, 65). Regarding the types of conditioning, Bye points out that “[t]he distribution of suppletive allomorphs may be governed by a variety of contextual factors, including the semantics, phonology, and exact identity of the selecting lexical item”. Similarly, Mascaró (2007) proposes that lexical representation of allomorphs constitutes a partially ordered set and the preference for a particular allomorph over its competitor is maintained by the faithfulness constraint PRIORITY (ibid: 718). There also appear recent OT studies of exceptionality which show that this phenomenon is patterned, the “arbitrary” list of words where allomorphs are lexically conditioned is in fact organized, and the choice of allomorph is motivated (Pater 2004, 2007, 2010).

While formal phonology encounters the semantic component of phonological phenomena through the analysis of overwhelming exceptions, this is not the only path. In Cognitive Linguistics, it is natural to search for semantic motivation of linguistic phenomena, including variation in the phonological shapes of morphemes.

A peculiar example of semantically conditioned allomorphy described in terms of Cognitive Linguistics comes from Modern Russian verbs that are formed by two rival semelfactive markers – the prefix *s-* and the suffix *-nu-*. Dickey & Janda (2009) argue that these affixes behave as suppletive allomorphs in the formation of verbal predicates meaning ‘do something once’: e.g. *čixnut* ‘sneeze once’ from *čixat* ‘sneeze’; *sxodit* ‘go someplace and come back once’ from *xodit* ‘walk’. The choice between these two markers of semelfactive Aktionsart is governed by morphological verb classes and by semantic classes of base verbs, which are defined independently in the Russian National Corpus (ibid: 241). In particular, verbal bases that denote sounds, speech acts, and physical impact or movements prioritize the *-nu-* suffix (e.g. *kvakat* ‘croak’ > *kvaknut* ‘croak once’; *kričat* ‘shout’ > *kriknut* ‘shout one thing’; *tolkat* ‘push’ > *tolknut* ‘push once’; *maxat* ‘wave’ > *maxnut* ‘wave once’), whereas motion verbs and bases that refer to behavior tend to form semelfactives by means of the prefix *s-* (e.g. *glupit* ‘act stupid’ > *sglupit* ‘act stupid once’). Dickey & Janda found that both robust patterns of distribution (morphological and semantic) tolerate certain degrees of overlap, and there are verbs that attach both semelfactive markers at the same time (e.g. *s-trux-nu-t* ‘do one cowardly thing’ < *trusit* ‘behave cowardly’) (ibid: 242). The allomorphy hypothesis pursued in Dickey & Janda 2009 was later tested and confirmed experimentally in Makarova 2009. Since morphological and semantic classes of base verbs in Russian are not independent from each other, these studies call into question the generative assumption about the modularity of morphological and semantic components of grammar (recall a similar theoretical outcome of the semantic conditioning analysis in Bye 2014).

2.4.5 Historical conditioning

It is worth mentioning that phonological, morphological, and lexical conditioning of allomorphs are related. Booij (1997: 37) mentions that in some cases phonologically conditioned allomorphy becomes morphologically conditioned allomorphy. This shift can take place over time, when stem allomorphy “reflects the phonological history of a language, but then is no longer transparent synchronically.” An example is the Vowel Shift in English which was once an active process but is no longer productive (cf. also an example from Frisian in Booij 1997).

Similarly, when the phonological processes that conditioned alternation of morpheme variants are no longer active, allomorphy becomes lexicalized and the selection of the right allomorph should be restated “in terms of morphological structure and analogical patterns” (Booij 2010: 252). Such allomorphy is also **historically** determined (Booij 2005: 170).

Allomorphy as a reflex of history can be found in alternation of English stems in the pairs *produce* – *product-ion* and *reduce* – *reduct-ion*, where the derivatives use *duct* as a base, which is the participial stem form of the Latin verb *ducere* (Booij 2005: 170). Likewise, the stem allomorphy in the English words *drama* vs. *dramat-ic*, *dramat-ist* and *Plato* vs. *platon-ist*, *platon-ism* exists due to the historical reflexes of these originally Greek words: “the long form is the underlying form, but the stem-final consonant was dropped in NOM.SG forms” (ibid: 170). As stated by Booij (2005: 170), “[t]he effect for the present day English is that the long form is to be used before suffixes of the non-native learned stratum, whereas the short form is to be used before native, Germanic suffixes, and with prefixes. For instance, the plural form of *drama* is *dramas*, not *dramats*”.

Summing up, examination of the literature reveals that the notion of allomorphy embraces a wide range of rather heterogeneous linguistic phenomena. Conditioning of allomorphs can refer to a wide scope of factors – phonological, morphological, lexical, semantic, and historical. The types of factors that condition allomorph selection have different weights, because they indicate different degrees of closeness between morpheme variants.

2.5 Two additional criteria for allomorphic status

Apart from two standard definitional criteria for allomorphic relationship – identical meaning and complementary distribution, different sources mention additional properties that support allomorphic status. In this section, I address two such characteristics – formal phonological similarity and a common historical source. Neither of these two properties is considered obligatory for allomorphs, yet different linguistic traditions value them differently.

2.5.1 Formal phonological similarity

Many sources mention phonological similarity as an expected or desirable characteristic of proper allomorphs. Matthews (1974: 114) posits that “the alternants of a morpheme regularly have much of their phonological make-up in common”. Bauer (2004: 15) claims that “according to some views of the morpheme, the allomorphs also have to be phonologically similar to each other”. However, most sources suggest that phonological similarity is not an obligatory property of allomorphs.

In European and American linguistic tradition it is common to recognize the phenomenon of **suppletive allomorphy**, that is the relationship between variants of a morpheme that are “not at all similar in pronunciation” and are not synchronically related by any automatic regular phonological rule (Haspelmath 2002: 28; Spencer 1991: 8). For example the suffixes of the English past participle *-ed* and *-en* are not phonologically similar and are analyzed as suppletive allomorphs (Haspelmath 2002: 28).

The problem is that there are no explicit criteria for rating formal shapes as similar or not. This has been pointed out by many linguists. Bauer (2004: 72) mentions

that “there is no well-defined cut-off point on the scale of phonological relatedness”: whereas the plural marker in *cat-s* and *dish-es* are clearly similar, the stems in *bear* and *birth* are less so. Likewise, Haspelmath (2002: 29) states that “it is not always easy to decide whether an alternation is phonological or suppletive”: the roots of English verbal forms *buy/bought* and *catch/caught* “are not as radically different as *go/went-t*, but they are not similar enough to be described by phonological rules either”. In order to capture different degrees of formal phonological similarity and relatedness of allomorphs, Haspelmath uses the terms **weak suppletion** (as in *buy/bought*) and **strong suppletion** (as in *go/went*, *good/better*) (ibid: 29). Similarly, Spencer (1991: 8) proposes parallel terms **partial suppletion** and **total suppletion**. The former describes cases of some phonological connection between two forms, while the latter refers to allomorphs that bear absolutely no phonological resemblance.

Suppletive allomorphs have a special status in many respects. Usually, suppletive allomorphs are conditioned morphologically (Haspelmath 2002: 29). Suppletive allomorphy can also be conditioned by phonological properties of the context, but the allomorph selection in such a case is not necessarily optimizing – it does not have to facilitate pronunciation (Paster 2006; Bye 2007). Bye 2007 therefore suggests that allomorphy is rather a selection, whereas phonological optimization occurs due to historical or coincidental factors.

Whereas most phonologists include suppletion in the scope of allomorphic relationships, some morphologists do not do so. For example, Fábregas & Scalise (2012: 16) view allomorphy and suppletion as different phenomena: “[i]t is generally assumed that the different allomorphs of a morpheme should have some phonological relation, since, in the traditional notion of allomorph, they are variants of one and the same lexical unit. For this reason, it is expected that one variant can be related to the other through some change in its shape, even if this change does not generally take place in that language. <...> In cases of suppletion it is generally believed that we are talking about two (or more) different lexical items that happen to share the same paradigm; as they are different lexical items, we do not expect that their shapes will be related by a phonological rule of any kind”. In addition, Fábregas & Scalise (2012: 17) mention that “in practice the boundaries between allomorphy and suppletion are fuzzy”, especially in the case of lexicalized allomorphs that exist due to some historical reflexes but still exhibit some formal resemblance.

Likewise, Booij (2005: 172) views the formal similarity of rival forms as a criterion for their allomorphic status: “When the degree of phonological similarity between competing affixes is considerable, they may be considered allomorphs of one morpheme. When the similarity is less, it is better to regard them as distinct morphemes, though in a relationship that resembles the suppletive one between the stem-forms *good* and *bet-* of the lexeme GOOD.”

Phonological similarity is traditionally viewed as an obligatory property of allomorphs in the Russian linguistic tradition (this regards mostly linguistics of the Soviet time but also many modern Russian linguists, although some adopt an alternative model, e.g. Mel’čuk 2001: v.4, 248-250). The difference between American / European understanding of allomorphy and Russian linguistic use of this term is explicitly addressed in the Linguistic Encyclopedic dictionary by Jarceva (1990) and some recent articles (Gerd 2004). In particular, Bulygina and Krylov (in Jarceva 1990: 311) state that American linguists stand on the ground of identical function and analyze suppletive variants as allomorphs of a single morpheme. By contrast, in Soviet linguistic tradition, if

there is no formal resemblance of rival forms, it is customary to assign them to distinct but synonymous morphemes.¹²

The difference between American and Russian linguistic traditions with regard to the notion of allomorphy is a topic of explicit discussion in Gerd 2004. Gerd attributes the irrelevance of phonological similarity for the American understanding of allomorphy to the descriptivist tradition which focuses on the distributional properties and prioritizes economic accounts with a minimal number of postulated abstract linguistic items like morphemes. In search for the most explicit definition of allomorphy that reflects its understanding in the Russian linguistic tradition, Gerd refers to a chapter "Introduction to morphemics" of the Russian Grammar (Švedova et al. 1970). The authors of this chapter, V.V. Lopatin and I.S. Uluxanov, specialists of Russian word-formation, define allomorphs as morphs that are characterized by 1) identical meaning and 2) formal (phonological) closeness, that is partial identity of the set of phonemes and their order.¹³ Note that this definition does not mention that allomorphs must be distributed complementarily.

In this account, phonological closeness is discussed in greater detail. For example, one phoneme of a polyphonemic morph can be replaced with another phoneme or several phonemes in another morph: e.g. the Russian suffixes *-ščik* ~ *-čik*, nominal roots *-rabot-* ~ *-rabat-*, verbal stems *-my-* ~ *-moj-*. Another option is when one morph is a part of another morph, as in the Russian suffixes *-iva-* ~ *-va-* ~ *-a-* and *-sk-* ~ *-esk-*. Yet another possibility of formal resemblance is the case when one of the middle phonemes of one morph is absent in the shape of the other morph: *-son-* ~ *-sn-* (Švedova et al. 1970: §31).

According to this definition of allomorphy, suppletion contradicts the notion of morpheme identity and is logically excluded from allomorphy. However, not all Russian linguists share this view. For example, Mel'čuk (2001: v.4, 248-249) discusses suppletive allomorphs as variants of a single morpheme.

Gerd (2004: 210) suggests that the issues of formal closeness, partial phonological identity of morphs, and especially its limits are highly controversial and therefore the notion of allomorphy is one of the most unclear linguistic notions. Gerd concludes that modern linguistics should rather eliminate the notion of allomorphy altogether and instead analyze units of morphophonological segmentation – morphonemes (ibid: 210). Because morphonemes are the units of phonological formal representations regardless of their semantics and morpheme identity, morphophonology takes a step back to *morphs* and in fact avoids a morphological analysis. Instead of solving the problems of unclear semantic or functional relatedness of formal shapes, this approach tries to

¹² The original text reads as follows: "Priznanie dvux različnyx morfov allomorfami odnoj morfemy svjazano takže s tem, kakoe soderžanie vkladyvaetsja v ponjatie morfemy. Tak, različnym obrazom traktujutsja slučai suppletivizma, naprimer morf *i-* vo franc. *ira* 'pojdet' i morf *all-* v *allos* 'pojdem'. Amerikanskije lingvisty obyčno isxodjat v takix slučajax iz toždestva funkcii i govorjat ob allomorfax, togda kak v sovetskom i v evropejskom jazykoznanii bolee rasprostranena drugaja točka zrenija: pri otsutstvii formal'nogo sxodstva meždu morfami govorjat o raznyx, xotja i sinonimičnyx morfemax, i, sledovatel'no, dannye morfy ne sčitajutsja allomorfami" (Jarceva 1990: 311).

¹³ The original text reads as follows: "Allomorfy odnoj morfemy – èto morfy, kotorye xarakterizujutsja sledujuščimi priznakami: 1) dannye morfy imejut toždestvennoe značenie; 2) oni obladajut formal'noj (fonematičeskoj) blizost'ju, pod kotoroj ponimaetsja častičnaja toždestvennost' sostava fonem i ix porjadka" (Švedova et al. 1970: §31).

discredit the very attempt to do so and is concerned with bare variation of formal shapes.

We can observe similar attempts in other theories to dissociate form and meaning in a linguistic analysis and address only the former. Consider, for example, Aronoff (1994) who proposes the notion of *morphome* as “a cluster of signs which are in complementary distribution, and thus may mean the same as morpheme in one of its many uses <...> However, while the meaning is important for the identification of a morpheme, it is not for the morphome” (Bauer 2004: 3-74) For instance, *-s* in *besides* and *-s* in the verbal *goes* belong to the same morphome even though they are two different morphemes.

Another version of such gravitation away from the issues of morpheme identity and allomorphy is the Separation Hypothesis of Beard (1995). According to Beard (1987: 1, 18; 1995: 44), only a lexeme should be considered a true linguistic sign – a pairing of form and meaning. In this regard, morphemes should be separated from lexemes: they lie outside the lexicon and are “purely phonological operations independent of semantic (grammatical) operations that they mark.” Because languages have homonymous morphemes on the one hand and morphological syncretism on the other hand, Beard (1987: 18) concludes that the meaning, or function, is not associated with the morpheme’s form directly but is rather associated with its context. Therefore, Beard draws a distinction between affixation as modification of the formal shape of the base and derivational operations as modification of its meaning.

To sum up, individual linguists and linguistic traditions differ in judging how relevant formal similarity is for allomorphic relationships. The absence of clear criteria for evaluating phonological closeness and the phenomenon of suppletion undermine the standard theory and make some linguists deny the validity of allomorphy as a notion. Suppletion is included in allomorphy by some linguists, but excluded by others. Suppletive allomorphs remain the most debatable members of the category even for those linguists who grant them allomorphic status. This suggests that phonological resemblance is an important property which defines best (or ideal) exemplars of allomorphy.

2.5.2 Common historical source

Etymological kinship is regarded as a criterion for allomorphy by those linguists who exclude suppletion from the scope of allomorphic relations. For example, Fábregas & Scalise (2012: 17) mention a shared historical origin as an expected property of allomorphs. However, they admit that there is no guarantee that historically related forms will necessarily be allomorphs from a synchronic point of view: “this criterion, useful as it is in lexicography and other linguistic disciplines, does not give clear results in the study of contemporary morphology.” For example, “in a Romance language, the relationship between a word borrowed from Latin and the result of the historical evolution of that same Latin word may be judged differently by speakers depending on their historical knowledge of this relationship” (ibid: 17).

2.6 Allomorphy: a term with Structuralist “baggage”

The theoretical notion of **allomorphy** belongs to a core set of notions employed in modern linguistics. It is used in various frameworks and theories of language. Judging from this fact one might conclude that this notion is theory-neutral. However, I argue

that this is not the case. The term *allomorphy* was coined within the Structuralist approach to language and was inherited by modern linguistics together with a whole set of Structuralist theoretical assumptions about language and its organization. I further suggest that allomorphy is a concept with Structuralist “baggage”. In this section I aim to look back at the theoretical context in which allomorphy entered the linguistic vocabulary and discuss the crucial implications and premises that this notion still carries.

How do we know that allomorphy came with Structuralism? Because the idea of allomorphy and the term itself evolved during a fruitful discussion on the pages of *Language* in the 1940s. The participants of the discussion were American descriptivists¹⁴ who worked on the development of a universal procedure for a morphemic analysis. In what follows I highlight three major articles that became milestones in the theory of allomorphy and played a crucial role in the history of this notion. In discussing these articles, I will set aside a number of specific issues which concern interfixes, empty morphs, and types of conditioning. Instead, I focus on two aspects of this discussion – the clarification of terminology and the criteria of allomorphic relations.

2.6.1 Descriptivists

2.6.1.1 Harris 1942

In 1942, Zellig Harris in his article “Morpheme alternants in linguistic analysis” makes a crucial distinction between two notions – ‘**morpheme unit**’ and ‘**morpheme alternant**’, what today is known as *allomorph*. Harris explicitly suggests that “the linguistic levels are organized as analogous to each other”. Accordingly, he draws a parallel between phonology and morphology: “if a phoneme is a class of allophones in complementary distribution, then it is logical that a morpheme unit is a group of alternants with the same meaning and complementary distribution” (Harris 1942: 171). In other words, Harris applies to allomorphy the same criteria that were employed for establishing allophony, despite the common knowledge that a morpheme is a linguistic unit essentially different from a phoneme. A morpheme is a minimal meaningful linguistic sign. It has a semantic content that a phoneme lacks. However, this difference between morphemes and phonemes does not prevent Harris from saying that “we can arrange alternants into units in exactly the same manner as we arrange sound types (positional variants) into phonemes” (ibid: 179). A question arises in this connection: why is the method of establishing phonemes explicitly viewed as a model that one should apply to other linguistic levels? The logic that motivates this view is the following: the Structuralist theory of phonemes, alternations, and features is seen as “rigorous” (ibid: 170). Being rigorous is certainly an advantage in a Structuralist approach to language, because Structuralism strives to be scientific. Therefore, the ambition is “to suggest a technique for determining the morphemes of a language, as rigorous as the method used now for finding its phonemes” (ibid: 170). In particular, the assumption about the analogous organization of phonology and morphology justifies the extrapolation of the two allophonic criteria to allomorphy.

¹⁴ I will use the term *descriptivists* to refer to post-Bloomfieldian American structuralists before the Generativist approach appeared in the late 1950s (Crystal 1991).

2.6.1.2 Hockett 1947

Five years later, in 1947, *Language* publishes a follow-up article titled “Problems of Morphemic Analysis” by Charles Hockett. The objective of this work is to revise and optimize the procedure of morphemic segmentation proposed by Harris. With regard to terminology, Hockett wants to replace the term *alternant* with the term *morph*. His argument is twofold: first, *morph* is shorter – it “eliminates the lengthy expressions ‘morpheme alternant’ and ‘morpheme unit’”; second, the term *morph* suggests the key analogy of linguistic concepts – **(allo)phone : phoneme = morph : morpheme** (Hockett 1947: 332).

Sharing with Harris the assumption about linguistic levels and their units, Hockett suggests a fundamental change. He proposes an important amendment to the distributional criterion. Instead of the requirement for complementary distribution of forms, Hockett proposes a **non-contrastive distribution**. The latter includes two possibilities: 1) complementary distribution, and 2) partial complementation. Hockett also recognizes environments where both alternants can occur in **free variation**, that is without bearing any semantic contrast (ibid: 328). This observation suggests that overlap in the distribution of alternants is possible, but their content should be the same. Hockett illustrates this statement with the forms of the English personal pronoun *I* and *me*. He argues that in some contexts these forms occur in free variation, for instance in the expressions *It’s me* vs. *It’s I*, and *for you and me* vs. *for you and I* (cf. subsection 2.2.4).

The difference between complementary distribution and non-contrastive distribution is crucial here. I suggest that this difference concerns the chosen “scope of definition” which can be negative (reductionist) or positive (inclusive). For example, consider two alternative definitions of a phoneme. A negative definition of a phoneme reduces the number of its various characteristics to a minimal set of features – that is only those features that are in contrast with other phonemes in a language. A positive definition of a phoneme introduces a different view, which is inclusive, as it includes a maximal (or at least most comprehensive) set of features – all features that are characteristic of a particular phoneme.

Note that the reductionist approach based on a negative type of definition might seem more scientific because it counts only those features that are contrastive. These features are accessible to us directly through the behavior of linguistic signs and the comparison of their overt shapes.

An inclusive approach arises from a positive type of definition and takes a broader view – it takes into account not only the formal behavior of forms, but also the content of these forms. This approach aims to elicit a comprehensive and exhaustive combination of all features what make a sign unique.

These two approaches to a phoneme were articulated in the history of phonology by two opposing phonological schools – the Moscow phonological school (the reductionist approach) and the Petersburg phonological school (the inclusive approach) in the 20th c.

Beyond phonology, the two approaches to selection of definitional characteristics of a linguistic phenomenon motivate fundamental differences in linguistic frameworks. In this light, Cognitive Linguistics represents an inclusive approach. This framework does not find it useful to reduce the linguistic content of a unit to formal oppositions. On the contrary, Cognitive Linguistics is concerned with both the internal content and how it is organized in linguistic phenomena. The opposition between “negative” and

“positive” definitions is discussed by Nida (1948) who advocates a different approach to allomorphy.

2.6.1.3 Nida 1948

In 1948, one year after the Hockett’s article, *Language* published a related paper “Identification of morphemes” written by Eugene Nida – a linguist who is now more famous for his pioneering methodology for Bible translation. As a linguist he worked on more than eighty languages in more than thirty countries. In this article Nida first introduces the term *allomorph* to the linguistic field by stating that “**morphemic alternants** can conveniently be called **allomorphs**”¹⁵ (Nida 1948: 420). Nida recycles the term *morph* and suggests that “[i]n the process of analyzing a language there might be occasion to use the term **morph** to designate a structural unit which had not as yet been assigned to any morpheme” (ibid: 420). Thus, the term *morph* is given the role of an intermediate middle layer in the procedure of assigning allomorphs to morphemes. The term *morph* is only concerned with form regardless of its meaning, whereas *allomorph* relates form to meaning, assigning a morph to the structural meaningful unit – the morpheme. In this sense, the term *morph* is less informative than *allomorph* and this might be the reason why this term has today almost gone out of use (Haspelmath 2002: 29)¹⁶.

The major contribution made by Nida to the theory of Bible translation was a meaning-based approach (Nida 1964, 1975, 1982, Neff 2002). Nida proposed that in translation one should first of all focus on translating thought-for-thought rather than word-for-word (the idea of *dynamic equivalence*).

Likewise, meaning matters in Nida’s theory of allomorphy. In particular, Nida is very careful with the criterion of identical meaning and instead speaks about sub-

¹⁵ Here it is important to do justice to Nida’s contribution to linguistic terminology, since not all reference works on the history of linguistics acknowledge him as the inventor of this term. For example, Matthews (1993: 81) claims that the term *allomorph* was invented by Hockett: “Hockett (1947) invented the terms ‘morph’ and ‘allomorph’ by explicit analogy with the ‘phone’ – a term that he had himself introduced a few years earlier (1942) – and the ‘allophone’.” This last term had been introduced by Whorf in an unpublished paper at the end of the 1930s.” However, Hockett never mentioned the term ‘allomorph’ in this article. Itkonen (2005: 85) also refers to Hockett: “[a]nd Hockett adds [cf. Hockett 1968] that it was he himself who coined the term ‘allomorph’, on the analogy of ‘allophone’. However, the first true occurrence of the linguistic term *allomorph* came in Nida 1948.

As Luschützky (2000: 452) points out, the term *allomorphy* (along with a synonymous term *allotropy*) was used long before linguistics in physical chemistry and crystallography in order to denote chemical substances that have similar crystal forms like diamond and graphite as forms of carbon: “In diesem Zusammenhang ist erwähnenswert, daß *Allotropie* und *Allomorphie* in der physikalischen Chemie und in der Kristallografie seit langem als Synonyme für chemisch gleich geartete Substanzen verwendet werden, die keinerlei Ähnlichkeit in ihrer Kristallform aufweisen (z. B. Kohlenstoff als Diamant oder Graphit)”. So, in linguistics the term *allomorphy* was arguably reinvented.

¹⁶ In particular, Haspelmath (2002: 31) states that “[m]any morphology textbooks introduce the distinction between *morph* and *morpheme* in this sense, but in actual practice linguists rarely use the term *morph*, and the term *morpheme* is often used in a concrete sense as well.” Therefore, Haspelmath himself adopts the common practice for morphologists and eliminates the term *morph* from the remainder of his book. Bauer (1988: 17) also point out that “many American sources do not distinguish between ‘morph’ and ‘allomorph’.”

morphemic differences in the semantic content of allomorphs. This especially matters when allomorphs are conditioned not phonologically but rather by other factors, for example morphology: “[m]orpheme alternants whose distribution is not phonologically definable exhibit **sub-morphemic differences in meaning**” (ibid: 431).

In discussion of semantics and distribution of linguistic forms, Nida draws attention to Bloomfield’s ideas: “[n]ote also Bloomfield’s assumption that there are ‘no actual synonyms’ – no items that are different in form but absolutely identical in meaning. If it is true that ‘selection of forms contributes a factor of meaning’, then the different selection (distribution) of allomorphs implies that they have different meanings” (ibid: 431). Nida concludes that “[f]rom the difference in their distribution they acquire a certain difference in meaning” (ibid: 433).

Nida shares the assumption about the analogical structure of linguistic levels: “accordingly, **allomorphs** are related to **morphemes** as **allophones** are related to **phonemes**” (ibid: 420). However, Nida points out that there is a difference between phonemes and morphemes. Morphemes belong to a higher level of linguistic organization where meaning comes into play. Therefore, when applying the procedure of phonological segmentation to the identification of morphemes, one should be aware of the fact that “morphemes are themselves meaningful units” (ibid: 433). This leads Nida to the following conclusion: “In phonemics we admit conditioned and free variation only when the meaning of the whole remains constant. In morphology we must admit sub-morphemic differences of both form and meaning within a phonetic-semantic distinctiveness of a single morpheme” (ibid: 434).

2.6.2 After descriptivists

The three approaches to allomorphy articulated in the three articles by Harris (1942), Hockett (1947), and Nida (1948) are rather different. Interestingly, it is Harris’ version of allomorphy, the most rigid of them, that persisted in the history of the field. The substantial fruitful discussion about the nature of allomorphic relations that took place on the pages of *Language* was not continued by the community of linguists. Quite the opposite, the advances achieved by American descriptivists in the 1940s are not reflected in the modern theory of allomorphy. Instead of the important amendments introduced to the theory of allomorphy by Hockett and Nida, the subsequent linguistic tradition focused on the core initial version of allomorphy that we find in Harris 1942.

Contemporary linguistics still operates with the definition that served as the point of their departure for discussion, with the original two categorical criteria. Thus, we find the same definition of allomorphy in a modern guidebook to morphology by Haspelmath (2002: 27): “[t]he crucial defining property is that they [allomorphs] have the same meaning and occur in different environments in complementary distribution”. In this regard, over the years that passed since the 1940s the reductionist approach to allomorphy prevailed and an unfortunate simplification of the allomorphic theory has taken place.

It is not an easy task to trace the history of linguistic ideas. However, for this discussion it is crucial to explain how it happened that the debate about non-standard properties of allomorphs (such as submorphemic semantic differences and distributional overlap), once started in *Language*, was abandoned.

Recall that it is Harris’ definition of allomorphy that was adopted by the field. His procedures of morpheme identification were based primarily on distributional criteria and, as pointed out by Matthews (2003: 81), they “were criticized by scholars outside

the American school” (e.g. Haas 1954). Crucially, Matthews asserts that “by the 1960s it was clear that the model in its strict form distorted the morphology of many languages” (ibid: 81).

In the 1960s, however, the general interest of mainstream American linguistics shifted from morphology to syntax and phonology, mainly due to the advent of generative grammar. Many surveys in the history of linguistics report on this fact (Beard 1995: 1; Bauer 2004: 6; Matthews 1993). Anderson & Dressler (2003: 82) claim that “morphology came to be thought of as a collection of relatively uninteresting but well-understood techniques for organizing a list of the words of a language”.

Regarding American linguistics, Beard (1995: 1) also states the fact that “morphology has all but vanished from the agenda of linguistic inquiry”. Crucially, Beard names two reasons for the decline of morphology in the history of American linguistics: first, the unresolved flaws in the Structuralist model of morphology including zero and empty morphemes, and second, ignorance of the semantic side of morphemes: “The second reason of the malaise of the discipline is that while the word and morpheme have two sides: the semantic and phonological, only for a brief period in the 19th century did research treat both sides evenly in any attempt to account for the relation between the two. The initial interest of the Greeks centered exclusively on the semantic side of words, their various categories and subcategories. Over the centuries interest slowly shifted, not to a balanced scrutiny of the meaning and sound of morphemes, nor to the crucial relation between the two, but rather to the exclusive study of the phonological side of morphemes that dominated the Structuralist school” (Beard 1995: 1).

Similarly, Matthews (1993: 1) points out that American structuralists were mostly interested in techniques for classifying data and ignored meaning. In this sense, the ideas about semantics expressed by Nida (1948) were rather an exception than a trend.

A natural question that arises in this regard is: how did it happen that Structuralism, the framework which introduced the idea of linguistic sign as a pairing of form and meaning in the first place, facilitated the privileged status of formal approaches and commitment to distributional criteria at the expense of attention to semantic issues?

According to Matthews, “[t]his is mainly the fault of Bloomfield, who was the First to Make a Science of Linguistics¹⁷” and claimed that one cannot study meaning scientifically. Indeed, Bloomfield’s theory of language was the most influential in American linguistics in the first part of the 20th century and shaped the linguistic views of American descriptivists, including Hockett, Harris, and Nida. Without Bloomfield’s conception of the morpheme, this discussion would be incomplete.¹⁸ That is why I will briefly outline his view as well as existing alternatives in the following section.

2.6.3 Before descriptivists

2.6.3.1 Bloomfield 1933: Abandoning issues of semantics as unscientific

Bloomfield’s theoretical program incorporated the notion of *morpheme* proposed by Baudouin de Courtenay (1881a; 1895), as well as major tenets of structural linguistics

¹⁷ Capitalization is used in the original.

¹⁸ Note that Bloomfield’s work is relevant here not because he had a theory of allomorphy but because he advocated a particular view of semantics discussed in 2.6.3.1.

outlined by Saussure (1916). We can firmly state this, first, because Bloomfield uses the term *morpheme* in *Language* (1933 [1961]) and, second, because in 1923 he wrote a review of the second edition of Saussure's *Course in General Linguistics*. Bloomfield's understanding and interpretation of Saussure's ideas are analyzed in detail in Harris 2001: 59-75.

According to Bloomfield, the major contribution of Saussure's work is a bare systematization of ideas that had already been previously expressed (Bloomfield 1923: 63). As pointed out by Harris (2001: 68), Bloomfield failed to appreciate the Saussurean concept of the linguistic sign. In his review of Saussure, Bloomfield tries to avoid the term *sign*, does not provide a clear explanation of *signifier* and *signified*, and describes language from a synchronic perspective as a *system of signals* (Harris 2001: 68): "once it is left out, however, the originality of Saussure's approach is lost and the CLC¹⁹ begins to sound like a rather simplistic résumé of ideas commonplace in nineteenth-century linguistics. Which is exactly how Bloomfield's review presents it" (Harris 2001: 69).

One of the reasons that Bloomfield failed to appreciate the very idea of the linguistic sign, according to Harris (ibid: 72), was Bloomfield's behaviorist perspective. From this perspective, a pairing of form and meaning was just an unimportant remnant of mentalism. This observation explains why Baudouin de Courtenay's notion of the morpheme was re-established by Bloomfield as well.

In *Language*, Bloomfield suggests that "[a] morpheme can be described phonetically, since it consists of one or more phonemes, but its meaning cannot be analyzed within the scope of our science" (Bloomfield 1933 [1961]: Ch.10, 161). In his view, "the meaning of a morpheme is a sememe. The linguist assumes that each sememe is a constant and definite unit of meaning, different from all other meanings, including all other sememes, in the language, but he cannot go beyond this. There is nothing in the structure of morphemes like *wolf*, *fox*, and *dog* to tell us the relation between their meanings; this is a problem for the zoölogist" (ibid: 162)²⁰. This statement illustrates the behaviorist approach adopted by Bloomfield during the years he spent at Ohio State University, being strongly influenced by the behaviorist psychology of A.P. Weiss (Matthews 1993: 64). Likewise, Bloomfield defines the meaning of an utterance in terms of relevant stimuli and reactions. As he puts it, "[t]he signals can be analyzed, but not the things signaled about" (ibid: 162). The focus of Bloomfield's attention is on the formal shape of morphemes and the distribution of allomorphs, whereas their semantic content, according to Bloomfield, is not available for a scientific linguistic analysis because it is not directly accessible.

Because Bloomfield believed that the study of semantics lies beyond the scope of linguistic science, he also took the next logical step in claiming that lexicon is "really an appendix of grammar, a list of basic irregularities" (Bloomfield 1933 [1961]: 274). Tracing the history of American linguistics, Matthews (1993: 14) suggests that exactly this assumption about the lexicon dominated "almost all American thinking from the 1930s until late in the 1970s". Indeed, we can encounter the same assumption more than half a century after Bloomfield's claim: "The lexicon is like a prison – it contains only the lawless, and the only thing that its inmates have in common is lawlessness." (Di Sciullo & Williams 1987: 3).

¹⁹ Harris (2001: 69) uses CLC as an abbreviation for *the Course in General Linguistics*.

²⁰ Bloomfield repeated this idea in many places in *Language*: "We assume that each linguistic form has a constant and definite meaning, different from the meaning of any other linguistic form in the same language" (Bloomfield 1933 [1961]: 158).

In this light, Bloomfield's theory of language facilitated underestimation and even neglect of the semantic component in subsequent studies of allomorphy. Since it was already established that morphemes were minimal meaningful linguistic units, Bloomfield assigned their semantics to the lexicon which was previously thought of as a storage of words (Bloomfield 1933 [1961]: 161-63). According to Beard (1995: 6), this assumption radically shifted the focus of studies to the formal side of morphemes, issues of Trubetskoy's morphophonemics, and allomorphy understood as a simple variation of phonological form.

2.6.3.2 Saussure 1916: Linguistic sign is a word

An important aspect of Saussure's version of Structuralism relevant for our discussion is that he was not concerned with morphemes. Anderson (to appear: 3) points out that Saussure was apparently familiar with both Baudouin's and Meillet's use of the novel term *morpheme*, but did not adopt this notion. Moreover, careful reading of Saussure indicates that he applied the idea of linguistic signs to whole words, but not to their structural parts (cf. Anderson to appear: 3; Harris 2001: 69). Beard (1995: 6) even states that Saussure "carefully avoided the term "morpheme" in his lectures and associated his definition of the sign only with words." Thus, morphemes are left out of his discussion of form-meaning relations.

Saussure stressed the idea of rules and likened the system of language to a chess game. According to this metaphor, the meaning of a linguistic sign comprises not "internal" idiosyncratic characteristics of the sign, but rather a set of contrasts, or "external" oppositions, that the sign is involved in.

A question arises: were there any other alternatives in the Structuralist movement that provided a more balanced view of the morpheme, where its functional, or semantic, side was not overlooked? The answer is yes: we can find a quite different conception of the morpheme in the works of Baudouin de Courtenay, the scholar who coined the term *morpheme* and first introduced it to linguistics.

2.6.3.3 Baudouin de Courtenay 1881: The "psychological autonomy" of the morpheme

At the end of the 19th century, Baudouin de Courtenay expressed truly innovative ideas that anticipated European Structuralism. He was the first to distinguish between phonetics and phonology²¹. He proposed the contrast between "static" and "dynamic laws" that rudimentarily anticipated, according to Jakobson (1971), the opposition of synchrony and diachrony²² (Heaman 1984: 29). Among other new terms, he first introduced the term and notion of *morpheme* (Baudouin de Courtenay 1881a, 1895; Stankiewicz 1972: 7; Trask 1996: 227; Jarceva 1990: 312). Likewise, the notion of *morpheme alternants* was first developed in his works.

Yet, Baudouin de Courtenay is very little known in the West, as opposed to Saussure and Bloomfield, whose books appeared much later. Stankiewicz, the translator

²¹ In Baudouin's terminology: *antropophonics*, i.e. examination of acoustic and physiological properties of sounds, and *psychophonetics*, i.e. the study of functions of sounds.

²² In Baudouin's terminology: the contrast between "static" laws and "conditions that form the foundation of life of sounds in a language at a given moment" on the one hand and "dynamic laws and forces" which constitute a historical development on the other hand (Heaman 1984: 29).

of Baudouin's works into English (1972), states that his ideas are scattered in numerous publications and are often expressed in languages like Russian and Polish that are hardly accessible for western linguists, whereas Saussure's ideas were collected and published by his students as a single book, and Bloomfield's program is laid out in detail in his major monograph. We can speculate that if the ideas of Baudouin de Courtenay about the morpheme and the nature of linguistic units would have become better known and developed in his time, the very understanding of allomorphy in modern linguistics could have been different.

Among these three founders of Structuralism, it is Baudouin de Courtenay who places most focus on the meaning of linguistic units and the "psychological", as he calls it, or mental, aspect of linguistic phenomena. Baudouin de Courtenay emphasized this idea in many of his works. In the article "A chapter from Psychophonetics" (in the original "Glava iz psixofonetiki"), Baudouin de Courtenay claims: "I consider myself an adherent of the linguistic school which emphasizes the psychological factor in all linguistic phenomena"²³ (Baudouin 1895; Stankiewicz 1972: 146). He further lists his principal commitments in the study of language: "[t]o treat language as an "organism", and linguistics as a natural science, is a fallacy without any scientific basis. Language is exclusively psychological. The existence and development of language is governed by purely psychological laws. In human speech, or language, there is not a single phenomenon which is not psychological" (Stankiewicz 1972: 213). A *phoneme* is understood by Baudouin de Courtenay not only as an abstract unit that can be realized in more than one way, but also as "the psychological equivalent of a speech sound". Likewise, Baudouin's definition of *morpheme*²⁴ also refers to the psychological aspect of this notion: "The *morpheme* is a part of a word which is endowed with psychological autonomy and is for the very same reason not further divisible. It consequently subsumes such concepts as the *root (radix)*, all possible *affixes, (suffixes and prefixes)*, *endings* which are exponents of syntactic relations, and the like"²⁵ (Baudouin de Courtenay 1895; Stankiewicz 1972: 153).

This definition comprises several important ideas. First, *morpheme* is proposed as an umbrella term for roots and affixes. Second, the notion of *morpheme* contains a psychological element (the so called "psychological autonomy") and thus refers to speakers and their linguistic competence. Under the psychological aspect Baudouin arguably understands some kind of mental representation of a linguistic unit that a given speaker of a language has. "Psychological autonomy" can be also understood as the meaning of a morpheme that is relevant for a speaker.

According to Baudouin de Courtenay, there are linguistic phenomena that justify the notion of *morpheme*, since they give evidence of the psychological reality and autonomy of this structural unit in speakers' grammars. In this regard, Baudouin de Courtenay discussed three processes that indicate speakers' ability to single out

²³ Translated by Stankiewicz. The original reads in Russian as "kotoroe vo vsech javlenijax jazyka usmatrivaet v pervuju očered' psixičeskij faktor" (Baudouin de Courtenay 1963: 266).

²⁴ Baudouin worked intensively on morphological issues in 1870s, when he was giving lectures at the University of Kazan' (Stankiewicz 1972: 35; Heaman 1984: 28) For the first time, the term *morpheme* appears in the program of lectures for Kazan' written in 1877 and published in 1881 (Baudouin de Courtenay 1881a [1963]). The term *morpheme* occurs first sporadically in Baudouin's writings, but in 1895 it becomes a term which he uses frequently in his theory of alternations (Baudouin de Courtenay 1895 [1963]).

²⁵ "Morfema – ljubaja čast' slova, obladajuščaja samostojatel'noj psixičeskoj žizn'ju i dalee ne delimaja s ètoj točki zrenija." (Baudouin de Courtenay 1895 [1963]: 272)

morphemes in the structure of a word, interpret them and modify them – these are reinterpretation (the Russian equivalent term is *pererazloženie*), analogy, and folk-etymology. Baudouin de Courtenay concludes that “linguistic creativity takes place not only in the field of syntax, i.e. the compilation of ready-made words into phrases and sentences, but also in the combination of morphemes into words” (Stankiewicz 1972: 36).

Baudouin de Courtenay defines *morpheme* as a unit of word structure, indivisible from the perspective of its psychological nature. Emphasizing the structural atomism behind the idea of *morpheme*, he also used parallel periphrastic synonyms like “morphological atoms”, “morphological components” of a word, and “roots in the broad sense” (Baudouin de Courtenay 1881a [1963]: 114; Stankiewicz 1972: 36).

Baudouin de Courtenay recognized the need for the notion of *morpheme*, while he was elaborating his classification of alternations and the procedure of establishing an alternation of phonemes in the first place. According to his theory, alternations of phonemes should take place within the same unit – a *morpheme*. In this light, sound changes are only comparable if they occur within the same morpheme (Baudouin de Courtenay 1895 [1963]). This is probably the reason why the term *morpheme* had such success. For example, Meillet adopted this term as a “joli mot” in his translation of Karl Brugmann’s comparative grammar (Stankiewicz 1972: 36; Anderson to appear: 2). In his discussion of phonological alternations within morphemes, Baudouin de Courtenay examines morphophonemic processes and the phenomenon of allomorphy – the variation of morpheme’s phonological shape. The notion of *morpheme* is thus used as an abstraction that can be realized in different ways which are united under the same meaning.

2.6.3.4 Summing up: Three versions of Structuralism and their implications for allomorphy

Wrapping up the discussion about conceptual origins of *morpheme* and *allomorphy*, I want to highlight two major ideas.

First, I tried to show that the understanding of allomorphy that prevailed in the history of linguistics is a natural outcome of underestimating the role of meaning in the organization of language. Surprisingly enough, this trend takes its roots in Structuralism (in particular, Bloomfield’s version of Structural linguistics), despite the fact that it is the Structuralist framework that established the very notion of the linguistic sign as a pairing of form and meaning.

Second, I have argued that Structuralism as a movement had at least three versions which stressed different aspects of the theory of language and therefore offered different implications for the concept of allomorphy. The Structuralism of Jan Baudouin de Courtenay (1881a,b), which was further developed in the Prague Linguistic Circle and in the Russian linguistics of the 20th century, produced the idea of the morpheme as a structural linguistic unit with “psychological autonomy”. The Structuralism of Ferdinand de Saussure (1916), most relevant for European linguistics, elaborated the semiotic theory of linguistic signs but applied it primarily to whole words, avoiding the discussion of smaller units like morphemes. The Structuralism of Leonard Bloomfield (1933), which shaped a long-standing American tradition of linguistic thinking, abandoned the issues of semantics and assigned morphemes to the lexicon. The idea of separation of the study of formal relations and the study of meaning was welcomed in the generative understanding of language. Matthews (1993: 3) points out that this motif

in different variants “runs through American linguistics from the 1940s onwards, and where similar views are current elsewhere it is largely under American influence.” We can spot this idea in the theory of linguistic modularity (Chomsky 1965; 1981), where the lexicon constitutes a module separate from the modules of syntax and phonology, and where morphology is assigned to the domain of lexicon.

As we see from the history of linguistic thought, the understanding of the morpheme and allomorphy is determined by a broader view on natural language and its organization, the role of meaning and principles of linguistic analysis. In this light, the drawbacks and flaws of the traditional understanding of allomorphy can be overcome if we reconsider this phenomenon from a new perspective that takes a different approach to categorization, semantics, and data collection.

2.7. Proposal of this dissertation

The goal of this dissertation is to adjust the concept of allomorphy to available data in order to make the theory more adequate and realistic. A question arises: what theoretical amendments are necessary to implement for the notion of allomorphy? In this section I explain how the concept of allomorphy can benefit if we reconsider it from the perspective of Cognitive Linguistics – a framework advocated in Lakoff 1987, Langacker 2008, Taylor 1995, Dąbrowska 2004, Nessel 2008, and Janda 2010. I will draw attention to four issues – 1) approach to data and data collection, 2) a distinction between Prototypical, Standard, and Non-Standard allomorphic relations, 3) the role of semantics, and 4) statistical modeling of multifactorial conditioning and distributional overlap.

2.7.1 Data-driven approach: From Data to Theory

The very notion of allomorphy, as well as its definition, criteria, and types of allomorphic conditioning is an intellectual product of the Structuralist movement of linguistic thought. Note that in the Structuralist era and indeed until recently it has been common to build a theoretical analysis on data that was often restricted to a set of manually collected examples, sometimes even constructed by linguists themselves.

Today, the rapid development of electronic corpora and technology have set the stage for different standards for data collection and require adjustment of common linguistic routines in order to take advantage of new possibilities for data collection. The key feature of the new linguistic era is open access to large amounts of data, most of which have not yet been analyzed. After the invention of electronic corpora linguists can no longer excuse themselves from accounting for limited or restricted datasets and avoiding counterexamples. Today it is possible to obtain a more comprehensive picture of any linguistic phenomenon, its functioning and distribution, just because there is much more data available.

Therefore, a primary requirement of modern linguistics is to adjust the concept of allomorphy to data that might contradict or go beyond the narrow scope of the traditional clear-cut definition. I propose a data-driven approach to investigation of this phenomenon, which implies that the data should be primary to the theory, and the approach should be agnostic. The goal is to incorporate data that shows near-allomorphic relations into the theory of this phenomenon. The first step towards this goal is to revisit common objectivist assumptions about the nature of linguistic categories and rethink allomorphy in terms of a Radial Category.

2.7.2 Allomorphy as a Radial Category: Standard & Non-Standard exemplars

It is useful to rethink the modern concept of allomorphy in terms of a Radial Category, which has its prototype and periphery, following the notions developed in Cognitive Psychology and Cognitive Linguistics in the past forty years. I will first explain the difference between an Aristotelian Category and a Radial Category and then turn to how we can apply the latter to allomorphy.

The absolute idealistic understanding of allomorphy that prevails in Modern Linguistics depicts this phenomenon of a natural language as an Aristotelian category – a classical category with clear-cut boundaries, defined by binary features, internally unstructured and with an equal status of all exemplars (Taylor 1995: 23). Either two linguistic formants are allomorphs or they are not. This conception of allomorphy leaves no space for an intermediate zone of membership, it rules out all ambiguous cases that might to some degree belong to this category and to some degree not. Instead, there are only two clear degrees of membership that the Aristotelian category of allomorphy allows for – a member and a non-member, that is allomorphy and non-allomorphy. The two defining characteristics of allomorphs – identical meaning and complementary distribution – are understood as binary features, they can be either [+] or [-], a matter of all or nothing (Taylor 1995: 23). The rival candidates either have a perfectly identical meaning or they do not; likewise, their distribution is either purely complementary or not. Therefore, all exemplars of allomorphy are equally good. There is no hierarchy or difference in their status.

This kind of understanding is characteristic of the traditional objectivist scientific view of categories that lies at the heart of both Structuralist and Generative approaches to language (Lakoff 1987: 161; Taylor 1995: 21-24). Lakoff (1987: 159) explains that “on the objectivist view, reality comes with a unique, correct, complete structure in terms of entities, properties, and relations”. In this light, the linguistic notion of allomorphy is a theoretical construct, a product of mental activity, an idealization, a category that exemplifies a more general fundamental way of scientific thinking and reasoning. Therefore, it might or might not capture the relationship of form and meaning that we observe in data. The traditional objectivist way of thinking is problematic not because there are no classical categories, but rather because this way of thinking assumes a priori that all categories are classical (Aristotelian), whereas there is a lot of empirical evidence that is in conflict with this view.

It has been shown in a number of experimental studies in cognitive psychology (Rosch 1973, 1975, 1977a,b, 1978 [1999], 1983, Rosch & Mervis 1975) that phenomena of human cognition, and in particular linguistic phenomena, have a radial structure, with a central prototypical member and marginal members that lie at the periphery of the category. Membership is determined by the similarity with the prototype. Because similarity is a graded property, a Radial Category allows various degrees of membership. Moreover, the members of a Radial Category are characterized by family resemblance – they can share some properties and differ in others. As a consequence, they can be closer or further away from the conceptual prototype. As opposed to the classical (Aristotelian) objectivist category, a Radial Category is internally structured (prototype vs. periphery), reflects gradience and scalarity, its members can be assigned numerical values and be compared with respect to the key features of the prototype. The prototype is “a conceptual core of a category” (Taylor 1995: 59), a schematic representation that can be instantiated in individual exemplars which are most representative of the category.

In this light, it is fruitful to look at the phenomenon of allomorphy as a Radial Category. The idea that concepts of linguistic theory can have a radial structure is not entirely new. It was mentioned by Taylor (1995: 59) and explicitly discussed with regard to allomorphy by Dickey & Janda (2009: 247) and Janda & Nessel (2010b). The idea is that we should reevaluate the cases that are in perfect compliance with both criteria for allomorphy as instances of the prototype rather than the only possible cases. Deviations from identical meaning and complementary distribution should be recognized and explored in detail. Deviations should be either assigned or not assigned to allomorphy depending on statistical measures of semantic similarity/divergence and significance of distributional differences.

If this revised version of the concept is to be applied, what should be the prototype of allomorphic relations? In the literature we observe a variety of rather heterogeneous phenomena, where the choice of allomorphs is conditioned by phonological, morphological, lexical, semantic, or historical factors. However, there is no discussion of which types of conditioning yield more prototypical allomorphs. Some meaningful conclusions can be made already by comparing the types of allomorph conditioning recognized in different sources. For example, semantic and historical conditioning of allomorphs as well as suppletive allomorphy are often not even mentioned in many reference works. This indicates their marginal status in the category of allomorphy.

In this regard, Bauer (2004: 71-72) points out that “in the most clear-cut instances, members of the set are clearly phonologically related to each other, and <...> the conditioning factor(s) for the complementary distribution can be stated in phonological rather than in lexical terms”. In other words, Bauer suggests that the core type of allomorphy is the one with regular phonological conditioning of allomorph selection and allomorphs which are clearly related via a phonological rule. In this light, lexically conditioned allomorphy is less clear, or “less allomorphy”, because it is not as regular as phonologically conditioned allomorphy is. Rather, lexical conditioning accommodates exceptions and might be a peripheral fact of marginal allomorph distribution. Bauer mentions, for example, that his conditions are clearly met in *cat-s*, less clearly in *ox-en* and *child-ren*, and even less in *bear* and *birth*. Crucially, Bauer concludes this discussion with the following statement: “but there is no well-defined cut-off point on the scale of phonological relatedness” (ibid: 72). This property is graded, and so is allomorphy. However, instead of further elaboration, this insight about the lack of clear boundaries triggered criticism of the very notion of allomorphy which some scholars try to abandon (Anderson 1992; Beard 1995; Gerd 2004).

Returning to Bauer’s claim, I agree that it is fair. Phonologically related and phonologically conditioned allomorphs are most likely to meet the requirements for identical meaning and complementary distribution. The more automatic the phonological conditioning of allomorphs, the less the chance of lexicalization and developing semantic differences. Therefore, I suggest the following definition of prototypical allomorphy:

Prototypical allomorphy is a relationship of phonologically similar and phonologically related forms which have identical meaning and complementary distribution conditioned by an automatically applied (regular and productive) phonological rule.

This definition bears several implications. Not all phonologically conditioned allomorphies are prototypical. At the same time, allomorphies that are morphologically conditioned and lexically conditioned are not prototypical in terms of a Radial Category structure, but can be closer or further away from the prototype. In addition, I propose the notion Standard Allomorphy as an intermediate notion that should capture the zone between the prototype and the periphery:

Standard Allomorphy is a relationship of forms that satisfy two criteria – identical meaning and complementary distribution.

Standard Allomorphy captures phonological, morphological, and lexical conditioning, if they satisfy the two criteria mentioned in the definition. A given allomorphy might be not Prototypical, but nevertheless Standard, if it meets the two criteria. Phonological similarity is not an obligatory characteristic. Therefore, suppletive allomorphs can also be Standard if they conform to the two criteria.

According to the concept of a Radial Category, apart from Prototypical and Standard allomorphies we are entitled to recognize Non-Standard allomorphies. I propose the following definition of Non-Standard allomorphy:

Non-Standard allomorphy is a relationship of forms that fail to satisfy either the criterion of identical meaning or the criterion of complementary distribution or both criteria. Despite their deviation from Standard Allomorphy, these forms exhibit strong evidence that binds them into a single perceptible morpheme. This evidence can be in terms of a strong semantic similarity (which exceeds their semantic divergence) and robust pattern of distribution. Both characteristics should be justified quantitatively in terms of statistical significance.

Non-Standard allomorphy covers those cases when we observe sub-morphemic semantic or stylistic (register) differences between allomorphs, an overlap in their distribution, or semantic conditioning of allomorph selection. Non-Standard allomorphy introduces the idea that allomorphy is a continuum rather than an all-or-nothing property. Non-Standard allomorphy refers to data that occupies an intermediate position between Standard Allomorphy and Non-Allomorphy.

Figure 4 visualizes the radial structure of the new concept of allomorphy that I adopt in this study. It offers a schematic representation of allomorphy as a Radial Category, which has a central Prototype, Standard members, and Non-Standard deviations that lie in the periphery of the category and have marginal status. In Figure 4, the sizes of the circles do not correspond to frequency of attestations that these types of allomorphy have in empirical data. This is a question for future research to find out the relative proportions of each subtype of allomorphy in the entire category – how large is the proportion of allomorphies that instantiate the Prototype, how many cases can be assigned to the Standard type and how large is the part constituted by Non-Standard allomorphic relations. It can be the case that Non-Standard allomorphies might predominate in this picture. If so, the very prototype should probably be revisited. However, this issue goes beyond the scope of my study, where I first of all aim at exploring the grey zone of intermediate cases that lie between Standard Allomorphy and Non-Allomorphy.

By contrast, Figure 5 visualizes the way that Allomorphy has been thought of in mainstream linguistics so far – as a Classical Category which is defined by binary

features and lacks any internal structure. Recall that many examples favored by textbooks as most representative of allomorphic relations, in fact, fail to satisfy at least one of the two criteria and should be relegated to the realm of Non-Allomorphy.

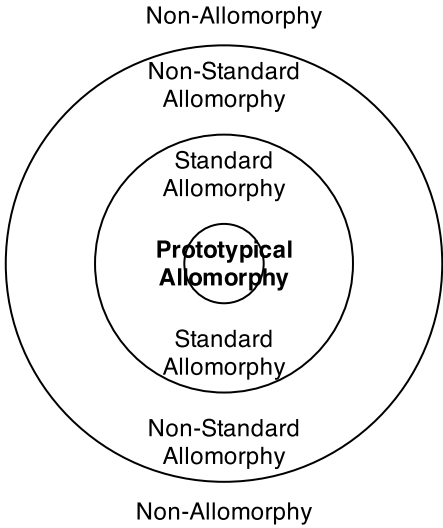


Figure 4: Allomorphy as a Radial Category.

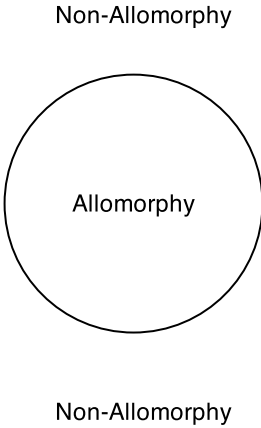


Figure 5: Allomorphy as a Classical (all-or-nothing) Category.

A comparison of Figures 4 and 5 clearly shows that rethinking allomorphy in terms of a Radial Category broadens the scope of this notion: it brings in a lot of relevant data concerning morpheme identity and form-meaning asymmetries that would otherwise be overlooked. Moreover, the understanding of allomorphy as a Radial Category is more realistic and makes the model more flexible: it can now capture both clear exemplars and complex deviations, making the former a standard that helps to improve our understanding of the latter. Thus, a Radial Category model of allomorphy does not dismiss the original definition altogether. Instead, it incorporates the two criteria into the center of the category.

2.7.3 The role of semantics

The phenomenon of allomorphy lies at the crossroads of three linguistic levels – phonology, morphology, and semantics. Attention to each of these aspects should result in a more balanced and multidimensional investigation of morphemes. However, in previous research the role of semantics in this trio was underestimated and therefore understudied (recall 2.4 and 2.6), lexical preferences of allomorphs often were “explained” via the specificity of the semantic module of language that was assumed not to be worth much interest and attention.

As opposed to the previous line of scholarship, in this dissertation I propose that it is fruitful to approach allomorphy from the perspective of Cognitive Linguistics – a framework where “meaning is recognized as the driving force of language” (Janda 1993: 310) and which views semantics as “playing a primary role in the organization of all linguistic phenomena” (Janda 1993: 310).

As soon as we start thinking about the semantic content of allomorphs, a number of questions arise. How can we apply the criterion of semantic identity if the candidates for allomorphy are highly polysemous? What if they share only a subset of their

semantic content, whereas in the remainder of submeanings they differ? Is it possible to measure the semantic similarity and divergence and assess which of the two predominates? What kind of semantic differences can be considered sub-morphemic and what kind of differences point to discrete morphemes? Is semantic conditioning of allomorphs a natural and common phenomenon or rare and exceptional? Can semantic factors compete with phonological and morphological factors in deciding the choice of allomorph, and if so, how to take them apart and estimate the relative weight of each factor? At present, these questions remain for the most part an unexplored terrain.

There are two major semantic phenomena in allomorphic relations that I will explore. First, semantics as an additional type of allomorph conditioning. This has been shown to be the case of Burushaski and Persian plural marking (Bye 2014: 44; Haspelmath 2002: 30) and the case of the Russian semelfactive marker (Dickey & Janda 2009). Moreover, semantics can be a factor that competes with phonology in conditioning allomorphs. I will show that the impact of all competing factors can be measured quantitatively and placed on a single scale of scores (cf. CART statistical models).

The second semantic phenomenon involved in allomorphic relations is semantic specialization, or partial split (divergence), of polysemous allomorphs. It is often not the case that the semantics of allomorphs is completely identical. Difference in form naturally triggers the process of semantic divergence and it can lead variants of a morpheme to develop differences in meaning. While allomorphs exhibit differences in their formal phonological shapes, they also naturally develop differences in use, since they complement one another in distribution. This divergence can cause allomorphs to gravitate towards different semantic centers of a shared semantic network. Co-existing forms can first start developing preferences for a particular contextual environment, then they might become highly frequent in this environment, at the next step – specialized for a particular context, and finally – idiomatically restricted to a closed list of words.²⁶ In this dissertation I suggest that semantics is not an invariable constant dimension in allomorphic relations. On the contrary, distributional routines bring about semantic specialization of morpheme variants for certain meanings and derivational patterns.

The process of semantic dissimilation of allomorphs takes place over time and results in partial semantic similarity and partial divergence. In such a situation, it is crucial to find out what predominates – semantic similarities or differences between allomorphs. In this regard, I adopt a system of methods elaborated in Cognitive Linguistics.

Cognitive Linguistics has shown that semantics is not a chaotic vacuum but rather a structurally and hierarchically organized cognitive space. Cognitive Linguistics provides a highly elaborated methodology for analyzing complex semantic issues in polysemous particles, prepositions, and affixes (Janda 1986; Janda & Nessel 2010a; Janda et al. 2013), such as: how to distinguish between the primary and secondary senses, how to single out the spatial embodied image schema which constitutes the prototypical sense, how to elicit the network of polysemy and establish the most prominent senses on the basis of different types of frequencies, etc.

In all case studies presented in this dissertation I will be concerned with measuring the degree of semantic similarity and divergence of the candidates for

²⁶ This process is well documented in the development of Slavonic variants of Russian words and affixes: Pilipenko (2001) shows that the uses which were most frequent for Slavonic morphemes have often become the only possible uses.

allomorphy. I promote the methodology of Radial Category Profiling, first proposed in Nessel et al. 2011. This methodology makes it possible to compare polysemous affixes qualitatively (meaning-by-meaning) and also quantitatively – in terms of amounts of individual lexemes where the affixal submeanings are manifested.

Likewise, the different status of senses can be taken into account. For example, sharing the spatial prototypical sense is a strong argument in favor of semantic similarity and allomorphy, whereas semantic difference in the prototypical sense, on the contrary, strongly points towards distinct morphemes.

2.7.4 Statistical modeling of Non-Standard cases

An important property of data that a Radial Category is meant to capture is the gradient nature of linguistic phenomena. It has been intensively studied in Cognitive Linguistics (Dąbrowska & Street 2006, Dąbrowska 2012) and recently gained a considerable amount of attention in the generative framework as well (Bard et al. 1996; Keller 2000; Sorace & Keller 2005). Gradience is usually revealed when large quantities of empirical data are analyzed – be it data collected from corpora or data elicited experimentally from speakers. Gradience of data means that the observed contrasts are not clear-cut. Rather, they tolerate some degree of overlap. Statistical modeling can both manage large quantities of data and identify robust tendencies in the distribution of gradient data.

Moreover, the need for statistical computation arises when we have to account for multifactorial dependencies, such as mixed allomorph conditioning that involves competing phonological, semantic, stylistic, and, possibly, also historical factors. In this dissertation I offer several examples of advanced statistical modeling of non-standard allomorphic relations, where a statistical program measures the relative impact of competing factors in allomorph conditioning and produces a visual graph, interpretable in terms of complex interactions of factors (cf. CART methodology).

Statistical analysis is the most promising technique that can bring new insights to exploring the zone of Non-Standard Allomorphy. In this regard, I offer a number of methods for capturing significant contrasts that do not exhibit perfect complementary distribution. I suggest that allomorphic status can be established on the basis of statistically significant difference in the distributions of candidates. The larger the effect size, the stronger the impact of the factor that governs the distribution. Thus, distributional overlap that violates ideal complementary distribution is not decisive of the allomorphic status.

Statistical modeling can thus be helpful in distinguishing between submorphemic differences on the one hand and different morphemes on the other hand.

2.8 Conclusions

In the beginning of this chapter I tried to show that even textbook examples do not satisfy the current definition of allomorphy in its strict sense. Furthermore, I have argued that allomorphy as a notion is a theoretical construct produced within the Structuralist approach to language. Bloomfield's version of Structuralism in general and his amendments to the notion of morpheme in particular facilitated disparagement of the semantic component of grammar. Throughout the 20th century, we observe multiple attempts to separate formal relations from meaning in mainstream linguistics (Matthews 1993: 3). The most rigid concept of allomorphy has prevailed over the years.

The definition of allomorphy currently used in linguistics is formulated in absolute terms, suggesting only two possibilities – allomorphy and non-allomorphy.

My proposal is four-fold. First, I pursue a data-driven approach to allomorphy which implies an accurate account of comprehensive sets of authentic data collected via corpora and experiments. A data-driven approach goes from empirical data to generalizations, trying to incorporate counterexamples and deviations into the model, instead of ignoring what is difficult to explain. Second, as a natural consequence, I propose that instances of Non-Standard Allomorphy should be recognized along with those of Standard Allomorphy. I suggest that it is fruitful to rethink the very notion of allomorphy in terms of a Radial Category with Prototypical, Standard, and Non-Standard members. By accepting Non-Standard Allomorphy, we overcome the obvious flaws of the narrow Structuralist definition. We should recognize distributional overlap and semantic differences as normal properties of allomorphs, as suggested already by Harris (1947) and to an even greater extent by Nida (1948). These two aspects present a challenge for the clear-cut definition, but their exploration is the place to start in our new approach. The third point of my proposal is the need to explore modification of semantics under allomorphic variation. The fourth part is the statistical methods that can capture several competing factors, their interactions, and gradience of data.

Chapter 3

Standard allomorphy in Russian prefixes: The case of RAZ- ~ RAS- ~ RAZO- ‘apart’

In this chapter²⁷ I provide a statistical account of an example of standard allomorphy in order to demonstrate that statistical modelling produces an outcome different from those that I find for non-standard cases in other chapters. I show how the same statistical methods (Decision Trees, Random Forest model, and Chi-square test) work when they target a distribution of formal shapes that are not affected by semantics.

While various quantitative methods continue gaining more popularity in linguistic studies (Janda 2013), the application of statistical tests for assessment of linguistic contrasts still remains debatable. “Language is never random”, says Kilgarriff (2005) questioning the appropriateness of statistical methods built on the assumption of independent observations. Since all linguistic phenomena and categories are interrelated, a purely random distribution, which corresponds to the null hypothesis, can in principle never be found in a linguistic dataset. A question arises as to whether a statistical model applied to the non-standard allomorphies in the remaining chapters would be able to capture the form-meaning relationship associated with standard allomorphy as well.

In this chapter I examine standard allomorphy in Russian morphology by looking at formal variation of the prefix RAZ- ‘apart’. Like other prefixes explored in this dissertation, RAZ- is very polysemous (Janda & Nessel 2010a). The phonological difference in its allomorphs is unambiguously reflected in orthography: RAZ- vs. RAS- vs. RAZO-. In addition, the factors that trigger the variation in its formal shape are not absolutely transparent. I show that a statistical analysis can reveal both standard allomorphy (cf. RAZ- ~ RAS-) as well as potential problems with the standard status of the observed distribution of related rival forms (cf. RAZ- ~ RAZO-).

3.1 Introduction

While most of this dissertation is devoted to the exploration of non-standard allomorphy, this chapter aims to explore what is traditionally described in terms of regular, or standard allomorphy in Russian prefixes. I look at two phenomena well represented among consonant-final Russian prefixes like POD-, IZ-, OB-. First, I address regressive voicing assimilation (#C_{voiced-} ~ #C_{voiceless-}) and undertake a case study of the devoicing of the prefix final consonant in the allomorphs RAZ- ~ RAS- ‘apart’. Second, I look at vowel-zero alternation in consonantal prefixes (#C- ~ #CV-) and examine a typical example of this phenomenon in the allomorphs of the same morpheme RAZ- ~ RAZO- ‘apart’. I show that the two phenomena are different in terms of how they match the notion of standard allomorphy.

In the first case study, I propose that RAZ- ~ RAS- exhibits an absolutely standard allomorphic relationship. In particular, this alternation meets both crucial criteria of

²⁷ The findings reported in this chapter were presented at the CLEAR group seminar on 30.11.2012.

allomorphy. With respect to semantics, the seven submeanings of RAZ- recognized within its polysemy do not affect the devoicing of the prefix coda. This supports the claim that the two alternants are semantically identical. Moreover, the distribution of RAZ- vs. RAS- is governed by an automatic phonological rule and seems to be exceptionless according to the data available.

The second case study makes it possible to conclude that the insertion of the vowel in RAZ- ~ RAZO- (as well as in other similar pairs of prefixal alternants) cannot be explained purely by phonological conditioning. I show that the choice of the allomorph is affected also by a morphophonological factor. Because the vowel/zero alternation is regularly attested in Modern Russian and both semantic and distributional criteria of allomorphy are met, the relationship of RAZ- ~ RAZO- lies in the scope of standard allomorphy.

The contrast between the two case studies appeals to the notion of standard allomorphy which was elaborated in Chapter 2. Under standard allomorphy I understand a structural relationship that holds between two or more morpheme alternants which satisfy both crucial criteria suggested in the literature (Matthews 1974; Haspelmath 2002; Booij 2005; Bauer 2001): in particular, standard allomorphs should 1) exhibit a perfect complementary distribution and 2) have identical meaning with no “sub-morphemic semantic differences” (in the sense of Nida 1948).

In this chapter I turn to the phonologically conditioned regular allomorphs and show how a statistical account makes it possible to distinguish between them.

3.2 Regressive voicing assimilation: #C_{voiced}- ~ #C_{voiceless}-

3.2.1 General remarks

Twelve of twenty-two Russian prefixes end with a consonant: VZ-, VOZ-, V-, S-, RAZ-, PRED-, POD-, OT-, OB-, NIZ-, NAD-, and IZ-. In Contemporary Standard Russian, all these prefixes, both monoconsonantal (type C-, like the prefix S-) and monosyllabic (type CVC-, like VOZ-), undergo regressive voicing assimilation. The assimilation is realized as devoicing in the case of VZ-, VOZ-, V-, RAZ-, PRED-, POD-, OB-, NIZ-, NAD-, and IZ-; and as voicing in the case of the prefixes S- and OT-. Only five prefixes reflect this allomorphic change in their orthography, namely those that end in the consonant z: VZ- ‘up’, VOZ- ‘up’, RAZ- ‘apart’, NIZ- ‘down’, and IZ- ‘out of’ (one could also add BEZ- ‘without’ to this list).²⁸

²⁸ In the history of Russian the spelling of prefixes ending in z was very unstable (see the overview from the oldest manuscripts to modern times in Kaverina 1999). Most of modern Russian orthography reflects the morphological representation of affixes, but the spelling of the prefixes in -z reflects the phonetic devoicing of their final consonant. This orthographic rule was first officially introduced during the reform of Russian orthography carried out by the Soviet government in 1918 (Kaverina 1999: 111-112). It is less known that the blueprint for this reform was developed before the revolution in 1904-1914 by the czar’s orthography committee at the Imperial Academy of Sciences by leading Russian linguists (the chair F.F. Fortunatov, the members J.A. Baudouin de Courtenay, A.A. Šaxmatov, R.F. Brandt and others). The orthography reform in general and the spelling of the prefixes ending in z in particular was meant to simplify and unify the Russian orthography and bring it closer to the pronunciation (Lopatin 2000). The modern spelling of z-final prefixes depends on the voiced vs. voiceless adjacent onset of the base and thus resembles the way these prefixes were spelled in the oldest Russian manuscripts (Kaverina 1999: 112).

Voicing assimilation creates automatic alternation of allomorphs. It takes place in the environment of obstruent consonant clusters created on morpheme boundaries: e.g. compare the alternants of the prefix IZ- in *iz-bežat* ‘avoid’ vs. *is-pit* ‘have a drink’, and variants of the prefix NAD- in *nad-lomit* ‘break’ vs. *nad-pisat* ‘superscribe’ (with the devoiced [d] > [t]). The voicing assimilation follows the following pattern (Neset 2008: 54-55):

- 1) “If the last member of an obstruent cluster is voiceless, then the preceding member is also voiceless” (Regressive devoicing);
- 2) “If the last member of an obstruent cluster is voiced, then the preceding member is also voiced” (Regressive voicing).

It is generally assumed that in the Russian prefixes listed above this process creates phonologically conditioned allomorphs with perfect complementary distribution and identical meaning. I now turn to a case study of one such allomorphic pair, namely RAZ- vs. RAS- ‘apart’.

3.2.2 Case study 1: RAZ- ~ RAS-

This study examines two hundred perfective verbs prefixed in RAZ- and RAS-. They are collected and tagged according to a number of parameters described below. The database is available online in an electronic Excel spreadsheet file “RAZ DATABASE” at <http://hdl.handle.net/10037.1/10078>. Note that for the first case study I exclude the ten verbs that feature the allomorph RAZO-. This allomorph is explored separately in the second case study in subsection 3.3.2.

Recall that the prefix RAZ- is highly polysemous. Thus the key question of this case study is the following: do sub-morphemic semantic differences play any role in the distribution of the voiced (RAZ-) vs. devoiced (RAS-) alternants of this prefix across verbs or not?

For the sake of the argument I adopt the semantic classification of the submeanings for RAZ- developed independently in the first corpus-based study of this prefix in Janda & Neset 2010a (cf. also Endresen et al. 2012: 255-258 and Janda et al. 2013: 42-52). This account describes the polysemy of the prefix RAZ- as a structured system and models it in terms of a cognitive radial network of interrelated submeanings. As shown in Figure 1, all submeanings of RAZ- are organized around the central spatial prototype ‘APART’²⁹. In Figure 1, the submeanings are represented as rectangular boxes. The lines that connect them visualize the motivational links between the submeanings. The numbers in parentheses are the quantities of individual verbal lexemes attested with a given submeaning of the prefix.

Let us briefly outline the polysemy of the prefix RAZ-. The prototypical meaning 1.‘APART’ refers to separation of individuals or parts of an object: *razognat* ‘disperse’ (< *gnat* ‘chase’), *razrubit* ‘chop apart’ (< *rubit* ‘chop’), *raskolot* ‘split up’ (< *kolot* ‘split’). A related submeaning 2.‘CRUSH’ implies that taking apart an object’s parts destroys its internal structure, as we see in the verbs *razdavit* ‘squash’ (< *davit* ‘squash’) and *rastoptat* ‘trample’ (< *toptat* ‘trample’). By contrast, the submeaning 3.‘SPREAD’ takes the prototypical meaning 1.‘APART’ in another direction: it does not imply destruction of the

²⁹ Small caps are used in order to refer to submeanings of the prefix in a radial category.

object but instead brings the focus to the object's edges. They move apart and thus make the object enlarge, or spread: e.g. *rasrastis* 'spread by growing' (< *rasti* 'grow'), *rasstelit* 'spread out a cloth' (< *stelit* 'lay a cloth'). The next submeaning, 4. 'SWELL', takes the spreading to a three-dimensional space: *razdut* 'swell' (< *dut* 'blow'), *raspuxnut* 'swell' (< *puxnut* 'swell').

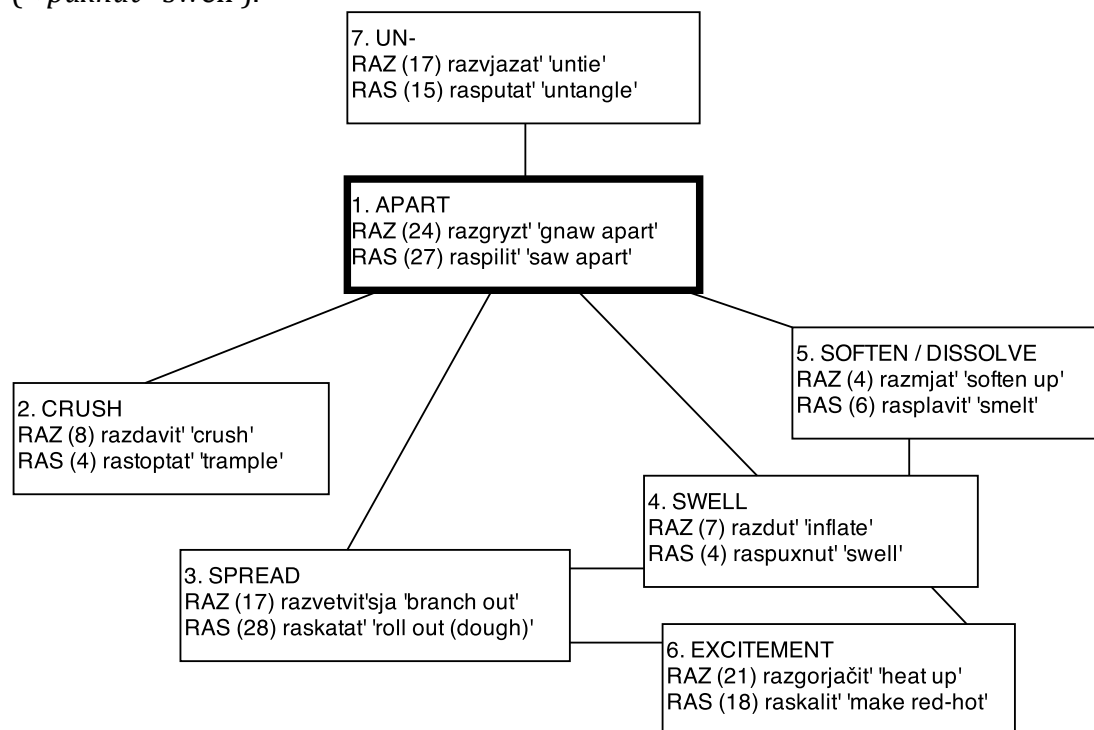


Figure 1: Model of polysemy of RAZ- and RAS-.³⁰

The submeaning 5. 'SOFTEN/DISSOLVE' applies spreading to the domain of substances that lose their original consistency and internal cohesion, as in *razmjat* 'soften up (by kneading)' (< *mjat* 'knead'), *rasplavit* 'smelt' (< *plavit* 'melt'), and *rastvorit* 'dissolve' (< *tvorit* 'create'). Since objects that spread and swell do so when heated, the submeaning 6. 'EXCITEMENT' is incorporated into this network as well. This submeaning refers to concrete heating (*razgorjaci* 'heat up', *raskaliti* 'make red-hot') as well as to metaphorical uses in the domain of emotions: *razveselit* 'exhilarate' (< *veselit* 'make cheerful'), *rassmešiti* 'make someone laugh' (< *smešiti* 'make someone laugh').

Finally, the submeaning 7. 'UN-', the most abstract of them all, is built on the same spatial prototype. While the simplexes *vjazati* 'tie' and *putati* 'tangle' refer to putting things together, their derivatives prefixed in RAZ- denote the reversal of these processes, namely taking things APART: *razvjazati* 'untie', *rasputati* 'untangle'. Similarly, the verbs *razgruziti* 'unload', *rasšifrovati* 'decipher', and *razmoroziti* 'defrost' denote undoing of what is designated by their simplexes: *gruziti* 'load', *šifrovati* 'cipher', and *moroziti* 'freeze'.

Summing up, the polysemy of the prefix RAZ- represents a coherent network of seven submeanings which are interrelated and motivated by a single prototype. As it appears from the examples provided in Figure 1, each of the seven submeanings can be expressed by both RAZ- and RAS-. In order to explore this issue further I compiled a database based on the one used in Endresen et al. (2012: 255-258). This database

³⁰ This radial category structure is adopted from Janda & Nessel 2010a; Janda et al. 2013: 42-52; Endresen et al. 2012: 255.

proved to be a representative sample of the verbs prefixed in RAZ-. A few words should be said on what is adopted in this database and on the amendments I have made for the purposes of the present study.

The database I adopt comprises all Natural Perfectives and those Specialized Perfectives³¹ that have over one hundred attestations in the Russian National Corpus. This shields the sample from marginal verbs that are less representative for the purposes of this study. In order to avoid duplication of data, this dataset merges reflexive verbs with their non-reflexive counterparts when the postfix *-sja* does not alter the lexical meaning of the non-reflexive base. In other words, the verbs like *rasselit'* 'settle apart' and *rasselit'sja* 'be settled apart', which differ only in terms of transitivity, are merged into a single lemma *rasselit'(sja)* and are taken as a single verbal lexeme.

There are two changes that I make in the adopted database. First, I exclude those few double entries that result from polysemous verbs. For example, the database originally included *razvit'* 'expand, develop' and *razvit'* 'uncurl, unwind' (both formed from *vit'* 'wind') as distinct verbal lexemes. For the study of semantics in Endresen et al. 2012 it was crucial to account for both submeanings of the prefix found in this derivative as well as for the connection of the two submeanings. However, for the purposes of the phonological analysis such double entries create a problem. They duplicate the data on attested combinations of RAZ- vs. RAS- with a particular onset type of the simplex base. Note that such double entries with RAZ- are rather few, with a total of seven verbs prefixed in RAZ- and RAS-, plus four verbs in RAZO-. In each case I preserved in the database the entry with the prefix submeaning which is more common for the given polysemous verb.

As a result, the final database contains two hundred verbs, including 98 lexemes prefixed in RAZ- and 102 lexemes prefixed in RAS-.

Second, in order to explore the interaction of semantic and phonological factors in the distribution of RAZ- and RAS-, I enrich the database with tags for several phonological characteristics. Each verb is additionally tagged for 1) the initial phoneme of the attached simplex base, 2) sonority of the onset (whether it is a voiceless obstruent, voiced obstruent, sonorant, or a vowel), and 3) the onset as simple vs. cluster. The final database has the structure shown in Table 1.

1	2	3	4	5	6	7	8	9	10	11
Lemma	Gloss	Allomorph	Initial Phoneme	Sonority Onset	Simple Or Cluster Onset	Base	Base gloss	Metaphor	Perf Type	Semantics
<i>rasšifrovat'</i>	decipher	ras	sh	voiceless	simple	<i>šifrovat'</i>	encipher	m	SP	undo
<i>rasšit'</i>	embroider	ras	sh	voiceless	simple	<i>šit'</i>	sew	m	SP	spread
<i>rasširit'(sja)</i>	broaden	ras	sh	voiceless	simple	<i>širit'(sja)</i>	expand	n	SP	spread
<i>rasševelit'</i>	set into motion	ras	sh	voiceless	simple	<i>ševelit'</i>	set into motion	n	NP	excite

Table 1: Sample of the database of the verbs prefixed in RAZ-.

³¹ I adopt the distinction between Natural and Specialized Perfectives proposed in Janda 2007. Natural Perfectives are perfective verbs that form an aspectual pair. When a Natural Perfective is formed, the prefix does not alter the lexical meaning of the base, as in *raskolot'* 'chop up' < *kolot'* 'chop up'. By contrast, in Specialized Perfectives, formed by the same prefixes, the prefix changes the lexical meaning of the simplex base, as in *raskryt'* 'open, uncover' < *kryt'* 'cover'.

Column 1 lists prefixed verbs (lemmas), column 2 provides their English glosses, column 3 specifies the allomorph of the prefix attested in the verb, and column 4 defines the initial phoneme³² of the base. The tags in column 5 describe the initial phoneme of the base according to the sonority hierarchy, featuring 102 bases starting with a voiceless obstruent, 55 bases starting with a voiced obstruent, 39 bases with a sonorant onset, and 4 bases starting in a vowel. Column 6 specifies whether a simplex base starts with a single consonant (143 verbs) or a consonant cluster (53 verbs) or has no onset (4 verbs). Columns 7 and 8 list the simplex bases and their English glosses respectively. Column 9 tells whether the use of the prefix is metaphorical and abstract (m) or non-metaphorical, concrete and spatial (n). The penultimate column (10) distinguishes between Specialized Perfectives (SP, 130 lexemes), where the prefix alters the lexical meaning of the verbal base, and Natural Perfectives (NP, 70 lexemes), where the prefix does not do so. The last column (11) tags the submeaning of the prefix recognized in the verb.

Let's say that we want to find out whether there are semantic differences between RAZ- and RAS-. Then we might want to know whether the two variants of the prefix have different preferences for certain submeanings. In other words, we should compare the radial category profiles of RAZ- and RAS-. A radial category profile is "the relative frequency distribution of the subcategories of a radial category" (Nesset et al. 2011: 21).

Table 2 summarizes the distribution of the two variants of the prefix RAZ- across the submeanings of the prefix. Here we see both the raw numbers of attested individual lexemes (also given in Figure 1) and the percentages of the total number of verbs prefixed with RAZ- and RAS-. It should be mentioned that the distribution of the two prefixes across verbs is perfectly complementary in the sense that there is not a single verb/verbal form that takes more than one allomorph.

Submeaning	# of verbs with RAZ-	RAZ- %	# of verbs with RAS-	RAS- %
1. APART	24	24.48 %	27	26.47 %
2. CRUSH	8	8.16 %	4	3.92 %
3. SPREAD	17	17.34 %	28	27.45 %
4. SWELL	7	7.14 %	4	3.92 %
5. SOFTEN/DISSOLVE	4	4.08 %	6	5.88 %
6. EXCITEMENT	21	21.42 %	18	17.64 %
7. UNDO	17	17.34 %	15	14.70 %
Total:	98	100%	102	100%

Table 2: Distribution of verbs with RAZ- and RAS- across submeanings of the prefix.

Table 2 shows the relative frequency distribution of the two allomorphs across the seven submeanings of the radial category. Figure 2 visualizes these values as bars, showing that the radial category profiles of RAZ- and RAS- are very similar although not identical.

Pearson's Chi-squared test indicates that the difference between these two profiles is not statistically significant: X-squared = 5.7, degree of freedom = 6, p-value =

³² Note that *sh* stands for *š*; *ch* for *č*; and *shch* for *šč* because diacritics are not recognized by the version of the statistical software employed in this study (R version 2.15.0.).

0.46³³. This means that there is a 46% chance that the distribution of RAZ- vs. RAS- across the submeanings is random. In other words, there is no statistically significant contrast in preferences of the two allomorphs for particular submeanings. This suggests that semantics plays no role in this distribution of the two allomorphs.

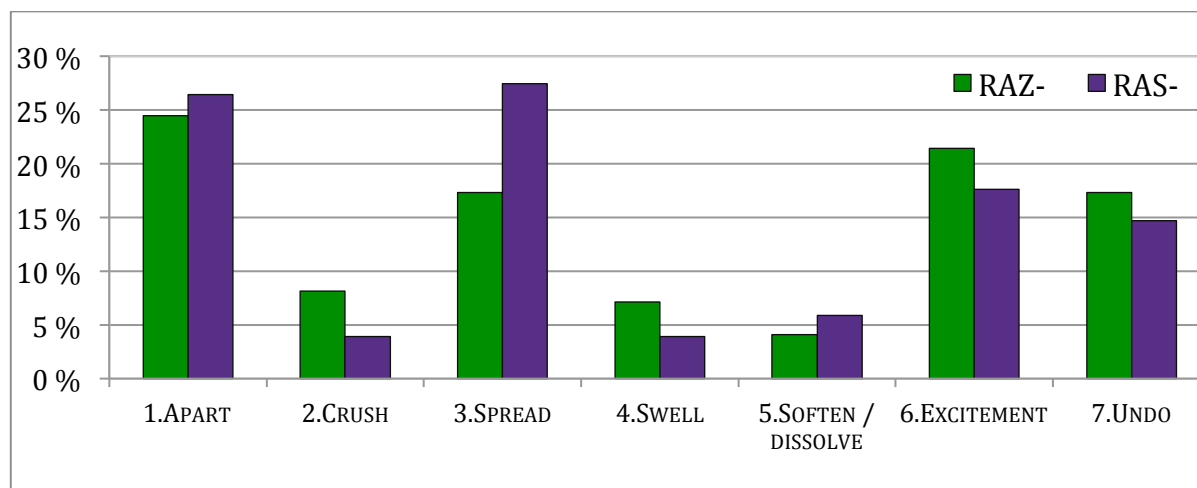


Figure 2: Radial Category Profiles of RAZ- and RAS-.

In order to address the impact of more than one factor on the distribution of RAZ- and RAS- and compare the role of semantic and phonological factors we need a more sophisticated statistical model. In this study I use the combined Classification Trees and Random Forests model. This model is comparable with logistic regression in accuracy (Baayen et al. 2013: 287) and is applied for multifactorial analysis in a large range of scientific fields (Strobl et al. 2009; for details on the model see Chapter 8).

The predicted variable in our analysis is Allomorph which is a categorical variable and has two values: RAZ- and RAS-.

The predictors (factors) can be subdivided into phonological and semantic:

Phonological factors: **SonorityOnset** (sonority of the initial phoneme of the base) with four values: voiceless obstruent, voiced obstruent, sonorant, and vowel;

SimpleOrClusterOnset (onset type of the base) with three values: simple, cluster, vowel;

Semantic factors: **Semantics** (submeaning of the prefix) with seven values that correspond to seven submeanings of RAZ- described above;

Metaphor (concrete or figurative facet of prefix submeaning in a given verb) with two values: metaphorical (m) and non-metaphorical (n);

³³ Recall that the largest p-value typically recognized as statistically significant is 0.05 (Cantos Gómez 2013: 49; Baayen 2008: 188). The values in the present analysis were calculated by the statistical software package R freely available at www.r-project.org. The code for the statistical analysis can be found at <http://hdl.handle.net/10037.1/10078> (file "R script RAZ"). Note that due to the small values of the subcategory 5. SOFTEN / DISSOLVE, the program produces a warning message. Exclusion of this subcategory from the analysis yields a very similar result without a warning: X-squared = 5.4, df = 5, p-value = 0.4.

PerfType (correlation type between the lexical meaning of the base and the submeaning of the prefix) with two values: Natural Perfectives (NP) and Specialized Perfectives (SP).

The graphical outcome of the Classification Tree analysis³⁴ is the plot given in Figure 3. The computational method employed for building Classification Trees is recursive partitioning (Strobl et al. 2009).

This means that a Classification Tree represents an algorithm that recursively subdivides the data into smaller sets and further subsets. The subsets of data are visualized as rectangles. The circle nodes of subdivisions are those factors that contribute to the model, showing that they can subdivide the data in a statistically significant way (Strobl et al. 2009: 325-334; Baayen et al. 2013: 265-66). After calculating all possible Classification Trees, the model chooses the most optimal tree, which visualizes how the data fed into the model should be partitioned in the most optimal way, having tested all the factors and their interactions to predict as much as possible of the dependent variable outcome. The most optimal Tree for the RAZ- vs. RAS- dataset is presented in Figure 3.

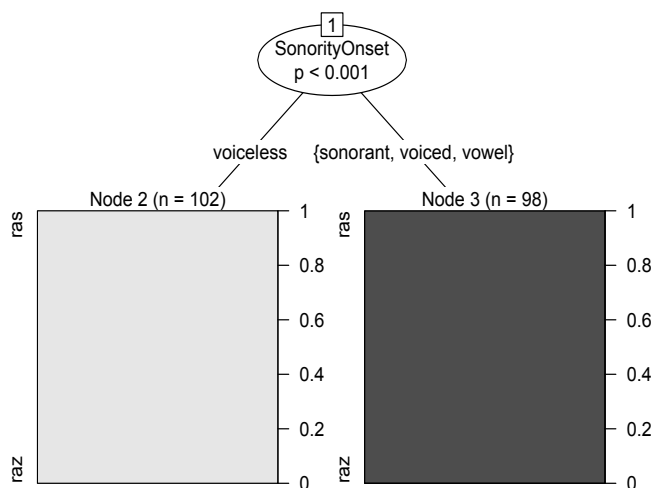


Figure 3: Classification Tree: verbs prefixed in RAZ- and RAS-.

Note that this Tree consists of only one split which subdivides the entire body of data into two groups – a group of verbs where the prefix is adjacent to a voiceless onset, and another group of verbs where the prefix is adjacent to initial sonorants, voiced obstruents, and vowels in the simplex base. In other words, this Classification Tree shows that the only factor that successfully predicts the choice between RAZ- and RAS- is the sonority of the base onset. Note that after calculating all possible splits of data that can produce the outcome allomorph, this model preserves only one factor as a statistically significant predictor (with p-value < 0.001). The Tree thus eliminates all factors that are found insignificant in predicting the allomorph of the prefix. The sonority of the onset accounts for the entire dataset without any exceptions. Another important thing shown by the Tree is that the distribution of RAZ- and RAS- is absolutely complementary, given that there is not a single simplex base that can combine with both allomorphs of the prefix.

While the Classification Tree visualizes only those factors that contribute to the model, the Random Forest analysis produces a graph where tested predictors appear with respect to their relative importance. In Figure 4, we can see that the five phonological and semantic predictors of the allomorph are ranked along a scale according to their relative importance scores calculated by the model. The higher the

³⁴ The formula used for this analysis locates the dependent variable Allomorph to the left of tilde and all possible predictor variables to the right of tilde: Allomorph ~ SonorityOnset + SimpleOrClusterOnset + Metaphor + PerfType + Semantics.

bar that represents a factor, the higher the score of its importance in the choice of the allomorph. The scores of importance in Figure 4 show irrelevance of semantic factors for the distribution of RAZ- and RAS-.

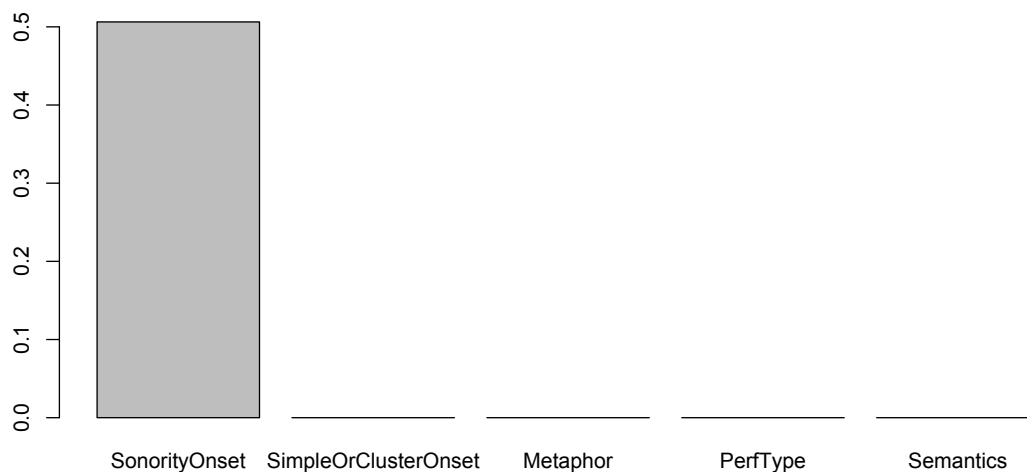


Figure 4: Variable importance scores for the phonological and semantic predictors of the choice between RAZ- and RAS-.

The statistical analysis differentiates between significant and insignificant factors for the distribution of prefixes across verbs. Recall that we included the semantics of the prefix as a factor that can possibly affect the distribution of RAZ- and RAS-. Taking semantics into account, the statistical model has shown that the difference between the distribution of RAZ- and the distribution of RAS- across the seven submeanings of these prefixes is not statistically significant. Thus, even though the distributions of RAZ- and RAS- are not identical, they lack any semantic contrast.

The only significant predictor of the choice of the allomorph here is the sonority of the onset base that the prefix attaches to. Thus, the alternation between RAZ- and RAS- is entirely phonologically conditioned. The lack of semantic contrast between RAZ- and RAS- as well as their perfect complementary distribution make them a good example of a standard allomorphic relation.

Moreover, the allomorphy of RAZ- and RAS- can be considered prototypical: the allomorphs exhibit not only 1) identical meaning (with no sub-morphemic semantic differences) and 2) perfect complementary distribution, but they are also 3) similar to each other in their formal shape; 4) phonologically conditioned; 5) motivated via a regular automatic phonological rule active in the language (in other words, RAZ- and RAS- are not suppletive allomorphs), and 6) etymologically related.

3.3 Vowel-zero alternation in consonant-final prefixes: #C- ~ #CV-

3.3.1 General remarks

Another phenomenon I want to discuss in this chapter is vowel-zero alternation in consonant-final Russian prefixes: #C- ~ #CV-, as it is called in (Townsend 1968: 76). By means of example consider the verbs *vz-letet* 'fly upward' (< *letet* 'fly') and *vzo-jti* 'walk upward' (< *idti* 'walk'), where the same prefix that denotes upward movement attaches in the shape of VZ- in the first case and in the shape of VZO- in the second case. The

vowel in the combination #CV- is always spelled as *o* but is unstressed and phonetically corresponds to [ə].

This pattern is prominent, productive, and can be found in all twelve consonant-final Russian prefixes that we observed in the previous study of voicing assimilation: OB-/OBO-, V-/VO-, VZ-/VZO-, VOZ-/VOZO-, IZ-/IZO-, NAD-/NADO-, NIZ-/NIZO-, OT-/OTO-, POD-/PODO-, PRED-/PREDO-, RAZ-/RAZO-, and S-/SO-.

This alternation is traditionally viewed as allomorphic variation in prefixes (Švedova et al. 1980: §851) conditioned by two rules. In both rules the vocalized allomorph of the prefix is triggered by the onset cluster of the root morpheme. This cluster CC+ is either 1) incompatible with the preceding coda consonant C₁ of the prefix (*C₁CC+ >> C₁VCC+) or 2) is itself an environment derived by means of a vowel-zero alternation (C∅C ~ CVC).

The first pattern can be illustrated by the already mentioned verb *vzo-jti* ‘rise’. Here the vocalized allomorph occurs due to the specific constraint of Russian phonotactics: the cluster **vzjt* on a morpheme boundary is ill-formed and has to be resolved. The grounds for the vocalized allomorph VZO- are thus purely phonological (phonetic) and the rule applies automatically.

The second pattern is more complicated. In the verb *vzobrat’sja* ‘climb up’ (< *brat’(sja)* ‘take’) the use of the allomorph VZO- cannot be caused by a problem with the cluster *vzbr* in Russian phonotactics, because there exist verbs like *vzbresti* ‘come into one’s mind’, *vzbryznut’* ‘splash’, and *vzbryknut’* ‘kick’ with the same consonant cluster *vzbr* on the morpheme boundary. What triggers the vocalized version of the prefix in *vzobrat’sja*_{PF} ‘climb up’ is the root *-br-* which alternates with *-ber-* (compare *vzberus’*_{PF.1PERSON.SG} ‘I will climb up’)³⁵. This means that the root itself contains a vowel-zero alternation. The zero in the consonant cluster of the root *-br-* triggers the vowel in the prefix. In other words, VZO- is conditioned not phonologically or phonetically, but rather morphophonemically, by means of the root morpheme which contains the vowel-zero alternation.

These two properties of environments that trigger a vowel in consonantal Russian prefixes are widely discussed in the literature in terms of rules (cf. Švedova et al. 1980: §851, Yearley 1995, Steriopolo 2007, Pesetsky 1979, Matushansky 2002)³⁶. In

³⁵ The alternation of a vowel and a zero in Russian is often referred to as a phenomenon of fleeting, or mobile, vowels (“beglye glasnye”). This alternation is a historical reflex of the Proto-Slavic alternation of *o* and *e* with so called jer vowels which disappeared in weak positions: e.g. compare Old Russian *beru* – *bьrati* ‘take’ (Vasmer 1971: v.1, 159) >> Modern Russian *beru* – *brat’* ‘take’. However, as we see in the verbs *vzobrat’sja*_{PF} ‘climb up’ – *vzberus’*_{PF} ‘I will climb up’ – *vzbirat’sja*_{IPF} ‘climb up’, the zero-vowel alternation extends beyond *e/o* ~ ∅ and involves the vowel *i* which does not come from a yer vowel and is usually associated with iterative verbal forms. The modern academic dictionary of Russian orthography (Lopatin 2007) suggests that the alternating vowels *e* and *i* should be viewed as fleeting vowels in 27 Russian roots where both of them alternate with a zero: *ber* ~ *bir* ~ *br* (e.g. *beru* – *sobiraju* – *sobrat’*), *der* ~ *dir* ~ *dr* (e.g. *deru* – *sdiraju* – *sodrat’*), *stel* ~ *stil* ~ *stl* (e.g. *stelit’* – *zastilat’* – *stlat’*), *zer* ~ *zir* ~ *zr* (e.g. *sozercat’* – *vzirat’* – *zret’*), etc. (cf. Sorokina 2012: 92 for discussion of *e~i~∅*).

³⁶ A great amount of attention is devoted to the issue of directionality in the vowel-zero alternation: does it represent an epenthetic vowel inserted into a cluster (#C- >> #CO-CC) or is it a deletion of an “underlying” vowel that produces a consonant-final prefix (i.e. #CЪ-CVC >> #C-CVC)? Looking at the surface forms of the prefixes, it is more natural to analyze the vowel in VZO- and RAZO- as epenthesis, because the default, neutral and most frequent forms of these prefixes are consonant-final: VZ- and RAZ-. Still, many scholars assume that it is the forms VZ- and RAZ- that are secondary and that they are produced via vowel deletion from VZO- and

my analysis, I stick to the surface representations of the prefixes and alternations available to learners from the surface input, staying away from stipulations on the underlying level. I also suggest that the alternation that we observe in prefixes should be called a *zero-vowel alternation* rather than a *vowel-zero alternation*, since the major allomorphs of the prefixes in question are consonant-final. Despite this fact, I follow the established tradition and adopt the latter term from Townsend (1968: 76).

What is more crucial for this study is that the allomorphic status of the vowel-zero alternation in consonant-final prefixes is not without problems.

First, in some prefixes this alternation is not entirely phonological but involves semantic differences. In particular, this concerns the pairs S-/SO- and V-/VO-. A seminal account of Russian morphophonology argues that phonologically conditioned vocalized allomorphs SO- and VO- should be distinguished from homonymous SO- and VO- that are not conditioned phonologically and constitute distinct morphemes (Itkin 2007: 230). According to Itkin, the prefix VO- found in factitive verbs of the type *vo-X-it* ‘make X’ (e.g. *voplotit* ‘embody’, *vočelovečit* ‘materialize in human body’, *vocarit’sja* ‘enthroned oneself, take the throne’) is morphologically distinct from the morpheme V- ‘IN’ (e.g. *vbežat* ‘run in’). Similarly, the prefix SO- that denotes concomitance of activity in verbs like *sopereživat* ‘share one’s worry’ (< *pereživat* ‘worry’) was argued to represent a separate morpheme (Itkin 2007: 230). I address this issue in Chapter 4.

Second, in some prefixal pairs the two allomorphs sometimes overlap in their distribution, being able to attach to the same simplex bases. This violates the criterion of perfect complementation. Illustrative examples are *o-zlit’(sja)* vs. *obo-zlit’(sja)* ‘embitter, make angry’, *s-kryt* vs. *so-kryt* ‘conceal’, where the prefixes O- and OBO- in the first case and S- and SO- in the second case appear as competing candidates that can attach to the same verbal stem.

Third, there are exceptions to the rules of distribution of the consonant-final and the vocalized versions of a prefix. For example, instead of the expected allomorph IZO- we have IZ- in the verbs *izbrat* ‘select’ and *izgnat* ‘chase away’. Similarly, the verbs *vozzvat* ‘appeal’ and *vozzrit’sja* ‘look at’ are formed by VOZ- instead of the expected VOZO-. Another example is the prefix RAS- in the verbal form *raspnu* ‘crucify (1person.sg)’ instead of the expected RAZO- (Itkin 2007: 229). On the other hand, sometimes a vocalized allomorph appears in those contexts where it is not expected: *so-beru* ‘bring together-1person.singular’, *so-zovu* ‘call together-1person.singular’. Unexplained is the use of OBO- in *obokrast’_{PF}* ‘rob’ and *obo-znat’sja* ‘take someone for someone else’, because in *obkradyvat’_{IPF}* ‘rob’ and *ob-znakomit’sja* ‘get acquainted’ the same clusters *bkr* and *bzn* on the morpheme boundary are not problematic and do not trigger vocalization of the prefix.

Thus, we must conclude that the vowel-zero alternation in Russian prefixes, although regularly attested, under closer examination does not fully fit into the definition of standard allomorphy: it is complicated by possible semantic differences, distributional overlaps and deviations from established rules. Moreover, the

RAZO- (Yearley 1995: 555; Gouskova 2012; Gouskova & Becker 2013). Because there are two yer-vowels, one front and one back, postulating an epenthesis does not make it possible to predict what kind of vowels alternate. The deletion analysis is in harmony with general historical facts, since the alternation in Modern Russian arose from the deletion of yers in certain contexts. However, in case of the prefixes that end in *z* it is problematic to postulate an underlying yer vowel at least because historically they did not contain a yer (Kaverina 1999). While the discussion of underlying representations may be interesting in itself, I will not discuss it further.

conditioning of the vocalized allomorphs of consonant-final prefixes can be only partially explained in terms of phonology because they involve morphophonemic dependencies on the shape and structure of the adjacent root morphemes. I will now turn to a specific case of the vowel-zero alternation observed in the pair RAZ- ~ RAZO- and examine how a statistical account can capture subtle motivations for their distribution.

3.3.2 Case study 2: RAZ/S- ~ RAZO-

In this case study I continue investigating the prefix RAZ- ‘APART’, now turning to its vocalized allomorph RAZO-. The data examined consists of 210 verbs, including the database of verbs in RAZ- and RAS- from the first case study, plus ten perfective verbs prefixed in RAZO-³⁷. They are collected from the Russian National Corpus based on the principles described above in 3.2.2.

In all ten verbs the allomorph RAZO- occurs exclusively in front of consonant clusters. Note that the non-vocalized allomorphs RAZ- and RAS- can occur in front of a cluster too: e.g. *razgrabit’* ‘plunder’ (< *grabit’* ‘rob’), *raspleskat’* ‘spill’ (< *pleskat’* ‘splash’). However, the consonant clusters that trigger the allomorph RAZO- bear an additional characteristic – they belong to root morphemes with a vowel-zero alternation. As an example, consider the perfective verb *razo-slat’* ‘distribute’ (< *slat’* ‘send’). Its imperfective counterpart verb is *ras-sylat’*: here the consonant cluster *sl* in the root morpheme is broken up by a vowel *y*, and therefore the prefix uses its non-vocalized allomorph RAS-. Yet if the root morpheme does not contain the vowel-zero alternation that we find in *razo-slat’* ~ *ras-sylat’*, the prefix can combine with the adjacent cluster *sl* without any vowel, as in *ras-slyšat’* ‘hear distinctly’ (< *slyšat’* ‘hear’) and *ras-sledovat’* ‘investigate’ (< *sledovat’* ‘follow’). Similarly, the five verbs in (1) are perfectives where the allomorph RAZO- is triggered by the onset consonant clusters *dr*, *gn*, *br*, and *rv* which contain a zero that alternates with a vowel within the same root morphemes:

- | | | |
|--|-----------------|--|
| (1) <i>razo-drat’</i> _{pf} – <i>raz-dirat’</i> _{ipf} | ‘tear apart’ | but <i>raz-drobit’</i> |
| <i>razo-gnut’</i> _{pf} – <i>raz-gibat’</i> _{ipf} | ‘straighten up’ | but <i>raz-gnevat’</i> (<i>sja</i>) |
| <i>razo-gnat’</i> _{pf} – <i>raz-gonjat’</i> _{ipf} | ‘disperse’ | but <i>raz-gnevat’</i> (<i>sja</i>) |
| <i>razo-brat’</i> _{pf} – <i>raz-birat’</i> _{ipf} | ‘dismantle’ | but <i>raz-brosat’</i> , <i>raz-brestis’</i> |
| <i>razo-rvat’</i> _{pf} – <i>raz-ryvat’</i> _{ipf} | ‘tear apart’ | * <i>raz-rv...</i> |

Otherwise the same clusters are compatible with RAZ- and RAS- and do not trigger the epenthetic vowel in RAZO- (although the cluster *rv* is not attested with the prefix RAZ-). Historically the clusters in these root morphemes contained a yer vowel that alternated with a full vowel in other forms of the paradigm. After the so called *fall of the yers*³⁸, the yer-vowels in weak positions were dropped, while the alternation of the root shapes remained. The absence of an overt vowel in particular consonant clusters is thus motivated historically by the fall of the yers. As a result, in Modern Russian from the overt shape of a consonant cluster alone one cannot predict whether RAZO- or RAZ/S- is used. Yet the reason for the occurrence of RAZO- is not fully opaque from the perspective of the Modern Russian phonology either: the vowel-zero alternation is

³⁷ I suggest that the verb *razočarovat’*(*sja*) contains the allomorph RAZ- but not RAZO- and is formed from the perfective verb *očarovat’*(*sja*). Note that five other verbs with č-initial stem attest the allomorph RAS-: *rasčuvstvovat’**sja*, *rasčlenit’*(*sja*), *rasčistit’*, *rasčesat’*, *rasčekanit’*.

³⁸ The sound change in the history of the Slavic languages, cf. Townsend & Janda 1996: 73-74.

available in the input of a learner. One does not need to postulate morphemes with underlying yer-vowels for Modern Russian because the vowel-zero alternation can be observed in the related modern forms and words.

So far I have looked at the verbs with RAZO-, where the vowel-zero alternation in the root can be established by comparing the forms within a paradigm. In two other verbs the same alternation can be found beyond the paradigm in related words. Consider the verb *razo-gret* 'warm up' (< *gret* 'heat'), related to *razgoret'sja* 'flame up' and *goret* 'burn'. The same cluster *gr* without vowel-zero alternation does not trigger RAZO- as evidenced by *raz-gryzt* 'gnaw apart', *raz-gruzit* 'unload', *raz-gromit* 'crash', *raz-grafit* 'draw columns', *raz-graničit* 'demarcate', and *raz-grabit* 'devastate'. The second verb is *razo-zlit* 'make angry', where the vowel-zero alternation in the root can be established via comparison of *zlit* 'make angry' with the adjective *zloj* 'angry' and its short predicative form *zol* 'angry'. This goes in line with the idea that vowel-zero alternation is a property of particular morphemes that form a sublexicon in Modern Russian (see Gouskova 2012 for discussion).

Another verb with RAZO- is *razo-mlet* 'grow languid'. From the modern Russian perspective the vowel-zero alternation in the root of this verb is hardly detectable. The simplex *mlet* denotes 'be overcome with delight, grow numb' and is etymologically related to the adjective *medlennyj* 'slow' (Vasmer 1971), although synchronically their semantic connection is rather obscure. Vasmer relates it to Proto-Slavic **mьdbl*, the source of Old Church Slavonic *mьdlěti*, Russian *mlet*, Ukrainian *mliti*, Belarusian *mlecь*, as well as corresponding verbs in other Slavic languages. Given that the vowel-zero alternation in *mlet* is not synchronically available, the epenthesis in RAZO- can be well explained by incompatibility of *z* with the adjacent cluster *ml*, since the combination *zml* is not attested in Modern Russian (**raz-ml...*). So, the use of RAZO- in this case can be explained phonetically rather than morphophonemically.

The last verb with RAZO- in the database is *razo-jtis* 'walk apart, split', where the epenthesis in the prefix is triggered by the incompatibility of *z* with the cluster *jt*. The form *razo-jtis* results from **raz-id-ti-s'*, which underwent devoicing of *d*, simplification of the geminate consonant *tt* into *t*, and reduction of *i* into *j*. In other words, RAZO- in this verb should be explained by phonetics.

Summing up, the ten verbs that exhibit the allomorph RAZO- can be accounted for by the two patterns previously discussed in 3.3.1. In eight verbs the epenthetic allomorph of the prefix occurs due to the fact that the root morpheme with an initial consonant cluster contains a vowel-zero alternation. As I showed, this alternation is available overtly in the forms of the same word or in words that are synchronically clearly related. In other words, in eight verbs the vocalization of the prefix is motivated by morphophonemic alternation in the root morphemes. In the remaining two verbs the allomorph RAZO- occurs due to phonological (phonetic) reasons, namely because of incompatibility of the coda consonant of the prefix with the onset cluster of the root.

Now, parallel to the first case study, let us test whether there is any semantic motivation for the choice of the vocalized allomorph of the prefix. In other words, how does the semantics of RAZO- fit into the model of polysemy that I outlined earlier in this chapter? Table 3 summarizes the distribution of the three allomorphs RAZ-, RAS-, and RAZO- across their submeanings:

Submeaning	# of verbs with	RAZ-	RAS-	RAZO-	Total:
1.APART		24	27	5	56
2.CRUSH		8	4	0	12
3.SPREAD		17	28	1	46
4.SWELL		7	4	0	11
5.SOFTEN/DISSOLVE		4	6	1	11
6.EXCITEMENT		21	18	2	41
7.UNDO		17	15	1	33
Total:		98	102	10	210

Table 3: Distribution of verbs with RAZ-, RAS-, and RAZO- across prefix submeanings.

Although in 50% of attested verbs RAZO- contributes the meaning ‘APART’, the data is too sparse (only ten verbs in RAZO-) for the chi-square test. The sparsity of data can be overcome in the Classification Tree and Random Forest analysis. It produces the most optimal Classification Tree presented in Figure 5 which accounts for all three allomorphs of the prefix RAZ-.

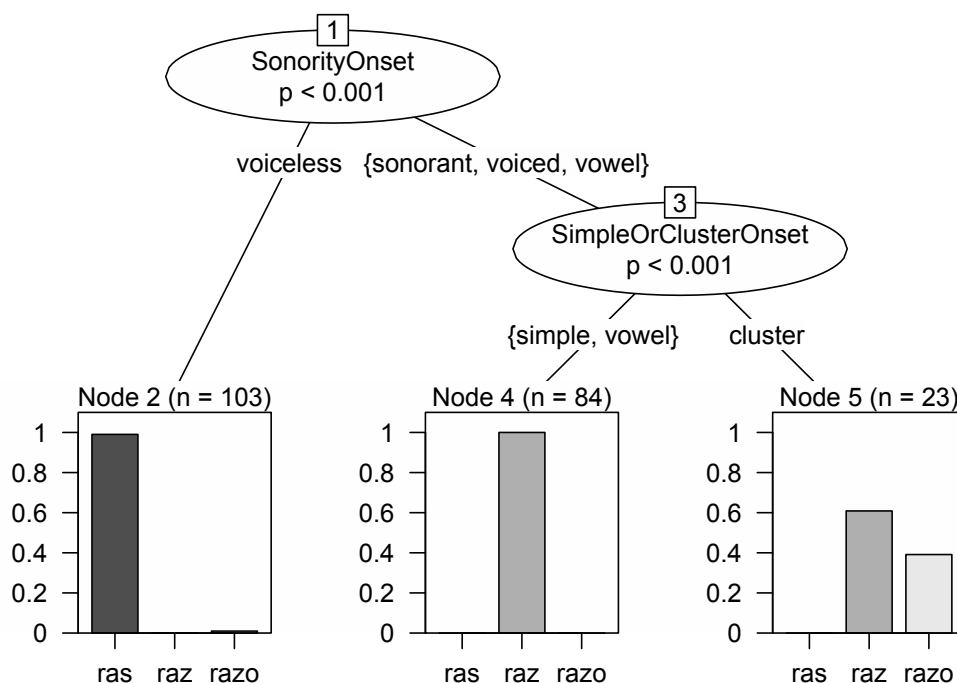


Figure 5: Classification Tree: verbs prefixed in RAZ-, RAS-, RAZO-.

The first split of the Tree highlights the sonority of the root’s onset as the most powerful predictor of the prefix shape. It reveals a tendency not mentioned in earlier accounts of the prefix RAZ-: if the root starts with a voiceless obstruent, it is more likely to attach the devoiced allomorph RAS- rather than the vocalized allomorph RAZO-. In other words, the occurrence of RAZO- concerns almost exclusively roots with voiced complex onsets. There are forty verbs with a complex onset in the root with a voiceless initial obstruent. All but one of them attach RAS- but not RAZO-. The only verb in RAZO- with a voiceless onset for the simplex base is *razoslat* ‘distribute’.

The second split of the Tree is made in the subset of verbs where the root starts in a vowel, sonorant, or a voiced obstruent. The decisive factor in this subset is the complexity of the onset, whether it is a cluster or a single consonant or a vowel. As we

see from the central graph in Figure 5, all 84 bases with simple onset or no onset attach RAZ-. By contrast, the group of 23 verbs with initial consonant clusters in the root remain problematic, because both RAZ- and RAZO- are possible here and the choice between them cannot be predicted by the shape of the root-initial consonant cluster alone. For this reason, the second phonological predictor (SimpleOrClusterOnset) is ranked at zero level in the scale of variable importance in Figure 6.

In addition, the Random Forest model shows that none of the tested semantic factors is involved in the distribution of the tree allomorphs.

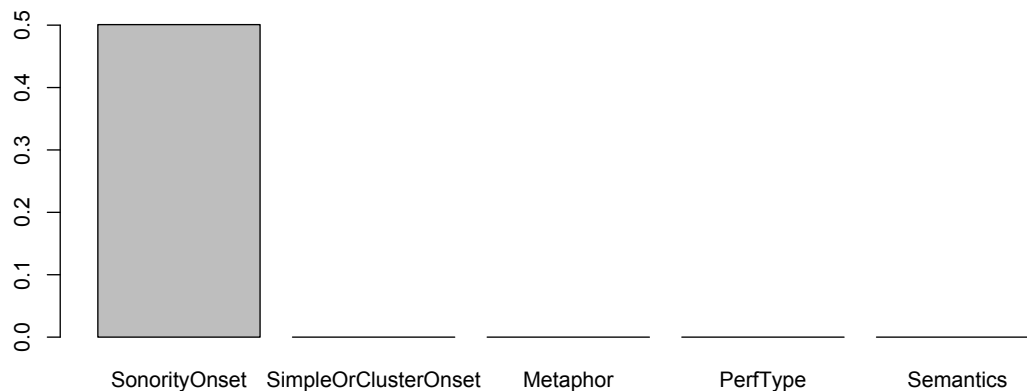


Figure 6: Importance scores for predictors of the choice between RAZ-, RAS-, RAZO-.

Let us now focus specifically on the opposition between the vocalized allomorph (RAZO-) and the non-vocalized allomorphs (RAZ-/RAS-). Taking into account the additional characteristics of the initial clusters of the root that trigger RAZO-, I include two additional factors in the model:

AlternationInRootCluster: does the onset cluster of the root contain the vowel/zero alternation? Three values: Yes (8 verbs), No (55 verbs), No cluster (147 verbs);

ClusterPossibleWithZ/S: is the onset cluster of the root compatible with the adjacent consonant of the prefix coda? Three values: Yes (60 verbs), No (3 verbs), No cluster (147 verbs).

In order to avoid collinearity³⁹, we exclude the factor SimpleOrClusterOnset from the model, because onset type of the root is already accounted for by the two new factors. The outcome of this new model is the Classification Tree shown in Figure 7.

³⁹ In statistics, *collinearity* of variables refers to an unwanted condition when the variables included in a model are highly correlated with one another. “As predictors become increasingly correlated, the statistical model becomes more unreliable.” (Cohen et. al 2003: 419).

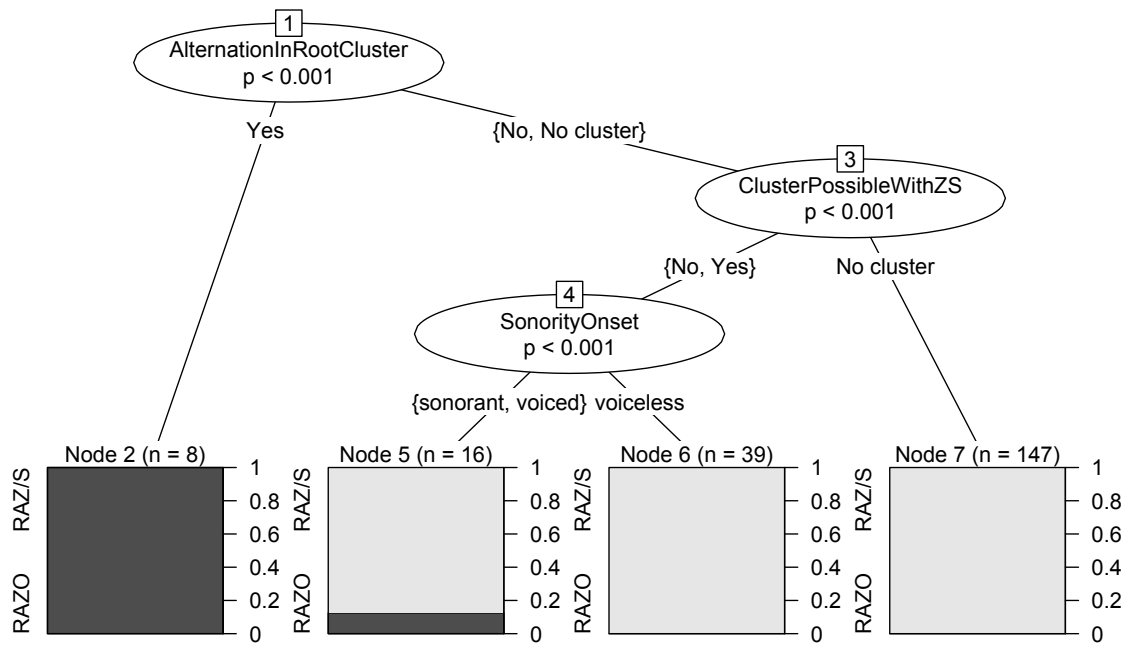


Figure 7: Classification tree: verbs prefixed in RAZ/S- vs. RAZO-.

The Tree indicates that the use of the vocalized allomorph RAZO- is determined by the interaction of three factors – vowel-zero alternation in the root cluster, compatibility of the initial root cluster with preceding z or s, and sonority of the onset. The first factor creates the major split of the Tree which accounts for eight lexemes with the vocalized allomorph visualized by the leftmost rectangle. The second factor creates the second split which captures 147 lexemes with simple onsets that attach non-vocalized allomorphs – either RAZ- or RAS- – depicted in the rightmost rectangle. Note that, as indicated in the Tree, it is more important for allomorph prediction whether the verbal base has a simple or complex cluster rather than whether its cluster is compatible with a preceding fricative obstruent or not. At last, the third factor, sonority of the onset obstruent, subdivides the remaining verbs into stems with voiceless stems that never trigger vocalization of the prefix (node 6) and stems that begin in sonorants or voiced obstruents. Among this latter group of 16 lexemes there are two that trigger RAZO-, namely *razomlet* ‘grow languish’ and *razojtis* ‘split’. In both of these verbs the stem is not compatible with the preceding fricative.

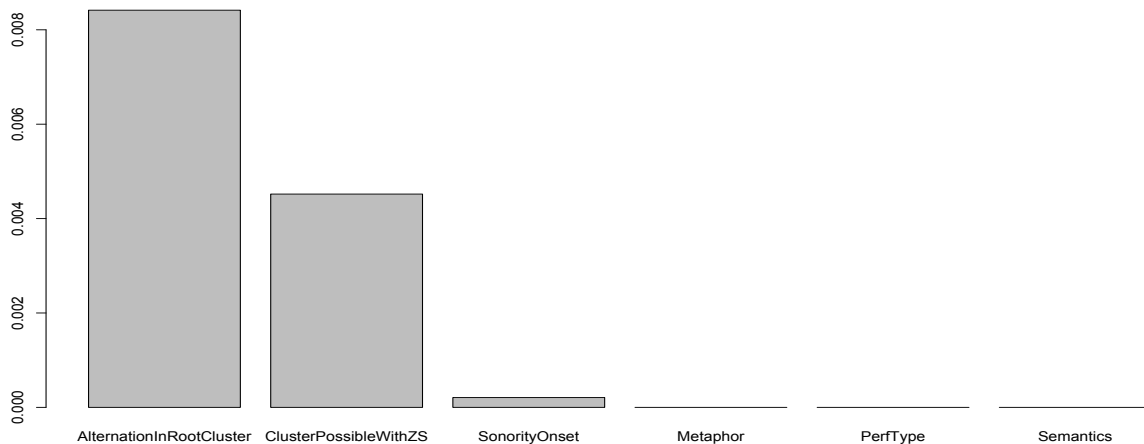


Figure 8: Importance scores for predictors of the choice between RAZ/S- vs. RAZO-.

Another important observation in this case study is that the first and major factor **AlternationInRootCluster** accounts for 80% of verbs with the vocalized allomorph RAZO-. Therefore, if we eliminate all other predictors, this factor can still predict vocalization and non-vocalization of the prefix correctly in 208 lexemes (in a total of 210 verbs). This is illustrated in the Tree shown in Figure 9. This distribution of verbs across RAZO- and RAZ/S- depending on the vowel-zero alternation in the initial cluster of the root is nearly complementary. Figure 9 thus resembles Figure 3 for RAZ- vs. RAS-, which complement one another perfectly in different phonological contexts. By contrast with RAZ-/RAS-, the vocalized RAZO- is triggered in most cases by idiosyncratic morphophonological properties of the adjacent root rather than by phonological compatibility with its onset.

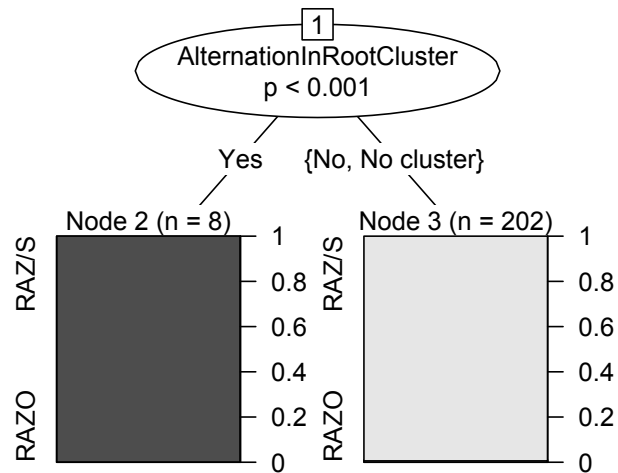


Figure 9: Classification Tree: RAZ/S- vs. RAZO-.

Crucially, the statistical account shows that nuances of semantic content do not affect the choice among the three allomorphs of the prefix RAZ-. Even though RAZO- is more frequently attested with the submeaning ‘APART’ (recall Table 3), in overall distribution this is not statistically significant and does not indicate a semantic contrast. The distribution of the three allomorphs across the seven submeanings is not identical, but the differences are not significant and this is what matters for establishing the allomorphic status of competing variants.

3.4. Conclusions

By contrast with the remaining chapters of this dissertation that explore unusual, non-trivial, and abnormal allomorphic relations, in this chapter I have looked at what is considered “normal”, or Standard, allomorphy. The contribution of this chapter is twofold.

First, I showed that statistical analysis is able to capture Standard (or Prototypical) allomorphic relations exemplified by means of the pair RAZ- and RAS-. The first case study establishes a baseline for evaluation of the results of all other case studies in this dissertation. The allomorphy of RAZ- and RAS- is characterized by perfect complementary distribution. Moreover, even though both prefixes can express seven different senses, they overlap in all of them and the difference in their distributions across verbs according to these senses is insignificant. This suggests that both RAZ- and RAS- share the same network of polysemy, where their distributional profile is the same. Compliance to both distributional and semantic criteria make these prefixes Standard Allomorphs. Moreover, their formal similarity and the transparent devoicing mechanism that relates these prefixes makes them exemplars of Prototypical Allomorphy in the sense that I proposed in Chapter 2.

Second, in the study of vowel-zero alternation in Russian prefixes, I pointed out violations of complementary distribution and identical semantics in the case of S-/SO-, V-/VO-, and O-/OBO-. Vocalized allomorphs of these prefixes sometimes attach to the

same bases and develop semantic contrasts. In this light, variation of Russian prefixes created by vowel-zero alternation goes beyond Standard Allomorphy although it is traditionally described as such in the literature. However, not all prefixes that feature vowel-zero alternation are equally problematic. I explored the case of the vocalized allomorph RAZO- and found that its relationship with RAZ/S- approximates Standard Allomorphy.

Whereas conditioning of RAZ- vs. RAS- is purely phonological, RAZO- is a product of an inherited but no longer active phonological alternation. In Modern Russian, the vowel-zero alternation applies to a closed list of particular morphemes and belongs to their morphophonological properties. The presence of vowel-zero alternation in the verbal root is not an immediately available characteristic, yet I argue that it is detectable through paradigmatic relationship of surface forms. What complicates the distribution of the vocalized allomorph RAZO- is the interaction of three factors which account for phonological and morphophonological properties of the verbal stem adjacent to the prefix. Together, these factors sort out vocalized and non-vocalized allomorphs of the prefix. However, in terms of semantics, RAZO- is not different from its non-vocalized alternants, and from this perspective it is a Standard Allomorph. Both multifactorial conditioning of vocalization and the irrelevance of semantics have been detected in the statistical analysis.

Chapter 4

Morphological status of the prefixes S- and SO- 'downward, with'

4.1 Introduction⁴⁰

The prefix S- is the second (after PO-) most productive marker of perfective aspect in Modern Russian, out of eighteen other prefixes (Mizoe 2011: 91; Janda et al. 2013: 15; Endresen et al. 2012: 243). This prefix has attracted much attention in the scholarly literature, yet one crucial issue has been largely overlooked: how to account for the modern controversial status of the two historical variants S- and SO-? In this chapter I focus on how their semantic and distributional properties challenge the traditional theory of allomorphy and present the first corpus-based study of S- vs. SO- based on a large dataset from the Russian National Corpus (www.ruscorpora.ru).

The puzzle of the prefix S- lies on the crossroads of phonology and semantics, strongly influenced by features inherited from its former use in Old Russian and Old Church Slavonic. The traditional account views all instances of S- and SO- as allomorphic variation, parallel to other consonantal prefixes that have vocalized allomorphs (recall Chapter 3) like RAZ-/RAZO-, POD-/PODO-, etc. (Švedova et al. 1980: §851; Tixonov 1985: 22; Kuznecova & Efremova 1986: 575). However, there is evidence that some uses of SO- are not triggered by phonology and have a distinct meaning CONCOMITANT ACTION (e.g. *so-suščestvovat'* 'co-exist' < *suščestvovat'* 'exist') that cannot be expressed by S- (Itkin 2007: 230). Moreover, SO- in this meaning is highly productive beyond the verbal domain, namely in nominal derivational morphology (e.g. *so-avtor* 'coauthor' < *avtor* 'author', *so-krizis* 'co-crisis' < *krizis* 'crisis'; cf. Valedinskaja & Golanova 2007: 179-181). Does this mean that SO- is a distinct independent morpheme that exists on its own? Or are these specific properties of SO- rather non-trivial sub-morphemic differences? How can the traditional theory of allomorphy account for this data?

The objective of this case study is to present a pair of related morphological markers that have a complex mutual status that is difficult to capture in absolute terms. I show that both traditional criteria of Standard Allomorphy (i.e. identical meaning and complementary distribution) are rather difficult to assess. First, instead of one meaning, in the case of S- and SO- one has to account for a whole network of meanings. Second, instead of a perfect complementary distribution we discover tendencies and overlap. Moreover, instead of straightforward phonological conditioning we face a mix of phonological, semantic, stylistic, and historical factors involved in the distribution of S- and SO- across verbs and other parts of speech.

⁴⁰ The results of the present study were presented at the conference "Cognitive Linguistics in the Triangle: Slavic and Beyond" at the University of North Carolina at Chapel Hill and at the 19th Congress of Nordic Slavists in Bergen in 2013. I am deeply grateful to both audiences for fruitful discussions.

In this chapter, I first establish the set of historical reflexes of the Common Slavic prefix *Sъ- attested in Modern Russian (4.2) and then I zoom in on the problem of S- vs. SO- and relevant previous accounts of their relations (4.3). The remainder of the chapter presents an alternative view built upon a usage-based study of Modern Russian verbs prefixed in S- and SO-. I describe data collection and methodology in 4.4 and then evaluate data with respect to the two traditional criteria of allomorphy: identical semantics (4.5) and complementary distribution (4.6). In 4.5, I offer a semantic analysis of the two prefixes in Modern Russian, where I argue that all their meanings are interrelated and form a structured network of polysemy. In order to compare the two prefixes according to all their semantic uses I propose a radial category and show that it accounts for the semantics of both S- and SO-. In 4.5.5, the two prefixes are compared in terms of their qualitative and quantitative profiles within a single network of meanings. Section 4.6 examines the distribution of the prefixes and shows that instead of straightforward phonological conditioning the choice of the prefix is driven by a mixture of factors. In 4.7, I summarize my findings from both semantic and distributional analyses with respect to empirical contribution, theoretical applications, and further research. In particular, I discuss the pros and cons of a monomorphemic account of S- and SO- and propose a solution in terms of a Non-Standard Allomorphy.

4.2 Historical variants of the prefix S-

Historical allomorphs of the prefix S- in Russian include not only S- and SO- but also SON-, SN-, and SU-. This is due to the historical fact that the source Proto-Slavic prefix *Sъ- (*SъN-) used in verbal derivatives was parallel to Common Slavic *sǫ- employed in nominal word-formation: nouns and adjectives (Vasmer 1971: v.3, 539). In this study, I focus primarily on S- and SO-.

The variant SU- is a historical reflex of the nominal Proto-Slavic prefix *sǫ-. It can be observed in Russian nouns like *suprug* ‘spouse’, lit. ‘together-chained’, *sustav* ‘joint’, as well as substandard *sused* ‘neighbour’, lit. ‘with-sitting’.

The variant SN- is historically attested in the verb *snjat* ‘take off (clothes)’ (<*sъn-jęti, cf. Vasmer 1971: v.3, 539). The variant SON- can be found in the noun *sonm* ‘assembly, a large group’ borrowed from Old Church Slavonic, where it was spelled as *сѣньмѣ* (Vasmer 1971: v.3, 717). In this word, the prefix is not analyzable from a synchronic perspective, because it has become fused with the root.

Note that the historical variants of the morpheme S- ‘downward, with’ should be distinguished from another morpheme S- found in the Russian words *zdorovyj* ‘healthy’, *smert* ‘death’, and *sčastje* ‘happiness’ (compare Old Church Slavonic forms of these words *сѣ-dravъ*, *сѣ-mrѣтъ*, *сѣ-častie*). In these words the morpheme *Sъ- comes from a different Indo-European source and is related to the Sanskrit formant SU- ‘good, welfare’ (Vasmer 1971: v.3: 791 ‘*xorošo, blago*’; Krasuxin 2008: 79, 80, 84). Another related parallel is the Greek formant EU- ‘good’ (compare borrowed to English *euphoria*, *euphemism*), which is the result of dropping the consonant *s* and the laryngeal from the Indo-European source root **h₁(e)su* (Beekes 2010: 484). Therefore, the Russian adjective *zdorovyj* ‘healthy’ (compare Proto-Slavic **сѣ-dorvъ*, Old Russian *сѣdorovie*) originally denoted ‘made out of a good tree’, where the root **dorvъ* is related to the

Modern Russian noun *derevo* ‘tree, wood’⁴¹ (Vasmer 1971: v.2, 90; Krasuxin 2008: 84). Likewise, the noun *sčastje* ‘happiness’ (Proto-Slavic **sъčestьje*) is not formed by the prefix S- ‘downward, with’, but is a combination of Sʙ- ‘good’ with **čestь* meaning ‘good lot, destiny’ (Vasmer 1971: v.3, 816) (as opposed to Old Russian *zlo-častie* ‘bad lot’). Finally, the semantically least obvious noun *smertʹ* ‘death’ originally had the meaning ‘good, virtuous, natural death (as opposed to death brought about by illness or murder)’⁴² as a combination of Sʙ- ‘good’ with **mьrtь* ‘death, human’ (Vasmer 1971: v.3, 685; Krasuxin 2008: 80). These are well accepted etymologies. Henceforth I will set Sʙ- ‘good, well’ aside from the analysis of Sʙ- ‘downward, with’ as a different and unrelated morpheme.

4.3 Problematic data and previous accounts

In the scholarly literature one can observe two opposed views on the morphological status of S- and SO- in Modern Russian. The first and most traditional account presented in the Academy Grammar considers all instances of the prefix SO- allomorphic variation of the morpheme S- (Švedova et al. 1980: §851). Major dictionaries of Russian morphemes and word-formation adopt this single-morpheme account as well (Tixonov 1985: 22; Kuznecova & Efremova 1986: 575). Moreover, S- and SO- are listed as a single morpheme by Townsend (1968: 132) and Zaliznjak & Šmelev (2000: 83). In most recent large-scale studies of the Russian prefix S- it is also assumed that S- and SO- represent a single morpheme (cf. Dickey & Janda 2009; Makarova 2009; Makarova & Janda 2009; Janda et al. 2013: 97-100; Janda & Lyashevskaya 2013). Yet the relationship of the two prefixes is far from simple.

An alternative account of S- and SO- was proposed by Ilja Borisovič Itkin⁴³ (2007: 229-231), who brought attention to the heterogeneous conditioning of the vocalized variant SO-. Itkin draws a distinction between two different SO- variants in Modern Russian. In line with previous tradition, he agrees that SO₋₁ is a positional allomorph of the morpheme S-. In particular, SO₋₁ is a phonologically conditioned variant that occurs in front of certain stem-initial consonants like *z*, *šč*, and clusters *žC* and *sC*. The relevant examples here include:

- (1) SO- in front of *z*: *sozercatʹ* ‘contemplate, gaze, observe’, *soznatʹsja* ‘confess’

⁴¹ Krasuxin (2008: 84) also points out that the Russian *zdorovyj* ‘healthy’ corresponds to Lithuanian *sudrùs* ‘firm, solid’ (*plotnyj, krepkij*) and Old Hindi *sudru* ‘mighty tree (“*mogučee derevo*”)’. Vasmer provides semantic parallels for association of wood with health and strength: German *kerngesund* ‘absolutely healthy’ < *Kern* ‘seed, heartwood of a tree’; Latin *rōbustus* ‘made-of-oak-wood, strong, healthy’ < *rōbur* ‘oak wood’. Compare also Russian *drevnij* ‘ancient, aged’ and English *true, trust* as well as German *treu* ‘reliable, correct’ that come from the Old European cult of trees seen as an example of permanence, longevity, strength, as described in Frazer’s monograph “*The golden bough*” (Krasuxin 2008: 85).

⁴² Compare with Lithuanian *sāvo smerčiù miřti* ‘die with one’s own death’ and the Russian expression *svoja smertʹ* ‘one’s own, natural death’, where the reflexive pronoun *svoj* ‘one’s own’ is also etymologically related to the root *s(u)*- ‘good’ (Vasmer 1971: v.3, 686; Krasuxin 2008: 80).

⁴³ As mentioned by Itkin himself (2007: 229), his analysis is largely based on the unpublished work of A.S. Kas’jan, whose contribution to the investigation of the Russian prefix S- should be acknowledged. Yet I refer to the account published in Itkin 2007.

SO- in front of šč: *soščipnut'* 'pinch something off', *soščurit'sja* 'blink, squint'

SO- in front of the cluster žC: *sožmurit'* 'close one's eyes tight'

SO- in front of the cluster sC: *sosvatat'* 'make a match'; *soskoblit'* 'scrape off'

In addition, SO₋₁ is morphophonologically conditioned when it occurs in front of a number of stems which are inherited from Old Church Slavonic and exhibit vowel/zero alternation and liquid metathesis (e.g. *sokratit'* 'reduce'; recall the same type of conditioning of the vocalized RAZO- discussed in Chapter 3).

By contrast, the second SO₋₂ is not triggered by such phonological and morphophonological contexts and constitutes a distinct productive morpheme with a distinct meaning. This meaning is defined as 'togetherness, concomitance' (the Russian term *sovместnost'*). According to Itkin, SO₋₂ can be observed in verbs like (2):

- (2) *sopereživat'* 'share one's worry with someone' < *pereživat'* 'worry'
sočuvstvovat' 'sympathize with someone' < *čuvstvovat'* 'feel'

Itkin points out that this distinct morpheme SO₋₂ is often used in nominal derivation:

- (3) *soavtor* 'co-author' < *avtor* 'author'
sobrat 'fellow man' < *brat* 'brother'
sodokladčik 'co-speaker' < *dokladčik* 'speaker, presenter'
sovremennyj 'contemporary, up-to-date' < *vremja* 'time'

Therefore, according to Itkin's account, the important distinctive property of this morpheme SO₋₂ is that it is neither an exclusively verbal prefix nor a perfectivizing marker, as opposed to the morpheme S- (Itkin 2007: 230).

In addition, Itkin (ibid: 230) mentions that the prefix S- is likewise able to carry the meaning 'concomitance' and to attach to non-verbal simplex stems, as shown in (4), but such examples are rare and peripheral for S-.

- (4) *sputnik* 'companion, fellow traveler, satellite' < *putnik* 'traveler'
smežnyj 'adjacent, contiguous, sharing a border' < *meža* 'border, boundary'

To sum up, in Itkin's model the morpheme S- has two allomorphs S- and SO₋₁, whereas the morpheme SO₋₂ has only one morph SO₋₂.

The crucial contribution of Itkin's account consists in recognizing the heterogeneity of different uses of SO- across the Russian lexicon. However, there are two problematic issues that one has to be aware of. The first problem with this account is mentioned by Itkin himself: there exist examples of "abnormal" use of the prefix SO- that this model cannot account for. In particular, the presence of SO- can be explained neither by phonological or morphophonological factors (SO₋₁), nor by the semantics of concomitance (SO₋₂) in a large number of words. Some of them are listed in (5):

- (5) *sobljusti* 'adhere to a rule' < *soveršit'* 'commit, perform'
sodejannyj 'committed' < *sodrogat'sja* 'quake'
sokrytie 'concealment' < *sotvorit'* 'create'
sokrovišče 'treasure' < *sokrušit'* 'destroy'

sotrjasti ‘concuss, shake intensively’

Another deficiency of Itkin’s account is that the meaning of concomitance is understood in a broad sense. As a result, this meaning is not restricted to the “legitimate” examples like *sopereživat’* ‘share one’s worry’, but it also claimed to be present in words where the prefix rather refers to a centripetal motion of bringing things together, to one location. Therefore, among Itkin’s examples of the ‘concomitant’ meaning (ibid. 230), there are many words that represent another meaning of the prefix, namely ‘centripetal motion’:

(6) <i>soprikosnut’sja</i> ‘adjoin, come in contact with’	< <i>prikosnut’sja</i> ‘touch’
<i>soedinit’</i> ‘unite, connect’	< <i>edinyj</i> ‘single’
<i>sopričislit’</i> ‘add’	< <i>pričislit’</i> ‘add’
<i>sobrat’</i> ‘gather’	< <i>brat’</i> ‘take’
<i>soprjač’</i> ‘conjoin, forge together’	< (<i>za/pri</i>) <i>prjač’</i> ‘buckle;harness’
<i>svesti</i> ‘bring together’	< <i>vesti</i> ‘lead’
<i>složit’</i> ‘add up, add together’	< <i>klast’</i> ‘put’

Exactly this sense of ‘centripetal motion’ is well attested for the prefix S-. In other words, including the centripetal meaning into the notion of concomitance undermines the semantic distinction between the morphemes S- (S-/SO₁-) and SO₂ that Itkin advocates.

Townsend (1968: 132) does not differentiate between the meanings of ‘concomitance’ and ‘centripetal motion’ for the Russian S-/SO- either. In particular, under the category ‘together’ Townsend lists two types of verbs without indicating any differences: first, the verbs where the prefix refers to concomitant action like *sosuščestvovat’* ‘coexist’ (< *suščestvovat’* ‘exist’) and *sootvetstvovat’* ‘correspond’ (< *otvet, otvečat’* ‘respond’), and, second, the verbs where the prefix denotes centripetal motion like *sozvat’* ‘call together’ (< *zvat’* ‘call’) and *sovместit’* ‘combine with’ (< *vmestit’* ‘contain’).

In other words, Townsend’s semantic category ‘together’ corresponds to Itkin’s category ‘togetherness, concomitance’ (“*sovместnost’*”), and both scholars exemplify this category with verbs in both S- and SO-. Yet, Townsend does not question their relations in terms of a single morpheme, as opposed to Itkin, who proposes a morphemic split. Instead of claiming SO- to be a separate morpheme, Townsend attributes such uses to the Church Slavonic heritage of SO- as well as the correspondence of some Russian SO- to the West European prefix CON- (COM-, COL-), as in *sonasledovat’* ‘co-inherit’ (Townsend 1968: 77; 132).

This controversy in the previous scholarship suggests that the mutual status of Russian S- and SO- is far from straightforward and requires a thorough analysis of a comprehensive dataset. It becomes clear that although the two meanings of concomitance and centripetal motion are very close if not adjacent to each other, their distinction might be beneficial for a fine-grained analysis that aims to compare the two prefixes in terms of their semantics and distribution.

Apart from Itkin (2007), recognition of the separate morphemic status of SO- in Modern Russian can be also found, although less explicitly, in Zemskaja (2006: 32, 35), Isačenko (2003/1965: 149), and in the recent morpheme dictionary by Efremova

(2010)⁴⁴. Similarly, in a recent study, Valedinskaja & Golanova (2007) assume that the prefix SO- is a distinct morpheme. Valedinskaja & Golanova (2007: 180) argue that in all novel coinages prefixed in SO- this prefix carries the semantics of concomitance, and, moreover, it is exactly the monosemantic content that made SO- so productive in Russian word-formation.

In Valedinskaja & Golanova 2007, the argument about the monosemantic nature of Modern Russian SO- is largely built around an example of so-called “independent” use of the prefix (the use of the prefix without a simplex verb)⁴⁵, given in (7).

- (7) *“Na putjax SO-obščenija” – tak nazyvaet poët knigu stixov. V nazvanii – sostojanie duši, kotoraja iščet sobesednika, edinomyšlennika... V žizni dlja nee tol’ko i cenno, čto “SO-” – sočuvstvie, sopereživanie, sozvučie, soglasie* (Predislovie k knige stixov Sjuzanny Serovoj “Na putjax SOobščenija”, 2003).

“On the routes of COM-munication” – this is the title that the poet gives to her book. This title expresses the state of a soul that is in search for a conversation partner, someone who thinks the same way... The only thing valuable for her in life is “COM-” – compassion (lit. co-feeling), commiseration, consonance, consent.’

In (7), the prefix SO- is used in its concomitance meaning as an umbrella term that stands for all its derivatives. Moreover, it is followed by the sequence of nouns prefixed with SO- that foregrounds the morphemic structure of these derivatives. The abundance of such derivatives highlights what they all have in common, namely the semantic component contributed by the prefix SO-: the concomitant character of an activity (as in *sočuvstvie* and *sopereživanie* ‘compassion’) or a state (as in *soglasie* ‘consent’ and *sozvučie* ‘consonance’). Foregrounding of the prefix takes place also in the title of the book: capitalization of the prefix in the word *SOobščenije* ‘COMMunication’ allows for reanalysis of the somewhat bleached morphemic structure of this word and the expression *Na putjax SO-obščenija* ‘On the routes of COMMunication’ as well. As a result, the fixed expression which commonly refers to railway lines is transformed into a label for interpersonal conversation and exchange.

Valedinskaja & Golanova argue that the independent use of a prefix makes use of its most central and entrenched meaning, and for the prefix SO- this turns out to be the sense of concomitant action. According to their argument, because SO- is sometimes spelled with a hyphen (e.g. *so-krizis* ‘co-crisis’), SO- overcomes its prefixal status and enters a class of analytical adjectives⁴⁶, similarly to other prefixes like *anti-* ‘anti-’, *super-* ‘super-’, *psevdo-* ‘pseudo-’, and other.

⁴⁴ Efremova (2010: 545) provides a list of submeanings for S- and SO- suggesting that the concomitance meaning is attested exclusively for SO-, while the other five meanings can be expressed by both prefixes.

⁴⁵ This phenomenon is well described the literature (cf. Krongauz 1998: 38-41; Zemskaja 2006: 24) and can be compared to the classical example from the poetry of V. Xodasevič: *Perešagni, pereskoči, Pereleti, pere – čto xočeš’ – No vyrvis’...* ‘Step over, jump over, Fly over, **over** – what ever you want – But get out...’ (cited according to Krongauz 1998: 39; boldfaced by me – A.E.)

⁴⁶ Analytical adjectives in Russian do not decline, as in examples like *kartofel’ fri* ‘French fries’, *on-lajn apteka* ‘on-line drugstore’. For discussion of fuzzy boundaries between some prefixes and analytical adjectives see Marinova 2010; Panov 1971; Golanova 1998.

Summing up, accounts which suggest that SO- is a distinct morpheme have several arguments: 1) SO- is productive in one particular meaning of concomitant action (Valedinskaja & Golanova 2007), 2) this meaning is very uncommon or even impossible for S- (Itkin 2007), 3) the use of SO- does not depend on the phonological context, as opposed to S- (Itkin 2007), 4) SO- is frequently used beyond the verbal domain, namely in nominal derivations (as opposed to verbal S-) (Itkin 2007), 5) verbal aspect is irrelevant for SO-, by contrast with perfectivizing S- (Itkin 2007).

These observations are reasonable, and in my analysis I will not disprove any of them. Rather, I suggest that these characteristics of SO- are nevertheless compatible with an allomorphic relationship, if we consider the possibility of Non-Standard Allomorphy. My major argument is centered around the semantic account of polysemy, where I try to show that SO- and S- are closely related. However, before I turn to the semantic analysis, I should first discuss a crucial issue of prefixal semantics that to a large extent underlies the willingness to connect one form with only one meaning and interpret this pairing as a distinct morphological unit.

One should be aware that an important role in this discussion on the status of SO- has been played by the theoretical premises linguists have about the structure of linguistic meaning in general and the semantics of the prefix S- in particular. It is a well-known fact about the prefix S- that the meanings it can express in Russian are extremely different – so different that one can wonder how on earth they came to be expressed by the same prefix.

An account of the prefix S- that was very influential in its time was introduced in 1957 by Olga Sergeevna Axmanova in the book “Očerki po obščej i russkoj leksikologii”. Axmanova argues that the prefix S- exhibits homonymy with a clear contrast of several meanings⁴⁷: 1) ‘movement from different directions to a single point’, 2) ‘movement downward’, 3) ‘removal of something off a certain location’, etc. Interestingly, Axmanova focuses her attention on verbs that combine these meanings and reveal them in different contexts. In particular, among other similar examples, Axmanova points to the verb *skinut’* (< *kinut’* ‘throw’), which can refer either to ‘throwing object(s) down’ or ‘throwing objects together into a pile’ (Axmanova 1957: 150). Yet, the conclusion she makes from this comparison is that the coexisting meanings are homonymic, that is unrelated, and therefore belong to different morphemes S₋₁, S₋₂, S₋₃, etc. Indeed, ‘bringing something together’ and ‘moving something downward and/or away’ might seem to have nothing in common. Therefore, Axmanova argues that here we deal with homonymy of several unrelated morphemes rather than with polysemy or something else.

The idea of prefix homonymy was so popular and the meanings of S- seemed so distant that this prefix even became a typical illustrative example of this phenomenon. Even Isačenko, considering different senses of the prefix S-, namely ‘away’ (*s-bežat’ iz tjur’mj* ‘escape from prison’), ‘down’ (*s-bežat’ (s gory)* ‘run down the mountain’), and ‘together’ (*s-bežat’-sja* ‘come together running’), claims that it is not possible to bring them to a common denominator and find a common ground, and therefore they should

⁴⁷ The original wording is the following: “S- daet **omonimičeskie, četko protivopostavljaemye drug drugu značenijsja**. Èto: 1) ‘dviženie s raznyx storon k odnoj točke’, 2) ‘dviženie sverxu vniz’ i 3) ‘udalenie čego-nibud’ s kakogo-nibud’ mesta.” (Axmanova 1957: 150; boldfaced by me – A.E.).

belong to distinct homonymous morphemes rather than to a single morpheme with rich polysemy (Isačenko 2003/1965: 149)⁴⁸.

The same idea of homonymic (read: unrelated) uses of the prefix S- is expressed in the major dictionary of Russian word-formation by Tixonov (1985: Vol. 1, 22), where the author explicitly claims that verbs are rich in homonymous prefixes like S-, where we find S-₁ in *sletet' s dereva* 'fly/fall down from the tree' (movement downward); S-₂ in *sxodit' v magazin* 'make a round trip to the store by walking', *sbegat' k sosedjam* 'make a roundtrip to the neighbours' (what is now recognized as semelfactive use); and S-₃ in *sdelat' delo* 'get things done', *sšit' kostjum* 'sew a suit', and *svit' gnezdo* 'build a nest' (now typically recognized as the resultative reading of S-).

Yet not everybody agreed on this somewhat simplified representation of the semantics of S-. An important contribution was made by Boguslavskij (1963/2001: 32) who pointed out that the meanings 'away' and 'downward' of the prefix S- are not so separate that they should be interpreted as homonymous and triggered by different contexts: cf. *s'exat' s gory* 'slide down the hill' and *s'exat' s dorogi v storonu* 'drive off the road to one side'.

Furthermore, Ignat'eva 1970 addressed the question of how the prefix S- (including SO-) can express two so diametrically opposite spatial meanings like 'removal downward' and 'conjoining movement of objects or their parts from different locations (at least two) to one place'. Ignat'eva analyzed the combinations of the two spatial meanings of the prefix S- with the semantics of the verbal simplexes. Ignat'eva found that the meaning 'removal downward' is "chosen" by the prefix when it attaches to verbal bases that denote **destructive physical movements**, illustrated in (8). By contrast, the other spatial meaning, 'conjoining movement', is encoded by the prefix S- when it combines with simplexes that refer to **constructive physical actions**, as in (9).

- | | |
|---|--|
| (8) <u>hit:</u> | <i>bit'</i> 'hit' > <i>sbit'</i> (<i>s nog</i>) 'knock down' |
| <u>decompose an object into parts:</u> | <i>rubit'</i> 'fell (tree)' > <i>srubit'</i> 'fell (tree)' |
| <u>remove from a surface with a sharp tool:</u> | <i>skoblit'</i> 'scrape' > <i>soskoblit'</i> 'scrape' |
| <u>remove from a surface by liquid:</u> | <i>lizat'</i> 'lick' < <i>slizat'</i> 'lick off' |

- | | |
|-----------------------------------|--|
| (9) <u>hit:</u> | <i>bit'</i> 'hit' > <i>sbit'</i> (<i>slivki, jajco</i>) 'whip up (cream, egg)' |
| <u>plait, net threads, ropes:</u> | <i>vjazat'</i> 'tie, knit' > <i>svjazat'</i> 'knit' |
| <u>use sticky substances:</u> | <i>kleit'</i> 'glue' > <i>skleit'</i> 'glue together' |
| <u>apply thermal processes:</u> | <i>kovat'</i> 'forge' > <i>skovat'</i> 'forge together' (ibid. 130). |

Some physical movements can potentially be both destructive and constructive in nature. Ignat'eva argues that the prefix "puts on" either the spatial meaning of 'removal' or 'conjoining' depending on a broader context that disambiguates the reading of the verbal base. Unprefixed verbs that can combine with both conflicting spatial meanings of S- represent several semantic groups, as shown in (10).

⁴⁸ The original wording: "Na samom dele každaja pristinavka upotrebljaetsja **v neskol'kix značenijax, kotorye ne vseгда možno svesti k obščemu znamenatelju**. Jasno, čto v glagolax *s-bežat' iz tjur'my* 'utiect', *s-bežat' zbehnút' (dolu)* i *s-bežat'-sja zbehnút' sa (dohromady)* **ne 'odna pristinavka s raznymi 'ottenkami' značenija, a tri omonimičeskie pristinavki**." (Isačenko 2003/1965: 149; boldfaced by me – A.E.)

- (10) hit: *bit* 'hit' > *sbit* 'knock down' & 'whip up (cream, egg)'
motion or transportation:
bežat 'run' > *sbežat* 'run away' & *sbežat'sja* 'come together running'
transposition in space:
tjanut 'drag' > *stjanut* 'drag off and down' & 'drag together';
transfer of dry loose substances:
sypat 'pour' > *ssypat* 'pour off somewhere' & 'pour all into one place',
taking a position in space: *sadit* 'seat' > *ssadit* 'unseat' & 'seat all in one place'

By so doing, Ignat'eva (1970) discovers a fine-grained correlation and interdependency between the semantics of the verbal simplex base and the meaning of the prefix. She identified the semantic types of verbs that unambiguously require either one or the other meaning of S-. In addition, Ignat'eva showed that this distribution is complicated by a grey area of overlap: the semantic groups of verbs that allow for both spatial guises of S-. Yet, the idea that the two spatial meanings of S- are determined by the meaning of the base is maintained in the article and points to their relation and submorphemic rather than homonymous status. At the same time, the two spatial meanings are interpreted by Ignat'eva as opposite and contradicting each other and the question of what they share apart from their derivatives is not discussed.

A breakthrough in the study of the prefix S- was made by several recent articles that promote a cognitive and corpus-based approach (Dickey 2005, Dickey & Janda 2009; Janda et al. 2013: 97-100; Janda & Lyashevskaya 2013). Because the framework of Cognitive Linguistics is particularly interested in how meaning is structured in language, its methodology and usage-based perspective are powerful in explaining the internal cognitive organization of complex semantic phenomena. This framework is particularly promising in understanding the nature of polysemous prefixes like S-.

Dickey 2005 presents a semantic model that accounts for the use of the prefix Sʙ- in Old Church Slavonic (OCS), the oldest recorded Slavic language which is representative for the Common Slavic stage. Instead of a list of discrete meanings like what we typically find in the lexicographic tradition (Efremova 2010: 545; cf. Kustova 2004: 9 for discussion), Dickey 2005 proposes a structured semantic network visualized in Figure 1.

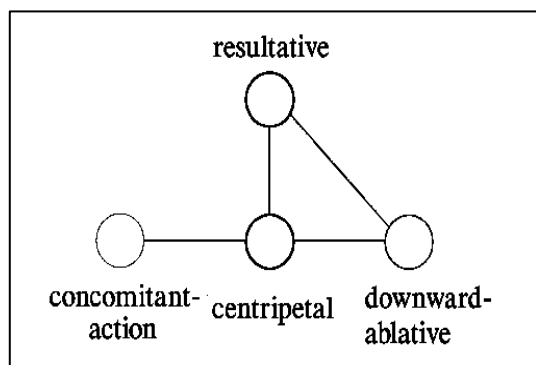


Figure 1: Model of polysemy of the prefix Sʙ- in OCS according to Dickey (2005: 11).

The nodes in Figure 1 correspond to four meanings recognized by Dickey on the basis the OCS data collected and interpreted by Słoński (1937: 225-253). The lines between the nodes depict semantic links between the submeanings. Dickey suggests the following labels and definitions for the four submeanings. The CENTRIPETAL meaning of S- implies "motion from many directions to a single landmark" (ibid. 9), as we see it in the verb *sʙbʙrati* 'bring together' < *bʙrati* 'take'. The DOWNWARD-ABLATIVE meaning refers to motion downward and away from a landmark, as in *sʙlězti* 'dismount, climb down' < *lězti* 'climb'. The RESULTATIVE meaning of S- can be observed in

the verb *sъзрѣti* ‘ripen’ < *зрѣti* ‘ripen’. The fourth meaning is CONCOMITANT ACTION (Słonsky uses the term *comitative*, *ibid.* 225) and implies being or doing something together and simultaneously, as in *съпожити* ‘live together with someone’ < *пожити* ‘live’.

The count of verbs in Słonski’s sample suggests that the CENTRIPETAL and the RESULTATIVE meanings were the most frequent and most salient for the prefix Sʙ- (47 verbs, or 31.5% each). They are marked with thicker lines of the circles in Figure 1. The DOWNWARD-ABLATIVE meaning was less frequently attested (23 verbs, or 15.4%), whereas the CONCOMITANT-ACTION meaning was present only in ten attested OCS verbs (6.7%) (Dickey 2005: 9).

As shown in Figure 1, the abstract RESULTATIVE meaning is related to both concrete spatial submeanings of CENTRIPETAL and DOWNWARD-ABLATIVE motion. Dickey convincingly argues that both of them served as the starting point for the RESULTATIVE use and contributed to its development. The verbs *съшити* ‘sew together’, *сътворити* ‘create’, *създати* ‘build’ illustrate how the RESULTATIVE meaning could derive from the CENTRIPETAL activity, whereas verbs like *съкратити* ‘shorten’ and *съкрыти* ‘hide’ suggest that it is possible to generalize the RESULTATIVE meaning of the prefix from its DOWNWARD-ABLATIVE use (for discussion cf. Dickey 2005: 10). A convincing argument in favor of the stronger connection between RESULTATIVE and CENTRIPETAL meanings comes from deadjectival and denominal inchoative verbs like *състарѣти се* ‘age’. According to Dickey, this verb represents the reflexive derivational pattern which is very productive in verbs with CENTRIPETAL Sʙ- like *сънити се* ‘come together’. Therefore, as visualized in Figure 1, the RESULTATIVE meaning is more closely related to the CENTRIPETAL than to the DOWNWARD-ABLATIVE meaning.

Crucially, Dickey includes both spatial meanings of CENTRIPETAL and DOWNWARD-ABLATIVE motion into a single network of Sʙ- and assumes “a general semantic link” between them: they both profile the transition from state 1 to state 2 (Dickey 2005: 11).

As shown by Słonski (1937: 225-253), both meanings coexisted in the semantics of two OCS verbs *съложити* ‘put together; put down’ and *съвести* ‘bring together; lead down’.

Janda and Lyashevskaya (2013: 17) express the idea that both spatial meanings of S-, CENTRIPETAL and DOWNWARD-ABLATIVE, are present in Modern Russian change of state verbs like *skondensirovat’* ‘condense’, *sgustit’* ‘thicken’, *sgnit’* ‘rot’, and *sostarit’(sja)* ‘age’. Janda and Lyashevskaya suggest that all these verbs imply becoming more compact and thus combine centripetal and downward movement.

Summing up, recent work by Dickey, Janda, and Lyashevskaya has contributed essential observations on the relationships among the various submeanings of the prefix S- and the structure of its polysemy. However, the status of SO- remains unclear and problematic in this discussion. Dickey 2005 does not discuss the mutual status of S- and SO- in Modern Russian, and neither do Janda and Lyashevskaya (2013). The question remains whether the non-trivial differences of “the two” SO- should be attributed to sub-morphemic differences and can be accounted for within a single model of polysemous S- or the CONCOMITANT ACTION use of SO- belongs to a separate morpheme.

This particular issue has not been examined yet on the basis of large corpus data and claims have been restricted so far to the observation of limited sets of examples (Itkin 2007; Valedinskaja & Golanova 2007). In this chapter, I adopt an agnostic view and make use of an extensive dataset culled from the Russian National Corpus. I present the data and methodology in the next section.

4.4 Data collection

In order to test whether the prefixes S- and SO- are allomorphs of a single morpheme or two distinct morphemes in Contemporary Standard Russian I have compiled a database. The database contains 998 prefixed verbal lexemes. The data set is culled from the Russian National Corpus, manually tagged for a number of parameters, and is available at <http://hdl.handle.net/10037.1/10078>. In the following, I describe how this data was collected and tagged.

4.4.1 Methodology

The choice of the Russian National Corpus as the source of data is motivated by the objective to collect the most complete and extensive dataset possible that would account not only for the lexemes present in the standard Russian lexicon but would also be representative in terms of innovative occasional formations that reflect the productivity and entrenchment of certain derivational patterns. Dictionaries and reference works are usually limited to standardized lexemes well established in the lexicon, and therefore lack occasionalisms and substandard marginal colloquial words. In this regard, the Russian National Corpus is a more promising and superior data source. It is a well-balanced collection of texts that provides data on both types of words (standard and marginal) and gives information on token frequencies of individual lemmas. Frequencies of words in the corpus can therefore help in distinguishing between standard vs. novel lexemes as well as in estimating a relative entrenchment of a word in a given usage.

Since in this study I aim to account for the mutual status of the prefixes S- and SO- in Modern Russian, I addressed the relevant part of the Russian National Corpus, namely its Modern Subcorpus. The latter contains over 123 million word attestations⁴⁹ and covers the texts created in 1950-2013.

Data collection took several steps. Because the corpus itself does not give access to a list of lemmas, as a first step I obtained a list of all verbs that start with the letter *s*.⁵⁰ The verbs were extracted from the frequency dictionary by Lyashevskaya & Šaroff (2009), where the inventory of lemmas is based on the Modern Subcorpus (at that time

⁴⁹ To be more exact, 123,411,770 words when the corpus was accessed for this study in August 2013.

⁵⁰ An alternative solution would be to extract all *s(o)*-initial verbs which have a simplex verbal base also attested in the corpus. However, this alternative leads to some unwanted restrictions on the data, for example it misses verbs like *s'ezdit'* 'travel to someplace and back', *symitirovat'* 'imitate once', and *sojti* 'walk downward, take off' because due to spelling and historical rules their unprefixated bases cannot be arrived at by simple omission of the prefix: namely *ezdit'* 'travel, drive', *imitirovat'* 'imitate', and *idti* 'go, walk'. In addition, a number of novel factitive verbs with semelfactive meaning would be left out, for instance *sblagorodničat'* 'behave in a noble way once' (from the adjective *blagorodnyj* 'noble'), *snecenzurničat'* 'say something obscene once' (from the adjective *necenzurnyj* 'obscene, indecent'), *sgitarit'* 'play guitar once' (from the noun *gitara* 'guitar'), since in the corpus there are no attested unprefixated verbs like **blagorodničat'*, **necenzurničat'*, and **gitarit'*. Due to these downsides, the alternative method of data extraction was rejected. Instead, all verbs that start with *s* (regardless of the attestation of the corresponding simplex base) were culled from the corpus. This labor-intensive, yet much more accurate, method required manual double-checking of every single verb in the corpus and finally yielded the most complete dataset possible.

covering texts from 1950-2007) of the RNC. The data was extracted automatically via the software management program MySQL and included 3,703 datapoints. After that each verb was double-checked manually in the corpus. As a consequence, a number of verbs were excluded, first of all numerous misspellings (e.g. *sogdavat* instead of *sozdavat* 'create') as well as unprefixed verbs like *solit* 'salt' and *signalit* 'signal'. Furthermore, I set aside rare verbs with the historical allomorph SU- like *sumnevat'sja* 'doubt'.

Another type of excluded verbs are those few that are derived from a whole prepositional phrase like *sumasbrodstvovat* 'go crazy', *sumasvodit* 'make someone go crazy' (coined by Nabokov), and *snogsšibat* 'surprize, make fall from astonishment'. Such verbs are derived from prepositional phrases and are not only structurally different from the remainder of the database, but also they unnecessarily duplicate the data, since the corresponding prefixed verbs *svesti* (*svodit*) 'lead downstairs, bring to zero, diminish', *sbresti* (*sbrodit*) 'go away from', and *sšibat* (*sšibat*) 'hit off and down' are already included.

Finally, because the objective of the database was to collect verbal derivatives formed by the attachment of the prefixes S- and SO-, a number of clearly denominal suffixed verbs had to be excluded, such as *sovestit* 'shame, appeal to one's conscience' (derived from the noun *sovest* 'sense of conscience'), *soborovat* 'concelebrate' (derived from the noun *sobor* 'cathedral'), *soveršenstvovat* 'bring to perfection' (from the noun *soveršenstvo* 'perfection'), and *sosedit* 'be a neighbour' (derived from the noun *sosed* 'neighbour').

However, many verbs are more problematic than such unidirectional derivations and often represent derivational constructions with multiple motivational links. There is no clear-cut boundary that separates denominal and deverbal derivations, especially when we consider verbs like *sožitel'stvovat*, which is associated with both a noun (in this case *sožitel* 'cohabitor' and *žitel* 'inhabitor') and a verb (*žitel'stvovat* 'live, reside, dwell'⁵¹). A parallel example comes from the verb *sotvorčestvovat* 'create something together' based on the nouns *sotvorčestvo* 'co-creation', *tvorčestvo* 'creation', and the verb *tvorčestvovat* 'create', which can marginally be derived from the corresponding noun.⁵² In a sense, such verbs are very similar to deverbal derivatives like *součastvovat* 'participate together with someone', which is formally derived from the simplex verb *učastvovat* 'participate', but nevertheless is semantically closely associated with the nouns *součastnik* 'copartner, accomplice' and *učastnik* 'participant'. For this reason the database includes *sožitel'stvovat* and *sotvorčestvovat* along with other verbs of this kind.

In addition to the exclusion of irrelevant datapoints discussed above, a number of further measures were undertaken. The major challenge was to solve the problem of

⁵¹ Although the verb *žitel'stvovat* 'reside, dwell' might seem marginal and defective, it is attested as an obsolete word in major dictionaries (e.g. Ožegov & Švedova 2001, Efremova 2006) and yields twenty-one attestations in the Modern Subcorpus of the RNC, with the search run on 15.09.2013.

⁵² There are two attestations of the verb *tvorčestvovat* 'create' in the Modern Subcorpus of the RNC (search run 15.09.2013): *My mirnye, esli nas ne trogat', veselye, esli ne obremenjat', sposobnye učit'sja i tvorčestvovat', esli nas ne kantovat*. [T. Savel'eva. Dviženie na omoloženie. (1997)] 'We are friendly, if you do not bother us, we are cheerful, if you do not burden us, we are able to learn and **create**, if you leave us alone.'

unnecessary duplication of data. The original list of culled verbs included perfectives together with their corresponding secondary imperfectives, often attested in both non-reflexive and reflexive forms. For example, the corpus contains aspectually related verbs like *sbrosit'* and *sbrasyvat'* 'throw down' together with their reflexives *sbrosit'sja* and *sbrasyvat'sja* 'throw oneself down'. Since the objective was to make a database that could serve as a useful tool for estimating how many lexemes represent different submeanings of the prefixes S- and SO- in the Modern Russian lexicon, it was important to avoid such duplication of data, especially given the fact that some of these forms are theoretically possible even if they are not attested in the corpus⁵³. In order to avoid data duplication, each secondary imperfective was merged with its perfective counterpart in a single lemma. For example, the verbs *sbrosit'* and *sbrasyvat'* were merged into one verbal lexeme *sbrosit'* 'throw down', which occupies a single entry in the database. The assumption behind this policy was that it is the perfective verb that is formed by the attachment of the prefix, while imperfectives are as a rule formed via attachment of suffixes. The imperfective verb occupies a separate entry in the database only if it lacks an attested perfective counterpart, for example *sosuščestvovat'* 'coexist'.

In order to avoid a surplus of identical data, a similar measure was applied to reflexives. If a reflexive verb was different from its non-reflexive partner only in terms of transitivity, they were merged as representing a single verbal lemma. In such a case, the postfix *-sja-* was placed in parentheses: for instance, *sbalansirovat'* 'balance' and *sbalansirovat'sja* 'balance oneself' were collapsed into a single entry *sbalansirovat'(sja)* 'balance (oneself)', parallel to other verbs like *sdelat'(sja)* 'make', *syskat'(sja)* 'find', *skopit'(sja)* 'save up', *sverit'(sja)* 'compare with, check up with', *sožmurit'(sja)* 'close one's eyes tight once', etc. In total, there are 300 such joint nonreflexive+reflexive entries in the database. A reflexive verb is presented in a separate entry only if 1) its meaning is different from that of the non-reflexive (e.g. *sbežat'* 'escape, run away; run downhill' vs. *sbežat'sja* 'come running from different locations to one place'); or 2) the non-reflexive counterpart is not attested in the corpus (e.g. *stabunit'sja* 'herd horses in a tabun' in the absence of **stabunit'*; *srinut'sja* 'get off in a rush' in the absence of **srinut'*). In case the reflexive counterpart is polysemous and refers to two meanings, where one is identical to its non-reflexive partner, while the other is different, both variants are taken into account. As an example, consider the reflexive verb *sbrosit'sja*: in the meaning 'throw oneself down' it was merged with the non-reflexive perfective verb *sbrosit'* 'throw down' into a single lexeme *sbrosit'(sja)*, while the other attested meaning of the reflexive 'chip in together, collect money' was added to the database as a separate entry.

After all the manipulations described above, the original unstructured list of 3,703 datapoints yielded an extensive elaborated database of 998 verbal lexemes prefixed with S- and SO-, representative for their distribution in Contemporary Standard Russian. The inventory of 998 verbal lexemes accounts for 1,113,239 actual attestations of these verbs in the RNC. Each lemma in the database is supplied with an English gloss, a glossed simplex base, token frequency (number of attestations) in the Modern Subcorpus of the RNC, and a few illustrative corpus examples of its use. The database is

⁵³ For example, even if some secondary imperfectives are missing from the database and the corpus, they are theoretically possible and can occur in language use, as evidenced by search engines like *www.yandex.ru* and *www.google.ru* and as convincingly shown in the study on Russian aspectual triplets by Kuznetsova & Sokolova forthcoming; cf. also Janda et. al. 2013: 163-179).

formatted as a Microsoft Excel file, which supports a variety of search and filtering options. Each verb was tagged for its aspect and prefix (exclusive S- vs. exclusive SO- vs. alternating S- ~ SO- within verbal paradigm). In addition, each verb was assigned one or more semantic tags according to the network of submeanings of these prefixes S- and SO- that I propose in 4.5.

4.4.2 S- and SO- in standard and marginal verbs

Each verbal lexeme in the database was manually checked for its token frequency in the Modern Subcorpus of the RNC. This information on token frequencies makes it possible to distinguish between novel occasionalisms with a single occurrence or few attestations in the corpus on the one hand and verbal lexemes highly frequent in the corpus and well established in the lexicon on the other hand. Both types of verbs are important for my analysis. Marginal occasionalisms provide evidence of the productivity of a derivational pattern and a given submeaning of the prefix, whereas standard words belong to the part of grammar that speakers share and build their linguistic competence on. In this subsection, I outline the overall picture and focus on the differences in the distribution of S- and SO- across verbs of these two types.

Token frequencies of individual verbs form a continuum, but it is nevertheless possible to group verbs together according to the overall number of their occurrences in the corpus. Although this subdivision of verbs into frequency groups might seem arbitrary, I suggest that for verbs prefixed in S- and SO- 10 attestations is a threshold that separates occasional marginal formations from standard verbs that belong to the shared Russian lexicon. Those verbs that cross the threshold of 10 corpus occurrences are counted as established standard lexemes. They represent the system, while marginal occasionalisms point to the direction where this system is moving. Low frequency verbs comprise almost half of the entire dataset, namely 44% (442 verbs with 1 to 9 attestations in the corpus). If we consider the two prefixes together, out of 1,000 verbs prefixed in S- and SO-, 196 verbs (19%) are attested in the corpus only once (hapax legomena), while 246 verbs (25%) have only 2-9 attestations.

Now a question arises as to whether the distributions of S- and SO- across standard and marginal verbs are significantly different or not. The parallel pie diagrams in Figures 1 and 2 show that the distribution of verbs is surprisingly similar. A chi-square test indicates that the distributions of the verbs prefixed in S- and SO- are not significantly different across frequency groups⁵⁴.

⁵⁴ The string of values as in `a=matrix(c(166, 29, 212, 34, 122, 16, 51, 10, 267, 70), ncol=2, byrow=TRUE)`, where I collapse the values for verbs with over 100 attestations, yields $X^2 = 8.5$, $df = 4$, $p\text{-value} = 0.07$.

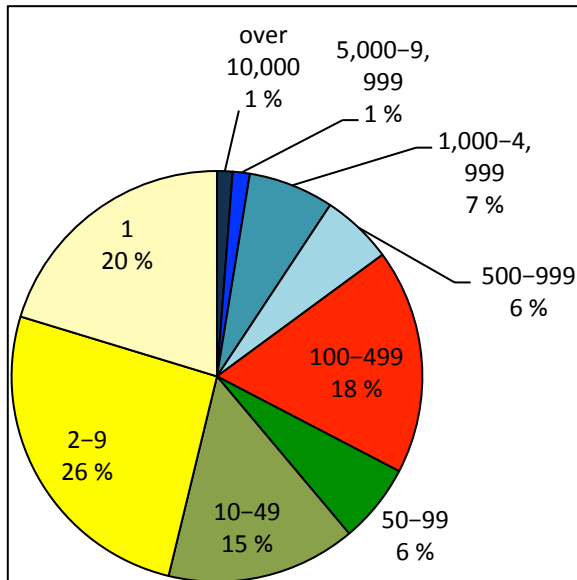


Figure 1: Verbs in S-

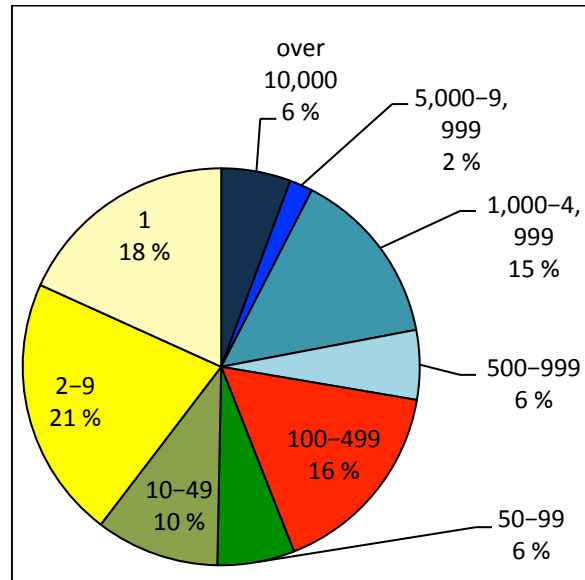


Figure 2: Verbs in SO-

Even if we collapse all subgroups into two large categories of marginal verbs (1-9 attestations) vs. standard verbs (over 10 attestations), as shown in Table 1 and Figure 3, we get very similar proportions of derivatives in S- and SO- across the two frequency types.

# attestations	S- (%)	SO- (%)	S-~SO- (%)	Total
over 10	440 (79%)	96 (17%)	20 (4%)	556 (100%)
1-9	378 (85.4%)	63 (14.4%)	1 (0.2%)	442 (100%)

Table 1: Distribution of S- and SO- across standard and marginal derivatives.

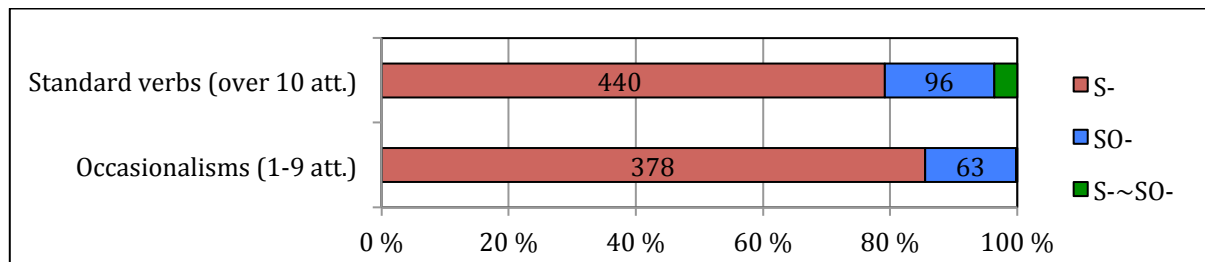


Figure 3: Distribution of S- and SO- across standard and marginal derivatives.

Indeed, in standard verbs the highly frequent prefix S- (79%) predominates over the restricted SO- (17%), and this overall proportion of S- vs. SO- prevails in marginal verbs (compare the parallel bars on Figure 3) with a slight difference in that the prefix S- is even more productive (85.4%), while SO- is even more restricted (14.4%). Yet this difference is not statistically significant.⁵⁵

⁵⁵ Merging verbs with 1-9 attestations vs. verbs with over 10 attestations, as shown in Table 1 (a2=matrix(c(440,96,378,63), ncol=2, byrow=TRUE)), yields values that point to statistical non-significance: X-squared = 2, df = 1, p-value = 0.15.

4.5 Semantic analysis

In my analysis of verbs prefixed with S- and SO-, I argue against three accounts discussed above:

- All SO- are regular allomorphs of the morpheme S- (Švedova et al. 1980: §851);
- Some SO- represent a distinct morpheme SO₂- (Itkin 2007: 230);
- There are three homonymous morphemes S₁-, S₂-, S₃- (Axmanova 1957: 150).

I propose that all submeanings of S- and SO- are interrelated and that the meaning CONCOMITANT-ACTION is well incorporated into the shared semantic network. However, I argue that the relations of S- and SO- are more complex than that of Standard Allomorphy because of strong specialization of the two prefixes for certain submeanings.

The central question of this section relates the data to the semantic criterion of allomorphy: do the prefixes S- and SO- have identical meaning or are they semantically different? Because both prefixes have more than one meaning, the central challenge is to compare the two prefixes with regard to all their submeanings and find out whether these meanings are related or not.

In what follows, I present a Cognitive Linguistics analysis of polysemy. I propose that, on the one hand, S- and SO- are semantically very similar and share a single network of related submeanings. Moreover, I show that each submeaning in this model can be expressed by both prefixes. On the other hand, within the shared semantic network S- and SO- have different “centers of gravity”, what I refer to as *radial category profiles* (recall the discussion of profiling in Chapters 1 and 3). In other words, the two prefixes are not fully identical in terms of semantics, yet very similar.

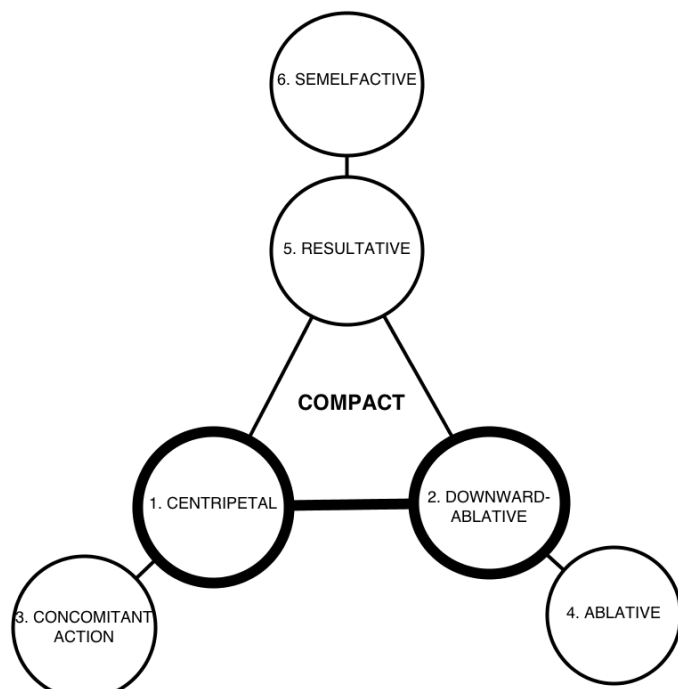


Figure 4: Radial category model of polysemy for the prefixes S- and SO- in Modern Russian.

I propose that all uses of S- and SO- can be described by a single model of polysemy shown in Figure 4. This model is a radial network which consists of six submeanings. The nodes represent submeanings of the prefixes identified on the basis of semantic comparison of prefixed derivatives with their simplex bases. The lines between nodes represent semantic links between submeanings. I discuss each of these links below.

In order to be consistent with previous accounts, I adopt the terminology and semantic distinctions established in Dickey 2005. Therefore, it is possible to compare my findings on S- and SO- in Modern Russian with the profile of S₁B- in OCS. In addition, I introduce

some formal criteria that help to analyze individual verbal lexemes and to distinguish between adjacent submeanings.

The network in Figure 4 has a radial structure, where all meanings are hierarchically organized around the two central spatial prototypes represented by thick circles. The prototypes are the two submeanings expressing CENTRIPETAL movement and DOWNWARD-ABLATIVE movement. The two prototypes are related to each other and motivate other meanings in the network. The remaining submeanings are related to these prototypes via the cognitive mechanisms of metaphor and metonymy, as I argue in this section.

Degrees of granularity in distinguishing prefix submeanings can vary. The model presented in Figure 4 serves best my major objective, which is to compare in detail the two prefixes in terms of their semantics.

4.5.1 CENTRIPETAL MOTION VS. DOWNWARD-ABLATIVE MOTION

At first glance, the two spatial prototypes are very different. The CENTRIPETAL meaning ('putting things together') encodes a movement of at least two trajectors from different locations to one place, as shown in Figure 5. The DOWNWARD-ABLATIVE meaning ('down & off') refers to the movement of the trajector downward and away from the landmark, as shown in Figure 6.

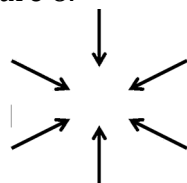


Figure 5: Trajectory of CENTRIPETAL movement.



Figure 6: Trajectory of DOWNWARD-ABLATIVE movement.

Both meanings can be expressed by S- and SO-, as evidenced by examples in (11), and also by the parallel preposition *s(o)* (12).

- | | |
|---|---|
| <p>(11) CENTRIPETAL
 <i>s-kleit</i> 'glue together' < <i>kleit</i> 'glue'
 <i>so-brat</i> 'bring together' < <i>brat</i> 'take'</p> | <p>DOWNWARD-ABLATIVE
 <i>s-prygnut</i> 'jump down' < <i>prygnut</i> 'jump'
 <i>so-stupit</i> 'step down' < <i>stupit</i> 'step'</p> |
|---|---|

- | | |
|---|--|
| <p>(12) <i>skleit</i> 'drug <i>s</i> drugom
 'glue together to each other'</p> | <p><i>sprygnut</i> 's dereva
 'jump down from the tree'</p> |
|---|--|

The two prototypical submeanings can be expressed by S- and SO- when they are attached to the same base verbs. A good example is the verb *sognat*'_{PF} – *sgonjat*'_{IPF} (< *gnat* 'chase') which denotes 'chase all into one place' in one context (13) and 'chase someone away and down' in another context (14).

- (13) *Vdali pastux **sognal** v kuču ovec.* [A. Iličevskij. Pers (2009)]
 'In the distance a shepherd **herded** the sheep **all together**.'

- (14) *Sam čert teper' menja ne **sgonit** s verxnej polki.* [I. Grekova. V vagone (1983)]
 'Now even the devil will not **chase** me **down from** the upper berth.'

The verb *sognat* ‘chase’ is not the only verb where the prefix has two readings. There are at least twenty-two other verbs that behave the same way. Table 2 provides a representative sample of such verbs, and an exhaustive list is available in the database.

#	CENTRIPETAL reading	Verb	DOWNWARD-ABLATIVE reading
1	‘chase all into one place’	<i>sognat</i> ’	‘chase someone away and down’
2	‘throw into a pile’	<i>svalit</i> ’	‘throw down’
3	‘bring together’	<i>svesti</i>	‘lead downward and away from’
4	‘bring together’	<i>svezti</i>	‘take downward and away from’
5	‘roll into one thing’	<i>svernut</i> ’	‘deviate off the main path’
6	‘drag to a place from many places’	<i>svolóč</i> ’	‘drag downward’
7	‘move to the center’	<i>sdvinut</i> ’	‘move away’
8	‘put together’	<i>složít</i> ’	‘lay away’
9	‘collide’	<i>stolknut</i> ’	‘push off, edge’
10	‘flow in’	<i>steč</i> ’	‘flow down’

Table 2: Verbs with CENTRIPETAL and DOWNWARD-ABLATIVE meanings of the prefix.

The data in Table 2 has been discussed to some degree in previous scholarship: the two meanings of the prefix can be expressed by the same prefixed verbs put in different contexts, where the context resolves the ambiguity of the derivative (Axmanova 1957, Ignat’eva 1970). However, one fact remains largely overlooked: the submeanings listed for each verb in Table 2 can sometimes be present in the same context simultaneously. For example, in (15) the prefix SO- in the participle of the verb *sognat*’ has both CENTRIPETAL and ABLATIVE meanings at the same time:

- (15) *Tolpa, sognannaja so vsech okrestnyx ulic, počti bezmolvno kolebalas’ vnizu.* [Ju.O. Dombrovskij. Obez’jana prixodit za svoim čerepom (1943-1958)]
 ‘The crowd, **gathered/chased away from** neighboring streets was below, almost silently moving downward’.

Likewise, in (16) both meanings of the prefix, CENTRIPETAL and DOWNWARD-ABLATIVE, are relevant for the content of the sentence.

- (16) *Tut že, u samogo zabora, byli svaleny v kuču doski.* [O. Pavlov. Karagadinskie desjatiny, ili Povest’ poslednix dnei. “Oktjabr”, 2001]
 ‘Right there by the fence, there were planks **thrown down into a pile**’.

How can we explain the coexistence of these two submeanings in the same contexts like (15) and (16), which involve the same verbs (Table 2) and the same prefixes S- and SO-? The first possible explanation that comes to mind is that the meanings CENTRIPETAL and DOWNWARD-ABLATIVE might come from different etymological sources. However, etymological dictionaries (Vasmer 1971 v.3: 539; Šapošnikov 2010 v.2: 296) do not talk about two distinct sources, but instead suggest a single source for both readings of S-: the Proto-Indo-European root **kom* ‘together, with, from’ which can be found among other successors in Latin *cum* ‘with’ and Greek *κατά* ‘down’.

It is fruitful to approach this case from the perspective of cognitive semantics. The two meanings share a single image schema that we can call COMPACT. This has been noted before, in Janda et al. 2013 and Janda & Lyashevskaya 2013: “The TOGETHER and DOWN meanings are related to each other because when something comes TOGETHER the size also goes DOWN as it becomes more compact, as we see in *sgustit’* ‘thicken.’” (Janda et al. 2013: 97).⁵⁶

Another good example is the verb *sžat’* ‘clench, squeeze’ which refers to making an object COMPACT. As shown in Figure 7, both CENTRIPETAL and DOWNWARD-ABLATIVE movement make an object more COMPACT. In this light, the trajectories of these two spatial movements are facets of a single image schema COMPACT. In other words, conceptually, the image schema COMPACT is present in both CENTRIPETAL and DOWNWARD-ABLATIVE movement which are foregrounded in different contexts. For example, the expression *sžat’ ruku v kulak* ‘clench a hand into a fist’ foregrounds the CENTRIPETAL reading of the verb and refers to a clenched fist that is smaller and more compact than an open palm. Similarly, the same verb can describe DOWNWARD-ABLATIVE movement of a spring (*sžat’ pružinu* ‘compress spring’) or a skyjack scissor lift (Figure 7) which are more compact when they are pressed all the way down.

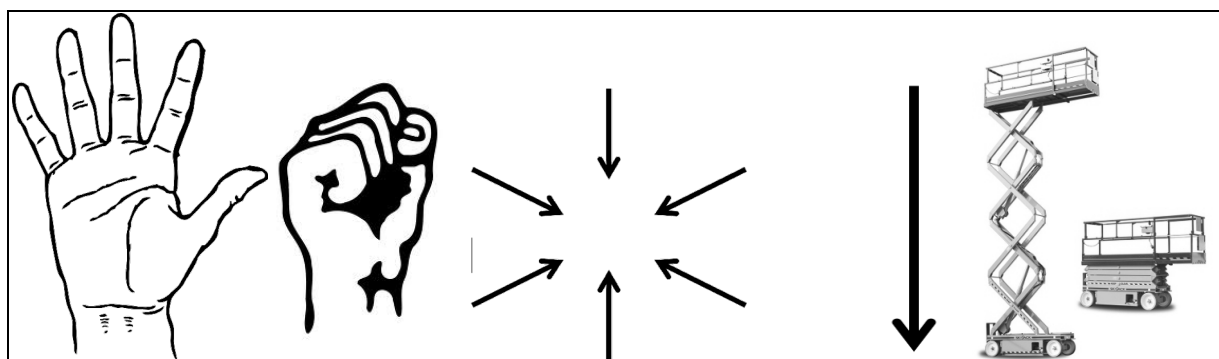


Figure 7: Two facets of the image schema COMPACT encoded in the verb *sžat’* ‘squeeze’.

There is a whole group of verbs like *sžat’* ‘squeeze’, where the semantic contribution of the prefix overlaps with the lexical meaning of the verbal base: *skompressirovat’* ‘compress’ (< *kompressirovat’* ‘compress’), *skondensirovat’* ‘condense’ (< *kondensirovat’* ‘condense’), *skonsolidirovat’* ‘consolidate’ (< *konsolidirovat’* ‘consolidate’), where *kon* etymologically comes from the same source as the prefix S-. Other verbs that encode the image schema COMPACT include lexemes that refer to size reduction (*spljuščit’* ‘flatten’, *spressovat’* ‘press down’, *suzit’* ‘narrow down’, *s’ežit’(sja)* ‘shrink’, *sdut’(sja)* ‘blow off, deflate’) and temporal reduction (*skorotat’* ‘shorten, while away’).

I suggest that we can describe the two submeanings CENTRIPETAL MOTION and DOWNWARD-ABLATIVE MOTION in terms of viewing effects, or construal (Langacker 1999: 206). The two meanings are two interpretations, or *construals*, of a single image schema COMPACT.

This effect can be compared to Rubin’s famous face/vase-figure in Figure 8, where you can see either the foregrounded faces or foregrounded vase depending on the

⁵⁶ Making something more COMPACT implies reduction in size. The relationship between reduction and downward movement is well established through the orientational metaphor LESS IS DOWN (Lakoff & Johnson 1980: 15).

color that you focus on – white or black. Whatever one sees, both construals are available.

Similarly, different verbal bases support different readings of the prefix. For example, sewing implies putting things together, and therefore it is natural for the base *šit* ‘sew’ to combine with S- in its meaning CENTRIPETAL MOTION: *sšit* ‘sew pieces together’.⁵⁷

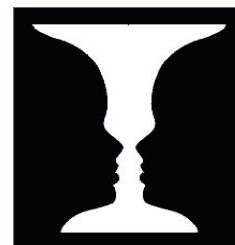


Figure 8: Rubin's face/vase figure.

By contrast, in *sbrit* ‘shave off’ the base *brit* refers to removing hair from a surface and therefore highlights the DOWNWARD-ABLATIVE MOTION meaning of the prefix.

If both readings of the prefix correlate with the verbal base, as in *vesti* ‘lead’, it is the context that highlights either CENTRIPETAL or DOWNWARD-ABLATIVE reading. Some contexts allow both readings, as I discussed above (recall (15) and (16)).

4.5.2 CENTRIPETAL MOTION VS. CONCOMITANT ACTION

The submeaning 3.CONCOMITANT ACTION is closely related to the prototypical meaning 1.CENTRIPETAL MOTION. Their image schemas are juxtaposed in Figure 9. Recall that CENTRIPETAL MOTION implies a change of location: two or more trajectors move from different locations to one place (*sblizit* ‘bring closer to each other’, *sostavit* ‘compose, put together’, *soeditit* ‘unite’), as visualized by two arrows in Figure 9. By contrast, the submeaning CONCOMITANT ACTION implies an image schema where both trajectors are in the same location from the very beginning of the situation and they proceed **together and identically** through a state or an activity, as depicted by two parallel lines in Figure 9. The CONCOMITANT ACTION meaning is expressed by the prefix in many imperfective verbs. Some representative examples are given in (17):

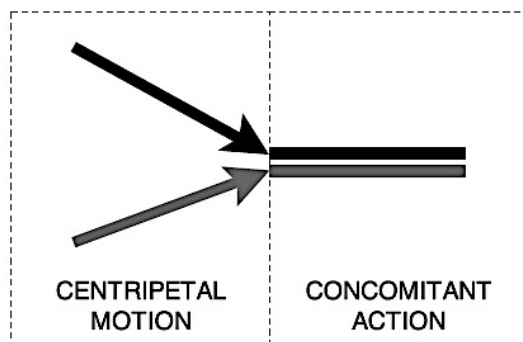


Figure 9: Relationship of image schemas of 1.CENTRIPETAL MOTION and 3.CONCOMITANT ACTION.

- | | | |
|------|---|-----------------------------------|
| (17) | <i>sosuščestvovat</i> ‘coexist’ | < <i>suščestvovat</i> ‘exist’ |
| | <i>sočuvstvovat</i> ‘sympathize with someone’ | < <i>čuvstvovat</i> ‘feel’ |
| | <i>sodejstvovat</i> ‘support someone’ | < <i>dejstvovat</i> ‘act’ |
| | <i>součastvovat</i> ‘participate together with someone’ | < <i>učastvovat</i> ‘participate’ |

The relationship between these two submeanings can be best explained in terms of metonymy: the trajectory of CENTRIPETAL MOTION is reduced to its endpoint, where the trajectors have arrived at the same location and proceed in parallel from there. Thus, the movement along the path is reduced to the endpoint of the path, which is often referred to as an “end-point metonymy” (Janda 2010: 18). Similarly, we observe this semantic phenomenon in the use of prepositions. Consider the English contexts for the preposition *over*: *John walks over the hill* vs. *John lives over the hill*. In the second

⁵⁷ Note that this verb has a more general sense ‘sew (e.g. a costume)’ where the meaning of the prefix is RESULTATIVE.

sentence, there is no movement along a trajectory but there is a state that takes place at the endpoint of the trajectory. In this light, I suggest that the meaning CONCOMITANT ACTION can be seen as a semantic reduction, or an end-point metonymy of CENTRIPETAL MOTION.

Notice that in the example with *over*, the sentence *John walks over the hill* refers to a dynamic change, whereas the end-point metonymy in the sentence *John lives over the hill* is encoded by the stative predicate *live*. Similarly, CENTRIPETAL MOTION refers a dynamic change of state, but its end-point metonymy, CONCOMITANT ACTION, implies a balanced state. This conceptual difference between the two submeanings in terms of force dynamics (Talmy 1985, 2000) corresponds to the opposite aspectual characteristics of the verbal predicates. When the prefix contributes the meaning CENTRIPETAL MOTION, it typically turns an imperfective base into a perfective derivative (e.g. *stavit'*_{IPF} 'put' > *sostavit'*_{PF} 'compile'). By contrast, when the prefix contributes the meaning CONCOMITANT ACTION, an imperfective base normally remains imperfective (e.g. *čuvstvovat'*_{IPF} 'feel' > *sočuvstvovat'*_{IPF} 'sympathize').

The submeaning CONCOMITANT ACTION is particularly prominent for the prefix SO- and has been claimed to be nearly exclusive for this prefix (Itkin 2007). Moreover, it is productively used in novel coinages like *sofinansirovat'* 'co-finance, help to finance' (< *finansirovat'* 'finance'), *sorazmysljat'* 'think together with someone' (< *razmysljat'* 'think'), *socirkulirovat'* 'coexist, co-spread' (< *cirkulirovat'* 'circulate'). SO- is frequently used in this meaning beyond verbs, especially in nouns like *sotvorčestvo* 'co-creation', *soavtor* 'co-author', and *sosed* 'neighbor'. The applicability of the CONCOMITANT ACTION meaning to nouns lends additional support to the idea about the stative, non-eventive nature of this meaning.

The prefix S- expresses the meaning CONCOMITANT ACTION very rarely, but is attested, for example, in the noun *sputnik* 'fellow traveler, satellite' (< *putnik* 'traveler', *put'* 'way'), which corresponds to the verb *soputstvovat'* 'accompany' (< *put'* 'way'). I suggest that S- carries this meaning in two standard verbs. The first one is a somewhat archaic verb *spospešestvovat'* 'assist, help, promote' synonymous to the verbs in SO- *sodejstvovat'* 'support, lit. act together' and *soputstvovat'* 'accompany'. A typical context of this verb is illustrated in (18):

- (18) *Izbegajte tragedij, i vam budet spospešestvovat' udača.* [A. Najman, G. Narinskaja. *Process edy i besedy. 100 kulinaryx sovetov. // «Oktjabr'», 2003*]
'Avoid tragedies and success will **keep** you **company** (lit. go together with you).'

The second verb that demonstrates that S- can express the CONCOMITANT ACTION meaning is a synonymous verb *sposobstvovat'* 'facilitate' which is not clearly analyzable, but is arguably related to *posobit'* 'help, assist'.

Apart from these verbs, there is a group of lexemes in (19) that are intermediate between the meaning CENTRIPETAL MOTION and the meaning CONCOMITANT ACTION.

- (19) *srabotat'sja* 'become closer while working together' < *rabotat'* 'work'
sžit'sja 'become used and connected to something' < *žit'* 'live'
sgovorit'sja 'agree on, make a joint plan with someone' < *govorit'* 'talk'

storgovat'sja 'agree on price, become partners, make a deal' < *torgovat'sja* 'bargain'
snjuxat'sja 'nuzzle together with someone, become friends' < *njuxatsja* 'nuzzle'
sygrat'sja 'become closer by playing together' < *igrat* 'play'
stolkovat'sja 'make an arrangement together with someone as a result of negotiations' < *tolkovat* 'talk'

All verbs in (19) represent the same pattern: they are formed by a prefix-postfix construction S-...-SJA, which structurally resembles motion verbs like *sbežat'sja* 'come together running'. However, in *sbežat'sja*, the participants do not affect one another while running. The verbs listed in (19) are different in this regard, because they describe mutual activities that imply close interaction between the participants. The general meaning of verbs in (19) is 'become closer to one another by doing something together', where the shared activity involves both participants on equal terms. In this sense, the prefix S- in such derivatives encodes two submeanings: CENTRIPETAL MOTION and CONCOMITANT ACTION. The latter meaning is also supported by the reciprocal postfix -SJA. A good example comes from a newly formed marginal verb *srastvorit'sja* 'co-dissolve' (< *rastvorit'sja* 'dissolve') illustrated in (20):

(20) *Nedostatočno slovo Božie uslyšať; ono dolžno proniknut' v naši glubiny, kak semja, padajuščee v dobruju zemlju, srastvorit'sja s ètimi glubinami, i prinesti plod – ne razmyšlenij, no žizni novoj, žizni večnoj, žizni Samogo Xrista i Duxa, živuščego v nas i dejstvujuščego v nas.* [mitropolit Antonij (Blum). O soxranenii slova Božija (1983)].
 'It is not enough to listen to the God's word; it [God's word] should find the way into the depths of our soul and, like a seed that falls on good soil, it [God's word] should **co-dissolve** (lit.) together with these depths, and bring forth fruit – not thoughts but a new life, eternal life, the life of Jesus Christ and the Holy Spirit, living in us and affecting us.'

The examples in (19) and (20) facilitate a close conceptual connection between the submeanings CENTRIPETAL MOTION and CONCOMITANT ACTION.

4.5.3 DOWNWARD-ABLATIVE MOTION VS. ABLATIVE MOTION

Similar to the relationship between CENTRIPETAL MOTION and CONCOMITANT ACTION, I suggest that the submeaning ABLATIVE MOTION represents a semantic reduction of DOWNWARD-ABLATIVE MOTION.

The prototypical submeaning implies the movement downward and away from the landmark, as shown in Figure 10. Recall that this meaning is manifested by the prefix in verbs like *s-prygnut* 'jump down' (< *prygnut* 'jump'), *sokol'znut* 'slide down' (< *skol'znut* 'slide').



Figure 10: Image schema of submeaning 2. DOWNWARD-ABLATIVE MOTION ('DOWN' + 'OFF').



Figure 11: Image schema of submeaning 4. ABLATIVE MOTION ('OFF').

By contrast, the schema of ABLATIVE movement is underspecified for the vertical dimension of the trajectory and merely directed away from the landmark, as depicted in Figure 11. In this sense, ABLATIVE movement is related to the prototypical meaning DOWNWARD-ABLATIVE MOTION via metonymy. However, we are dealing with *part-whole* metonymy, not *end-point metonymy*.

The ABLATIVE meaning is attested for both prefixes S- and SO- and is manifested in many verbs that refer to concrete spatial motion, as in (21):

(21) <i>sčistit'</i> 'scrape off'	< <i>čistit'</i> 'clean'	<i>soskoblit'</i> 'plane away'	< <i>skoblit'</i> 'plane'
<i>smestit'</i> 'displace'	< <i>mesto</i> 'place'	<i>soskresti</i> 'scrape off'	< <i>skresti</i> 'scrape'
<i>sliznut'</i> 'lick off'	< <i>liznut'</i> 'lick once'	<i>sorvat'</i> 'tear off'	< <i>rvat'</i> 'tear'
<i>steret'</i> 'erase'	< <i>teret'</i> 'rub'	<i>sostrič'</i> 'cut off'	< <i>strič'</i> 'cut'

Other applications of the ABLATIVE S- and SO- include verbs that refer to thievery and exile, the concepts that imply spatial removal: *staščit'* 'steal' (< *taščit'* 'drag'), *soslat'* 'send away' (< *slat'* 'send'). Likewise, S- and SO- in the meaning ABLATIVE form predicates that describe getting rid of something and hiding something away: *sbyt'* 'get rid of; work off' (< *byt'* 'be') vs. *sokryt'* 'hide away' (< *kryt'* 'cover').

One can often observe that the same verb can have both DOWNWARD-ABLATIVE and just ABLATIVE reading. For example, the verb *sbežat'* 'run' can refer to running downstairs (DOWNWARD-ABLATIVE) or away from home (ABLATIVE), where the vertical dimension is irrelevant. Such verbs speak in support of the close relationship between the two meanings.

4.5.4 RESULTATIVE VS. SEMELFACTIVE

Both submeaning 5.RESULTATIVE and submeaning 6.SEMELFACTIVE represent metaphorical "compression" of the activity designated by the verbal base. Both of them map the spatial image schema COMPACT to the domain of events and activities. Both meanings imply "reduction", or "compression", of the "substance"⁵⁸ of an activity into a single countable portion.

Both RESULTATIVE and SEMELFACTIVE are aspectual meanings, but they designate different kinds of perfectivity. The RESULTATIVE use implies perfectivity via **completion** of an action, as in the verbs *sdelat'* 'complete; produce' (< *delat'* 'do; make') and *svečeret'* 'become dark' (< *večeret'* 'darken'). By contrast, SEMELFACTIVE corresponds to another type of perfectivity via **singularization** of an action (Makarova & Janda 2009: 79). Semelfactive verbs refer to a single quantum of the activity designated by the base, as in *sxitrit'* 'act cunningly one time' (< *xitrit'* 'act cunningly') (Zaliznjak & Šmelev 2000: 118; Nessel 2013).

Nessel (2013: 6) points out that resultative verbs denote "a goal-oriented process that culminates in a change of state". The resultative meaning highlights achieving a natural result and is apparently most tangible in accomplishments, because they are durative, telic (Smith 1997: 46), and imply intermediate stages on the path directed towards the culminating point, the change of state. Therefore, for accomplishment predicates the diagnostic context of the resultative meaning is of the following type: *On el, el, el, i nakonec s"el* 'He was eating, eating eating and finally he ate it'. In my database,

⁵⁸ Lakoff & Johnson (1980: 30) suggest that humans conceptualize events and actions metaphorically as objects, and activities as substances. Janda (2004: 26) refers to this ontological metaphor in her account of the system of Russian aspect (IMPERFECTIVE IS A FLUID SUBSTANCE).

there are several semantic groups of verbs that denote accomplishments and typically manifest the resultative use of the prefix. Such verbs typically refer to:

- (22) EATING: *s'est* 'eat', *sglodat* 'gnaw', *strjamkat* 'eat', *sklevat* 'pick',
sožrat 'devour'
 CHEMICAL PROCESSES: *skisnut* 'sour', *sgnit* 'rot', *skvasit* 'sour',
somlet 'melt; relax'
 DESTRUCTION: *spalit* 'burn down', *srubit* 'cut down', *sžeč* 'burn', *sloamat* 'break',
sokrušit 'destroy'
 COOKING: *svarit* 'cook, boil', *sgotovit* 'cook', *speč* 'bake',
sostrjapat 'cook, concoct'
 CREATING: *sdelat* 'make', *sotvorit* 'create', *skonstruirovat* 'construct',
sozdat 'create'
 FACTITIVES: *srovnjat* 'level off', *sgorbatit* 'bend', *stixnut* 'calm down',
sostarit 'age, make old', *sosvatat* 'make a match'
 INCHOATIVES: *stemnet* 'darken', *sduret* 'become stupefied',
sozret 'ripen'

Note that both S- and SO- can express the 5.RESULTATIVE submeaning and are attested in each of the groups listed above.

The RESULTATIVE use of the prefix S- is so frequent and prominent that this prefix, like the prefix ZA-, should be recognized as a default perfectivizer in Modern Russian. It has been shown that S- is prominently used in loan verbs like *sformirovat* 'form' and *skopirovat* 'copy' (Mizoe 2011). In addition, the Russian National Corpus attests another relevant phenomenon: in many marginal coinages S- is used instead of a prefix that is standardly associated with a particular verbal base: e.g. *sprazdnovat* instead of *otprazdnovat* 'celebrate', *speč* 'bake' instead of *ispeč* 'bake', *spaxat* instead of *vspaxat* 'furrow', and many others. Such formations where the "usual" prefix is replaced with the prefix S- point to the high productivity of the RESULTATIVE S- and support the idea that this prefix functions as a default perfectivizer in Modern Russian. Such verbs are rarely cited and have never been collected. This study makes it possible to collect them and draw more attention to them. For a comprehensive sample of verbs where S- replaces ten other Russian prefixes and expands its RESULTATIVE use see Appendix 2.

Being a default productive perfectivizer presupposes a very generalized and semantically bleached use of the prefix. However, contrary to this assumption, in many resultative verbs the choice of the prefix S- is motivated by its spatial prototypes. Consider the verb *sdelat* which has a very generic semantics 'do'. Apart from completion of an action, this verb designates producing a whole product by assembling its parts (compare synonymous *smasterit* 'craft' and *soorudit* 'construct'). Therefore, the choice of S- in *sdelat* is very natural in the light of its spatial CENTRIPETAL meaning. Moreover, the CENTRIPETAL meaning of the prefix is present in the background of verbs that imply putting things together in the process of creation or cooking. In addition to the relevant illustrations in (22), consider the verbs *svit* 'make a nest' (< *vit* 'plait'), *sšit* 'sew' (< *šit* 'sew'), *smešat* 'mix' (< *mešat* 'mix'), and *sosčitat* 'count all together' (< *sčitat* 'count'). Similarly, the resultative verbs in (22) that denote eating and destruction are based on the DOWNWARD-ABLATIVE and ABLATIVE meanings of the prefix. The spatial cognitive motivation of the RESULTATIVE use of S- is mentioned in previous scholarship (Dickey 2005: 10; Mizou 2011: 40; Janda & Lyashevskaya 2013: 29) and is an important argument against the hypothesis of "pure perfectivizers" (so-called "čistovidovye

pristavki”). Even one of the default perfectivizers, S-, discloses its spatial semantics maintained in the background of many resultative verbs.

As opposed to resultatives, semelfactive verbs typically do not imply a change of state as a culmination point of the event (Nesset 2013: 3). Rather, semelfactives refer to “instantaneous, single-state events consisting of a single point, with no associated change of state” (Smith 1997: 246). In the present study, I follow Nesset (2013) in approaching the semelfactive aktionsart in terms of a radial category with the prototype characterized by four crucial properties: uniformity of sub-events, instantaneousness of the action, non-resultativity, and single occurrence. Following this approach, I recognize that some semelfactive verbs are more prototypical than others and can differ in how they meet the four requirements.

In order to identify semelfactive verbs, I used a number of diagnostic contextual markers of semelfactivity including punctual adverbs like *mgnovenno* ‘instantly’ and *rezko* ‘abruptly’ and adverbial expressions of unexpectedness like *vdrug*, *vnezapno*, *neožidanno* ‘suddenly’. It is worth mentioning that many semelfactives (especially if they refer to behavior like *sglupit* ‘act stupidly once’) do not necessarily imply a single event in a series of repeatable sub-events, as opposed to prototypical semelfactive verbs like *maxnut* ‘wave once’. Instead, semelfactives often refer to a single occurrence, unique in its kind, and, therefore, contextual markers like *odin raz* ‘once’ are also often diagnostic.

In the spatial domain, the 6.SEMELFACTIVE meaning of the prefix S- is very prominent with indeterminate verbs of motion, where the derivatives like *sxodit* ‘S-walk’ and *sletat* ‘S-fly’ refer to a completed round-trip someplace and back (Isačenko 2003/1965: 271; Dickey & Janda 2009: 7). Another domain of verbs that frequently manifests the 6.SEMELFACTIVE S- is the semantic domain of human behavior (Isačenko 2003/1965: 267; Dickey & Janda 2009: 13). Examples of standard verbs are numerous: *symprovizirovat* ‘improvise once’, *s’jazvit* ‘taunt once’, *srobet* ‘act timidly once’, *sparodirovat* ‘parody once’, *sglupit* ‘act stupid once’, *spljasat* ‘dance one time’, *sfoťkat* ‘take a picture’, etc.

The prefix SO- is also attested in semelfactive verbs: compare the verbs *solgat* and *sovrat* ‘tell lies once’, as well as *sosnut* ‘sleep for a short while one time’, *sostroit* (*grimasu/glazki*) ‘make a grimace; flirt with eyes once’, and *sodrognut’sja* ‘shake once’.

Both S- and SO- are attested in novel coinages of semelfactive verbs. Representative examples of SO- are verbs like *sozloradničat* ‘say something mischievously once’, *soskromničat* ‘behave too modestly once’, *sostirnut* ‘do the laundry fast once’. Novel verbs in S- are more numerous and can be illustrated with coinages like *snecenzurničat* ‘say something obscene once’, *sostorožničat* ‘take a cautious attitude once’, *smaxinatorstvovat* ‘do something illegal one time’, *sgrubijanit* ‘say rude things once’, as in (23):

- (23) **Sgrubijanil** on tebe? Èto s nim ne pervyj slučaj. [F.Knorre (1968)]
‘Did he **say rude things** to you? This is not the first time he did it’.

Apart from verbs that demonstrate clearly RESULTATIVE or unambiguously SEMELFACTIVE use of the prefix, some verbs have both readings or allow different interpretations⁵⁹. For example, it is generally difficult to distinguish between the resultative and the semelfactive reading in verbs that refer to achievements, because achievements are typically instantaneous (Smith 1997: 46), and can be both resultative

⁵⁹ For discussion of various challenges encountered in previous scholarship see Isačenko 2003/1965: 270; Dickey & Janda 2009: 10; Nesset 2013: 3.

and semelfactive: e.g. *sogrešit'* 'commit a sin; sin once' < *grešit'* 'sin'. However, even in the case of accomplishment predicates, the resultative and semelfactive interpretations sometimes largely depend on the context. For example, the verb *srepetirovat'* 'rehearse' used in (24) refers to a thorough preparation and arrangements which profile the RESULTATIVE reading of the verb. By contrast, the context in (25) highlights the SEMELFACTIVE reading, because the rehearsal is brief, occasional, and single.

(24) [Ž]izn' vystavki, vključaja "obmeny opytom", èkskursii, vstreči, nagraždenija, predstavala splošnym prazdnikom – xorošo postavlennym, tščatel'no **srepetirovannym**. [M. Čegodaeva. Socrealizm: Mify i real'nost'. (2003)]

'The life of the exhibition including exchange of experience, tours, meetings, and awards appeared as nothing but a holiday, well conducted and thoroughly **rehearsed**.'

(25) [M]y na vsjakij slučaj **srepetirovali** takoj ètjud: vyxodit korol': Gde Gamlet?.. [I. Ševelev, Ju. Ljubimov. Vdol' obryva po-nad propastju... (1998)]

'Just in case we **rehearsed** the following scene. The king comes out: Where is Hamlet?..'

Moreover, the RESULTATIVE and the SEMELFACTIVE readings can be assigned to different senses of the same verb. For example, the verb *skvasit'* 'sour, ferment' is RESULTATIVE when it refers to a chemical process (*skvasit' moloko* 'make sour milk'), but has a SEMELFACTIVE reading, when it is metaphorically used in the expression *skvasit' rožu* 'make a bored facial expression'.

Verbs that allow both 5.RESULTATIVE and 6.SEMELFACTIVE interpretations suggest that there is a close conceptual connection between the two submeanings. However, in the radial category they might have different status. Recall that in Figure 4, 5.RESULTATIVE is located closer to the spatial submeanings 1.CENTRIPETAL MOTION and 2.DOWNWARD-ABLATIVE MOTION than 6.SEMELFACTIVE. There are several reasons for this. First, the 5.RESULTATIVE submeaning is often present in the prefix together with 1.CENTRIPETAL (e.g. *sfokusirovat'* 'focus') or 2.DOWNWARD-ABLATIVE meaning (*svorovat'* 'steal'). Second, historical studies suggest that 6.SEMELFACTIVE use of these prefixes occurred later than the 5.RESULTATIVE use. In particular, according to Dickey and Janda (2009: 19) the first attestations of the semelfactive use of S- date back to the 14th cent., whereas the resultative use of S- is attested already for Old Church Slavonic (Słoński 1937: 225-253; Dickey 2005: 9). Dickey and Janda (2009: 19) point out that the productivity of the semelfactive S- increased in the 16th cent. in Middle Russian. Therefore, the semelfactive use of S- is considered a Russian innovation, whereas West and South Slavic languages did not develop prefixed semelfactives (Dickey & Janda 2009: 19).

4.5.5 Radial Category Profiling

Now that we have established that the submeanings of S- and SO- form a coherent system and are motivated by a single spatial image schema COMPACT, the question remains: how can we compare the two prefixes with regard to these submeanings?

I suggest that S- and SO- are different in terms of which meanings they express most often. In order to assess the relative frequency distribution of the two prefixes across their submeanings in the radial category I apply the *Radial Category Profiling* methodology (Nesset et al. 2011: 21). This methodology relies on the idea that the

number of verbs that represent a certain submeaning of the prefix reflects the strength of association of the prefix with that submeaning (and its relative salience) as opposed to other submeanings in the network of polysemy.

Table 3 is based on standard verbs in the database and summarizes type frequencies (numbers of verbs) of S- and SO- in six submeanings. Table 3 is organized as follows. The leftmost column 1 lists the submeanings, columns 2 and 3 present the raw numbers of standard verbs prefixed exclusively in S- and SO-. Column 4 addresses standard verbs that have paradigms with alternating prefixes S- ~ SO-. In subsequent calculations I leave these verbs aside. Columns 5 and 6 provide percentages of standard verbs prefixed exclusively in S- and SO-.

It is worth mentioning that in this analysis some verbs in the database received more than one semantic tag because they have different readings that manifest different submeanings of the prefix: e.g. 1.CENTRIPETAL and 2.DOWNWARD-ABLATIVE in the case of *ssypat* 'strew into one place; strew down & away', or DOWNWARD-ABLATIVE and 4.ABLATIVE in the case of *s'exat* 'slide down; move out', or 1.CENTRIPETAL and 5.RESULTATIVE in the case of *sšit* 'sew together; sew up (a piece of clothing)'. Without taking into account different uses of the same verb, the analysis would have been incomplete, and the calculation of the overall distribution would be non-representative. Therefore, 440 standard verbs in S- received 515 semantic tags, 96 verbs in SO- received 105 semantic tags, and 20 verbs in S- ~ SO- received 27 semantic tags. The calculations presented in Table 3 are based on the numbers of tags for the verbal lexemes.

Prefix submeaning	S-	SO-	S- ~ SO-	S- %-	SO- %
1.CENTRIPETAL	114	28	10	22%	26%
2.DOWNWARD-ABLATIVE	64	6	4	12.6%	6%
3.CONCOMITANT ACTION	2	21	0	0.4%	20%
4.ABLATIVE	60	6	6	12%	6%
5.RESULTATIVE	182	33	7	35%	31%
6.SEMELFACTIVE	93	11	0	18%	11%
Total:	515	105	27	100%	100%

Table 3: Distribution of standard verbs with S- and SO- in the database.

The relative frequency distribution of S- and SO- in standard verbs is visualized in Figure 13:

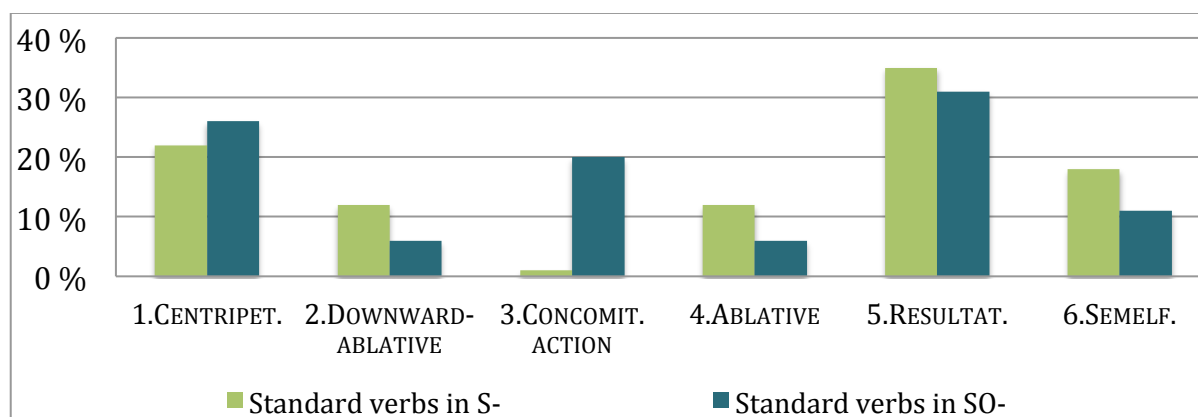


Figure 13: Radial Category Profiles of the prefixes S- and SO- in standard verbs (%).

The overview of semantic distribution of S- and SO- in Table 3 and Figure 13 suggests several generalizations.

First, all six submeanings can be expressed by both prefixes. This is a strong argument for a close semantic relationship between S- and SO-.

Second, raw numbers show that SO- is a more restricted prefix, as opposed to S-, which occurs more often (compare 96 verbs in SO- vs. 440 verbs in S-).

Third, S- and SO- are distributed across their submeanings non-identically. For S-, the most frequent submeanings are 5.RESULTATIVE (182 tags=35%), 1.CENTRIPETAL (114 tags=22%), and 6.SEMELFACTIVE (93 tags=18%). For SO-, the center of gravity includes the submeanings 5.RESULTATIVE (33 tags=31%), 1.CENTRIPETAL (28 tags=26%), and 3.CONCOMITANT ACTION (21 tags=20%). In other words, S- and SO- are both prominent in 5.RESULTATIVE and 1.CENTRIPETAL, but they differ with regard to 6.SEMELFACTIVE and 3.CONCOMITANT ACTION. Recall that the use of S- in the meaning 3.CONCOMITANT ACTION is very marginal. By contrast, SO- in the 6.SEMELFACTIVE use is possible but rather infrequent. In addition, S- and SO- are quite different with regard to the submeanings 2.DOWNWARD-ABLATIVE and 4.ABLATIVE: the prefix S- is twice as frequent in these submeanings as SO-.

Note that the 3.CONCOMITANT ACTION and 6.SEMELFACTIVE submeanings are not attested in verbs that have alternating S- ~ SO- in their paradigms. I suggest that these submeanings developed as later innovations as opposed to the other four submeanings. In other words, the use of S- and SO- in submeanings 3.CONCOMITANT ACTION and 6.SEMELFACTIVE reflects their partial semantic specialization that has emerged with time. However, a detailed account of their historical development is beyond the scope of this study.

Novel coinages indicate productive use of the 6.SEMELFACTIVE S- and the 3.CONCOMITANT SO- and thus provide additional support for the prominence of these submeanings in the repertoires of these prefixes.

Because S- is attested in the meaning 3.CONCOMITANT ACTION very rarely, this creates the strongest contrast between the two prefixes, but also complicates the statistical analysis which does not tolerate too sparse values. For the first trial, I excluded the verbs with 3.CONCOMITANT ACTION meaning of the prefix and subjected the remaining values to a chi-squared test. The result indicates that the difference in the distributions of S- and SO- is statistically non-significant: X-squared = 8, df = 4, p-value = 0.08. For the second trial, the verbs that represent 3.CONCOMITANT ACTION were included in the calculation. This time, the difference in the distributions between S- and SO- was found highly significant: X-squared = 101, df = 5, p-value < 2.2e-16, but R sent a warning message suggesting that the calculation might be imprecise due to the paucity of data. In order to solve this problem, I conflated frequencies of similar submeanings 2.DOWNWARD-ABLATIVE and 4.ABLATIVE, but the warning message on the imprecision of such a calculation remained.

For this reason, the data was instead subjected to the Fisher exact test run for the distributions of S- and SO- in each submeaning. In this analysis, I adopted the methodology widely known as collocation analysis. This method of statistical analysis was proposed by Stefanowitsch and Gries (2003) for calculating the strength of association between a word (collexeme) and a construction where the word appears.⁶⁰ Janda and Lyashevskaya (2013) used this method in order to measure the strength of attraction between five most frequent perfectivizing Russian prefixes and semantic types of verbal bases that combine with these prefixes.

⁶⁰ The model and assumptions of collocation analysis were recently criticized by *Schmid & Küchenhoff* (2013). However, this does not undermine my analysis, because I operate with much smaller values than a usual collocation analysis. Moreover, I explore a finite set of data where the numbers are not hypothetical but real.

Similarly, in the present study, I adopt the methodology used in collostructional analysis in order to estimate the strength of association between the prefix and each of its submeanings. I make use of type frequencies of verbs that represent the “combination” of the form (S- vs. SO-) and the semantic content (six submeanings). For each trial, this type of analysis requires a contingency table with numeric values presented in Table 4: A is the number of verbs that contain S- in the submeaning X; B is the number of verbs that contain a prefix other than S- (i.e. SO- in this study) in the same submeaning X; C is the number of verbs that contain S- in submeanings other than X; and D is the number of verbs that contain a prefix other than S- in submeanings other than X.

	S-	Not S- (=SO-)
Submeaning X	A	B
Not-Submeaning X	C	D

Table 4: Contingency table for the analysis of S- and SO-.

Collostructional analysis compares the distribution observed in the data with the distribution expected due to chance in a hypothesized model (Cantos Gómez 2013: 208). The Fisher exact test is used in order to compute the probability of the observed distribution to occur due to chance (according to the null hypothesis). In the present case, the smaller this probability (p-value), the stronger the association between the prefix and the submeaning.

The analysis targets the distribution of S- and SO- in each submeaning. I focus on verbs that contain exclusively S- or SO-. Calculations are applied to numbers of tags assigned to prefixes in prefixed verbs (recall that some verbs represent more than one submeaning of the prefix). The results are listed in Table 5.⁶¹

Prefix submeaning	Values A, B, C, D	P-value
1.CENTRIPETAL	114; 28; 401; 77	0.3
2.DOWNWARD-ABLATIVE	64; 6; 451; 99	0.06
3.CONCOMITANT ACTION	2; 21; 513; 84	1.997E-15
4.ABLATIVE	60; 6; 455; 99	0.08
5.RESULTATIVE	182; 33; 333; 71	0.5
6.SEMELFACTIVE	93; 11; 422; 94	0.06

Table 5: Results of Fisher exact test for each trial.

Table 5 indicates that a statistically significant association between the prefix and the submeaning is found only for 3.CONCOMITANT ACTION: judging from the distribution, S- is repulsed from this meaning, whereas SO- is strongly attracted to it.

Note that in the case of 2.DOWNWARD-ABLATIVE and 6.SEMELFACTIVE (and to a lesser degree 4.ABLATIVE), *p*-value approaches the threshold for statistical significance (assumed to be 0.05; cf. Cantos Gómez 2013: 210) suggesting that there might be a preference for S- over SO- in these submeanings. In other cases, non-significant *p*-values can be interpreted as indications of neutral form-meaning relationships. In other words, the submeanings 5.RESULTATIVE and 1.CENTRIPETAL do not show any significant preference for S- vs. SO-.

⁶¹ P-values are computed via the Fisher exact test by using the online calculator available at <http://research.microsoft.com/en-us/um/redmond/projects/mscompbio/FisherExactTest/>

What does the statistical analysis tell us? The first test (chi-square test) where I excluded 3.CONCOMITANT ACTION suggests that in most verbs (96.4% of verbs with S- and 80% of verbs in SO-) the choice of the prefix is determined not by any semantic factor but by other factors, arguably phonology and morphophonology. In this light, the situation with S- and SO- resembles what we saw in the analysis of RAZ- and RAZO- in Chapter 3. The second test (Fisher exact test) singles out the submeaning 3.CONCOMITANT ACTION as a domain where S- and SO- are not randomly distributed in terms of semantics. However, overall this domain of semantic divergence comprises a minor part of the category (23 verbs) compared to the zone of semantic similarity.

Summing up, the two prefixes cannot be merely called semantically identical or non-identical. Rather, they exhibit a non-trivial complex relationship, with evidence for an extensive semantic overlap and a strong semantic specialization in a subpart of the radial category.

4.6 Distribution of S- and SO- and conditioning of the prefix

The database contains 998 verbal lexemes. This overall number includes 818 verbs in S- (like *sdelat'* 'make, do'), 159 verbs in SO- (like *sotvorit'* 'create'), and 21 verbs that have alternating S- ~ SO- in their paradigms (like *sognut'*_{PF} – *sgibat'*_{IPF} 'bend').

4.6.1 Prefix alternation inside paradigm

There are nine verbal lexemes where the perfective infinitive contains the prefix S-, but perfective finite forms contain SO-:

(26)	<i>s-šit'</i> _{PF.INF}	–	<i>so-šju</i> _{PF.1PERSON.SG}	'sew'
	<i>s-teret'</i> _{PF.INF}	–	<i>so-tru</i> _{PF.1PERSON.SG}	'rub off'
	<i>s-lit'</i> _{PF.INF}	–	<i>so-lju</i> _{PF.1PERSON.SG}	'pour together or down'
	<i>s-žat'</i> _{PF.INF}	–	<i>so-žmu</i> _{PF.1PERSON.SG}	'press down from all sides'
	<i>s-vit'</i> _{PF.INF}	–	<i>so-v'ju</i> _{PF.1PERSON.SG}	'braid, plait, make a nest'
	<i>s-bit'</i> _{PF.INF}	–	<i>so-b'ju</i> _{PF.1PERSON.SG}	'knock down'
	<i>s-bit'-sja</i> _{PF.INF}	–	<i>so-b'ju-s'</i> _{PF.1PERSON.SG}	'gather in a flock, stray'
	<i>s-čest'</i> _{PF.INF}	–	<i>so-čtu</i> _{PF.1PERSON.SG}	'count'
	<i>s-žeč'</i> _{PF.INF}	–	<i>so-žgu</i> _{PF.1PERSON.SG}	'burn'

Prefix alternation in these verbs is triggered by the vowel/zero alternation in the stem discussed in Chapter 3. Because the occurrence of the vocalized prefix SO- is determined by the presence of this alternation in the adjacent stem, this conditioning of SO- is morphophonemic.

In 12 verbs, the perfective infinitive and finite forms contain SO-, whereas the imperfective counterpart forms (infinitive and finite forms) have S-. This opposition involves the verbs listed in (27).

(27)	<i>so-slat'</i> _{PF.INF}	–	<i>s-sylat'</i> _{IPF.INF}	'send away'
	<i>so-rvat'</i> _{PF.INF}	–	<i>s-ryvat'</i> _{IPF.INF}	'tear off'
	<i>so-žrat'</i> _{PF.INF}	–	<i>s-žirat'</i> _{IPF.INF}	'devour'
	<i>so-drat'</i> _{PF.INF}	–	<i>s-dirat'</i> _{IPF.INF}	'tear off'
	<i>so-gnut'</i> _{PF.INF}	–	<i>s-gibat'</i> _{IPF.INF}	'bend'
	<i>so-gnat'</i> _{PF.INF}	–	<i>s-gonjat'</i> _{IPF.INF}	'chase away; chase all to one place'
	<i>so-mknut'</i> _{PF.INF}	–	<i>s-mykat'</i> _{IPF.INF}	'close, join'

<i>so-brat'</i> PF.INF	– <i>s-birat'</i> / <i>so-birat'</i> IPF.INF	'collect'
<i>so-zvat'</i> PF.INF	– <i>s-zyvat'</i> / <i>so-zyvat'</i> IPF.INF	'call together'
<i>so-jti</i> PF.INF	– <i>s-xodit'</i> IPF.INF	'walk down (e.g. the stairs); get off'
<i>so-jti-s'</i> PF.INF	– <i>s-xodit'-sja</i> IPF.INF	'come together'
<i>so-pnut'</i> PF.INF	– <i>s-pinyvat'</i> IPF.INF	'kick off and downward'

In these verbs (all but *so-jti* and *so-jti-s'*), the vocalized variant SO- is conditioned morphophonologically: it is triggered by the same vowel/zero alternation in the adjacent stem, as in (26). Interestingly, in (27), the morphophonological alternation in prefixes and stems is employed to mark the contrast between perfective and imperfective partner verbs.

The correlation of the vocalized prefix and a non-vocalized stem as opposed to the non-vocalized prefix and a vocalized stem is very consistent. Yet, in two cases we observe a deviation from this pattern: both S- and SO- are possible in the imperfectives *s-zyvat'*/*so-zyvat'* 'call together' and *s-birat'*/*so-birat'* 'collect'. The alternatives are not entirely equal with respect to each other. The "deviating" forms in SO- *sozyvat'* and *sobirat'* are more neutral and more frequently used⁶² than their obsolete counterparts.

In two verbs from (27), *so-jti* 'get off' and *so-jti-s'* 'come together', the vocalized SO- appears not due to the morphophonological factor discussed above, but is rather triggered by a consonant cluster on the morpheme boundary. This type of conditioning is apparently purely phonological. Recall that phonological conditioning of SO- also applies in verbs that have this prefix throughout the entire paradigm, where SO- appears before particular stem-initial consonants (*z*, *šč*) and clusters (*žC*, *sC*), as illustrated earlier in (1) in 4.3.

In (26) and (27), all verbs but one belong to the standard lexicon and have over 500 attestations in the corpus. The only marginal verb in this list is *so-pnut'* 'kick off and down'. This verb follows the pattern of standard verbs, because in this case SO- is attached to a stem that paradigmatically contains a vowel/zero alternation (compare *pnut'* PF – *pinat'* IPF 'kick').

To sum up, the alternation of S- and SO- in paradigms of verbal lexemes, as well as phonological and morphophonological conditioning of SO- strongly suggest that it is an allomorph of S-. I will now turn to the cases of distributional overlap that undermine this generalization.

4.6.2 Prefix variation in minimal pairs

There are 15 verbal stems in the database that can combine with both S- and SO-. Minimal pairs show different types of relationships between the two prefixes.

One type of difference between the two verbal alternatives is the **degree of conventionalization**. For example, in the pair (28), both derivatives are perfective and represent the same CENTRIPETAL meaning of the prefix, but the verb in SO- is standard, whereas the verb in S- is marginal. The difference in their status is reflected in their corpus frequencies, 7,022 attestations and 6 attestations respectively.

⁶² In the Modern Subcorpus of the RNC accessed in 07.2014, *sozyvat'* has 41 hits, whereas *szyvat'* has only 5 hits. Likewise, *sobirat'* has 2,588 attestations, whereas *sbirat'* appears only in 8 occurrences.

- (28) *soedinit'(sja) – s''edinit'(sja)* 'join, bring together' < *edinit'* 'unify', *edinyj* 'single'
 (29) *sopret' – spret'* 'rot, sweat' < *pret'* 'rot, sweat'

Because there is no semantic contrast in (28), the two alternatives are theoretically interchangeable. A similar example comes from the pair in (29), where the verb in SO- is the conventionalized variant (52 corpus attestations), whereas the derivative in S- is marginal (1 attestation). Both S- and SO- contribute the same RESULTATIVE meaning, and the verbs have the same gloss. The derivatives with S- in (28) and (29) are arguably formed via a reduction of the prefix-final vowel in the conventionalized verb form. In other words, SO- is truncated into S-, and this truncation is possible because the two alternatives in these pairs are semantically identical.

As opposed to (28) and (29), in minimal pairs (30)-(33), the verbs in S- are standard and highly frequent, whereas the verbs in SO- are sub-standard (but not marginal because they have more than 9 hits in the corpus): *sodelat'(sja)* – 54 atts., *sodvinut'(sja)* – 13 atts., *sokryt'(sja)* – 346 atts., *sožeč'* – 12 atts. In each pair of verbs, the prefixes have the same meaning. In (30), both prefixes have RESULTATIVE meaning. In (31), both prefixes carry CENTRIPETAL and ABLATIVE meanings. In (32), both prefixes denote ABLATIVE, and in (33) – RESULTATIVE.

- (30) *sdelat'(sja) – sodelat'(sja)* 'make, do' < *delat'* 'do'
 (31) *sdvinut'(sja) – sodvinut'(sja)* 'move to the center, move away' < *dvinit'(sja)* 'move'
 (32) *skryt'(sja) – sokryt'(sja)* 'hide away from' < *kryt'* 'cover'
 (33) *sžeč' – sožeč'* 'burn down' < *žeč'* 'burn'

I suggest that in pairs (30)-(33) the derivatives with S- and SO- are different not only in degree of conventionalization (or entrenchment) but also in terms of **register**: verbs in S- are neutral, while verbs in SO- are stylistically marked and more restricted. For example, in (30), *sodelat'* is mostly used in religious discourse. In (32), *sokryt'(sja)* has an elevated bookish flavor evidenced by examples (34) and (35):

- (34) *Ona sokrylas' v nedrax biblioteki* [Ju. Davydov. Sinie tjul'pany. (1988-1989)]
 'She **disappeared** in the depths of the library.'
 (35) *Tajna postuplenija sokryta ot menja do six por* [N. Ščerbak. Roman s filfakom. (2010)] 'The secret of admission **remains concealed** from me up to the present.'

Corpus examples in (36) and (37) show the same contexts as in (34) and (35) but use the verb *skryt'(sja)*. These examples suggest that the derivatives in S- and SO- in the pair (32) have near-identical semantics and are interchangeable. Crucially, the choice of the prefix is determined by the register that a speaker intends to convey in the text.

- (36) *Ona skrylas' v pljuševyx nedrax kvartiry'...* [V.P. Kataev. Almaznyj moj venec. (1975-1977)] 'She **disappeared** in the plush depths of the apartment.'
 (37) *Kogda on byl pervym mudrecom, postig tajny, skrytye ot drugix.* [V. Piščenko. Zamok užasa (1991)] 'When he was the first wiseman, he got ahold of secrets **concealed** from others.'

Another type of difference between the two alternatives in a minimal pair is often determined by **historical facts** that became part of Modern Russian morphophonology. Consider the pair of verbs in (38), where both derivatives are standard verbs and both prefixes contribute the RESULTATIVE meaning to the base.⁶³ The prefixed verbs in this pair are not interchangeable.

- (38) *sxoronit'*(*sja*) 'hide in a safe place; bury' – *soxranit'*(*sja*) 'keep, preserve'
 < *xoronit'* 'bury' ~ *xranit'* 'keep, preserve'

Note that the simplex verb itself has two alternative allomorphs: the Russian *xoronit'* and the Slavonic *xranit'*. These variants of the stem are historically related and exhibit the *oro* ~ *ra* alternation which is widely attested otherwise in stem allomorphs of parallel Russian and Slavonic origin (e.g. *vorota* ~ *vrata* 'gate'; cf. Itkin 2008: 270). The prefix SO- combines with the stem *xranit'*, which bears Slavonic morphophonological features (metathesis and vowel lengthening). Likewise, S- is attached to the parallel Russian stem *xoronit'* characterized with pleophony typical for East Slavic. In other words, the choice of the prefix in this pair is largely predetermined by the stems carrying certain morphophonological characteristics, which date back to historical reflexes of different origins (South Slavic vs. East Slavic). The prefix SO- is compatible with the Slavonic stem because SO- is itself associated with Slavonic linguistic heritage in Russian. In (30)-(33), this property of SO- reveals itself as a higher register of derivatives, in (38) it comes out as a morphophonological compatibility with stems of the same origin.

In (39), I provide a similar example of morphophonological conditioning of the prefix by the stem which exists in two variants, Slavonic *vraščat'* and Russian *vorotit'*, both with the meaning 'turn, rotate'. Here we observe the same *oro* ~ *ra* alternation and characteristic properties of the Slavonic and Russian stem allomorphs as in (38). Similarly to (38), SO- is attached to the Slavonic stem, and S- is combined with the Russian stem. In both derivative the prefixes carry the meaning DOWNWARD-ABLATIVE.

- (39) *svorotit'* 'displace' – *sovratit'* 'seduce' < *vorotit'*/*vraščat'* 'turn, rotate'

The only difference in the pair (39) is that the verb prefixed with S- refers to concrete spatial movement, whereas the verb with SO- has an idiomatized metaphorical use: it maps the spatial schema of DOWNWARD-ABLATIVE movement to the domain of human relations. As a result, the verb *sovratit'* conceptualizes seduction as a metaphorical deviation from the "proper" behavior. Thus, in (39) the choice of the prefix is conditioned morphophonologically, but this factor additionally creates a contrast in terms of the domain (**concrete or abstract**) that the derivative refers to.

In (40), the morphophonological conditioning of the prefix brings about the contrast in terms of **register** and specialization for **collocations**.

- (40) *svolóč'* 'drag downward; drag to one place from many places' – *sovleč'* 'take off (the clothes, a mask)' < *voločit'* 'drag' / *vleč'* 'attract'

In the Russian verb *svolóč'*, S- can refer to either CENTRIPETAL (41) or DOWNWARD-ABLATIVE (42) movement which applies to all sorts of objects.

⁶³ S- in *sxoronit'* 'hide away' can be alternatively interpreted as ABLATIVE but has RESULTATIVE meaning in *sxoronit'* 'bury'.

(41) [i]x **svolokli** vniz, privezli v oddelenie. [M. Gigolašvili. Čertovo koleso. (2007)]
'They **dragged** them downstairs, transported them to a police-station'

(42) *Lejtenant svolók na kuxnju grjaznuju posudu.* [V. Kornilov. Demobilizacija (1969-1971)] 'The lieutenant **dragged** all dirty dishes to the kitchen.'

By contrast, the Slavonic verb *sovleč'* is obsolete and elevated in register. Moreover, this verb is restricted to the DOWNWARD-ABLATIVE reading of the prefix and to the arguments that refer to garments, bonds, or masks, as illustrated in (43).

(43) [Každyj] *pytaetsja sovleč' s sebja puty svoej sud'by.* [Je. Surkov. Spor o čeloveke (1990-2000)] 'Everyone tries to **take off** the manacles of one's fate.'

The minimal pair in (44) provides another peculiar example of morphophonological compatibility where the opposition of stem allomorphs is accompanied with **semantic contrast** between S- and SO-.

(44) *sprovodit'* 'accompany' – *sprovadit'* 'bundle someone away, get rid of' < *provodit'* 'accompany; say goodbye to'

The opposition of Slavonic and Russian variants of the stem is more tangible if we look at the imperfective partners of the verbs in (44). The verb *sprovodit'* 'accompany' has an imperfective partner verb *soprovoždat'*, which contains the consonant cluster *žd*, the characteristic Slavonic reflex of the Proto-Slavic **dj*. By contrast, the imperfective partner of *sprovadit'* 'get rid of' is *sprovaživat'*, which features the typical East Slavic reflex *ž* of the same historical **dj*. Again, we observe that SO- is attached to the verb associated with the Slavonic stem, whereas S- combines with the Russian stem.

The semantic contrast in (44) appears from different meanings encoded in the prefixes. In *sprovodit'* 'accompany', SO- has the meaning CONCOMITANT ACTION, but in *sprovadit'* 'bundle someone away, get rid of', S- contributes the meaning ABLATIVE. Therefore, the choice of the prefix in (44) is conditioned by both the morphophonology of the adjacent stem and the semantics of the prefix.

Another example of **semantic contrast** of S- and SO- comes from the minimal pair in (45).

(45) *sderžat'(sja)*_{PF} 'restrain, hold off' – *soderžat'(sja)*_{IPF} 'contain' < *deržat'* 'hold'

The verb in SO- has arguably a more abstract general meaning 'contain'. This semantics results from a combination of SO- in the meaning CONCOMITANT ACTION and the semantics of the base 'hold'. The prefix-verb construction 'SO-hold' thus refers to "holding together" different components. It is symptomatic that SO- in this meaning does not alter the aspectual characteristic of the base: the verb remains imperfective. Slavonic prefixes are usually weak perfectivizers, as I show by comparing PERE- and PRE- in Chapter 6. In this light, the use of SO- in the meaning CONCOMITANT ACTION might be a trace of Slavonic origin.

By contrast, the prefix S- turns the same base *deržat'* 'hold' into a perfective derivative and contributes the meaning that can be attributed to the concrete domain of spatial relations, CENTRIPETAL MOTION or ABLATIVE MOTION: e.g. *sderžat' konja* 'restrain a horse', *sderžat' slezy* 'hold back one's tears', *sderžat' protivnika* 'hold off the enemy'. In

addition, the contrast between S- and SO- in (45) can be described in terms of force dynamics (Talmy 1985, 2000). The verb in S- implies a dynamic motion directed against the opposed antagonist force, whereas the verb in SO- refers to a static and balanced relationship between the container and its contents. Thus, in (45) the choice of the prefix is determined by the target semantics associated with the derivative.

Summing up, in minimal pairs the prefixes S- and SO- show contrast in conventionalization degree, register, concreteness/abstractness, collocational restrictions, and semantics. The choice of the prefix can be conditioned by morphophonological properties of the stem and by specialization of the prefix for high/neutral register or/and particular semantics (e.g. CONCOMITANT ACTION meaning of SO-).

4.6.3 S- and SO- stacked in one verb

Most verbs in the database attach only S- or SO-. Interestingly, there are two verbs attested in the corpus that contain both prefixes at once. One of them is a marginal novel verb *sosprygnut'* 'jump down once' formed from the base *prygnut* 'jump once'. It represents the prefix SO- stacked over the prefix S-.

In this verb, SO- is phonologically conditioned, because it appears before a cluster that starts in the fricative obstruent *s* (compare *soskočit'* 'jump down, jump off' < *skakat'* 'jump'; recall (1)). The use of SO- in this verb is semantically redundant: the prefix S- attached to *prygnut* 'jump once' alone contributes the DOWNWARD-ABLATIVE meaning to the base. In a sense, SO- duplicates the same meaning and apparently puts additional emphasis on it. This intensifying effect of double prefixation is apparently employed in (46) for communicative purposes:

- (46) [V]spomnil ja staryj anekdot pro čeloveka⁶⁴, kotoryj poobeščal <...> čto sprygnut s Isakievskogo sobora. Kogda on zabralsja tuda, to zajavil: "Ob **sosprygnut'** ne možet byt' i reči. Vopros v tom, kak slezt' vniz." [A. Gorodnickij, "I žit' ešče nadežde" (2001)]

'I recalled an old joke about a person who promised to jump down from the St. Isaac's Cathedral. When he climbed up there, he announced: "**Jumping down (lit. down-down-jump)** is out of question. The question is how to climb downstairs.'

In addition to representing an overlap of S- and SO- in a single derivative, this example is relevant for the discussion of prefix stacking. In particular, if we assume that in (46) SO- carries the same meaning as S- DOWNWARD-ABLATIVE, this novel verb provides a piece of counter-evidence against the traditional assumption about prefix stacking. It has been argued that only a so-called superlexical prefix⁶⁵ can stack over another prefix (Svenonius 2004: 206, 229; Romanova 2004: 257, 265). Contrary to this account, however, SO- is stacked in (46) in its lexical (spatial) meaning DOWNWARD-ABLATIVE.

⁶⁴ I replace the reference to nationality with a general term for purposes of political correctness.

⁶⁵ Superlexical prefixes are defined as prefixes that have functional aspectual and quantificational meanings (semelfactive, ingressive, delimitative, distributive, etc.) and originate outside a verbal phrase. By contrast, lexical prefixes have spatial meanings and are structurally internal to a verbal phrase (Svenonius 2004). Proponents of this distinction admit problems with assigning specific uses of some prefixes to these two classes. For example, Romanova (2004: 269) discusses attenuative uses of Russian PRI- and POD- pointing out that they "seem to combine" properties of both classes.

The other verb that contains both prefixes S- and SO- is *sosborit'* and refers to making tucks and gathers on the fabric. This verb is formed from the prefixed base *sborit'* 'gather the fabric into tucks' which is related to the simplex *brat'* 'take'.⁶⁶ Thus, the derivative *sosborit'* is analyzable as SO-S-'take', where both prefixes carry the meaning CENTRIPETAL ACTION. In other words, this is another counter-example that features stacking of the lexical prefix.⁶⁷ In addition, this example provides evidence for the distributional overlap of S- and SO-.

4.6.4 Prefix variation in poetry

In poetry, the choice between verbs in S- and SO- is often made for the sake of maintaining rhythm. In particular, we can observe that a more standard derivative in S- can be replaced for metrical purposes by a less common parallel verb in SO-. Crucially, this replacement does not cause any misunderstanding. The method is quite old, as evidenced by (47), and is frequently used by modern poets, as illustrated in (48).

(47) *Za serebrjany monety **sokujut** tebjje braslety.* [F.M. Rešetnikov. *Gde lučše?* (1868)]
'For silver coins they will **forge** you chain bracelets.'

(48) *V prisutstvii železa i stekla My kofe p'jem, **sodvinuv** dva stola.* [A.P. Mežirov. «Verijskij spusk v snegu...» (1961-1981)] 'In the presence of iron and glass We are drinking coffee after **having put together** two tables.'

On the one hand, this phenomenon is another area of distributional overlap of the two prefixes and violates the criterion for complementary distribution. On the other hand, it shows that the S- and SO- are interchangeable and thus semantically identical which satisfies the criterion for identical meaning.

4.6.5 Verbs that exclusively attach SO-

What factors trigger the prefix SO- in verbs that contain this prefix in all forms and lack parallel derivatives in S-? There are 82 such verbs in the database (15% of all 556 standard verbs in S- and SO-). I analyzed each of them with regard to what factor conditions the appearance of SO-. Table 6 shows the distribution of different types of conditioning of SO- in standard verbs.

Conditioning type	# of verbs	%
Phonology	29	35%
Morphophonology	9	11%
Semantics (CONCOMITANT ACTION)	18	22%
Other (CENTRIPETAL; RESULTATIVE of Slavonic heritage)	26	32%
Total	82	100%

Table 6: Distribution of factors that condition the prefix SO-.

⁶⁶ The derivation here is more complex because there are two more steps in between two prefixations.

⁶⁷ A related non-trivial phenomenon involves verbs where the prefix SO- is stacked over a prefix of similar semantics. For example, in the verb *sopričislit'* 'add, assign something to', SO- carries the meaning CENTRIPETAL ACTION and is stacked over the prefix PRI- in the meaning ADD: SO-PRI- 'count/number'. This verb provides another example of stacking the lexical prefix SO- over another prefix.

According to my calculations, phonological conditioning is attested most frequently, in 35% of cases. Phonologically conditioned SO- occurs before stem-initial fricative obstruents *s*, *z*, *šč* and clusters *žC* and *sC* (recall Itkin 2007 and examples in (1)).

Morphophonology motivates the use of the prefix SO- with Slavonic stems, as in *sovladat'* 'cope with', *soglasit'sja* 'agree with', *sokratit'* 'reduce'. Note that morphophonology is a more powerful factor in conditioning SO- than it appears from Table 6. Morphophonological compatibility with the stem governs the use of SO- in 20 verbs with alternating S- ~ SO- and in most of 15 minimal pairs discussed above.

Semantic conditioning of SO- refers to those cases when this prefix clearly carries the meaning CONCOMITANT ACTION and is not triggered by phonological context. For example, in *součastvovat'* 'participate together with someone', *sotrudničat'* 'collaborate', *sopereživat'* 'sympathize'. These cases yield 22% of the data.

Under the "other" conditioning type I group those verbs in SO- that are not explained by the three previous factors. In most of these verbs, the prefix SO- implies putting things together and thus carries the meaning CENTRIPETAL but not CONCOMITANT ACTION, as whown in (49).

(49) <i>so-pri-kosnut'sja</i> PF 'touch with'	< <i>prikosnut'sja</i> PF 'touch'
<i>so-udarit'sja</i> PF 'collide with something'	< <i>udarit'(sja)</i> PF 'hit'
<i>so-po-stavit'</i> PF 'compare'	< <i>postavit'</i> PF 'put'
<i>so-ot-nesti</i> PF 'relate to'	< <i>otnesti</i> PF 'carry to'
<i>so-ob(-)razovat'</i> PF 'coordinate'	< <i>obrazovat'</i> PF 'create'
<i>so-iz-merit'</i> PF 'evaluate with regard to each other'	< <i>izmerit'</i> PF 'measure'
<i>so-v-mestit</i> PF 'combine with'	< <i>vmestit'</i> PF 'contain'
<i>so-v-past'</i> PF 'coincide with'	< <i>vpast'</i> PF 'fall in'
<i>so-prjač'</i> PF 'conjoin'	< <i>prjač'</i> IPF 'harness, connect'

In (50), I list verbs where the meaning of SO- is related to CENTRIPETAL, but from the synchronic perspective should be rather considered RESULTATIVE:

(50) <i>so-ob(-)razit'</i> PF 'figure out, coordinate'	< <i>obraz</i> 'image'
<i>so-orudit'</i> PF 'build, construct'	< <i>orudie</i> 'tool', <i>orudovat'</i> IPF 'handle'
<i>so-tvorit'</i> PF 'create'	< <i>tvorit'</i> IPF 'create'
<i>so-obščit'</i> PF 'inform, share news with'	< <i>obščij</i> 'shared'
<i>so-iz-volit'</i> PF 'allow'	< <i>izvolit'</i> PF 'deign'
<i>so-činit'</i> PF 'create'	< <i>činit'</i> IPF 'do'
<i>so-bljusti</i> PF 'follow regulations'	< <i>bljusti</i> IPF 'comply with, maintain'
<i>so-grešit'</i> PF 'sin'	< <i>grešit'</i> IPF 'sin'
<i>so-krušit'</i> PF 'destroy'	< <i>krušit'</i> IPF 'destroy'

Note that verbs in (49) and (50) are mostly perfective, as opposed to verbs that feature SO- in the meaning CONCOMITANT ACTION. However, most of these perfective verbs are formed from perfective base verbs, so the prefix SO- does not alter the aspect of the stem. Moreover, in most of these perfective verbs SO- is stacked over another prefix. Both characteristics are typical properties of Slavonic prefixes (compare PRE- in Chapter 6).

I conclude that SO- in (49) and (50) is historically conditioned, because these verbs belong to Slavonic heritage. Crucially, this data shows that Slavonic SO- is not restricted to the meaning CONCOMITANT ACTION, but is often attested in the meanings CENTRIPETAL and RESULTATIVE. In this light, the Slavonic SO- is not monosemantic (recall

Valedinskaja & Golanova 2007) and shares with S- all three submeanings (CENTRIPETAL and RESULTATIVE frequently and CONCOMITANT ACTION marginally). This undermines the account that views Slavonic SO- as a distinct morpheme unrelated to Russian S-.

In this section I have looked at factors that condition SO- in standard lexicon. In standard verbs, phonology governs only 35% of SO-, whereas the remaining 65% are conditioned by factors grounded in history. Synchronic morphophonology accommodates reflexes of historical alternations and correlations of stem allomorphs. Semantic factors foreground those submeanings that are most entrenched for SO- in the Slavonic sublexicon of Modern Russian.

In future research, it would be fruitful to examine marginal verbs in SO- and compare the distribution of conditioning factors. The choice of the prefix is especially relevant for novel coinages, where conditioning factors compete. Just to give a flavor of this competition, consider the perfective verb *svredničat'* 'behave maliciously once' formed from the base *vredničat'* 'behave maliciously'.

- (51) *Spičečno-korobkovyj televizor mgnovenno svredničal, zajaviv, čto «Žiliet» lučše.*
[A. Izmajlov. Trjukač (2001)] 'The matchbox size TV immediately **played a nasty trick** by claiming that *Gillette* is better.'

In this verb, the base has a Slavonic stem (*vred* 'harm' ~ *vered* 'abscess') that should trigger the prefix SO-. However, in this case morphophonology is outcompeted by semantics, which suggests that the meaning SEMELFACTIVE can be best communicated by S-. This example suggests that the system is flexible and conditioning factors do not have absolute power.

4.7 Conclusions

In this chapter, I have presented a case study of the Russian prefixes S- and SO-. I carried out the first corpus-based study of their relationship and analyzed their semantic and distributional patterns from the perspective of Cognitive Linguistics.

I found that the complex behavior of these prefixes challenges the traditional conception of allomorphy based on the absolute criteria of identical meaning and complementary distribution. In the case of S- and SO-, both criteria for allomorphy are difficult to assess.

The analysis of prefix distribution and conditioning shows that S- and SO- exhibit non-trivial behavior, which partly conforms to the distributional criterion of allomorphy, and partly violates it. On the one hand, the two prefixes often alternate in forms of the same verb and are conditioned by phonological and morphophonological factors, similarly to RAZ- and RAZO- discussed in Chapter 3. However, only these patterns can be uncontroversially attributed to Standard Allomorphy. On the other hand, we encounter examples of minimal pairs, overlap in prefix stacking, and prefix replacements in poetry that violate the requirement for complementary distribution. Moreover, instead of straightforward monofactorial conditioning we deal with a bundle of competing relevant factors which include historical, register, and semantic parameters. Arguably, the combination of different trends of distribution takes place because SO- combines uses of Russian and Slavonic origin. However, this does not resolve the issue with SO- from a synchronic perspective.

The other categorical criterion of allomorphy, namely whether the candidates are semantically identical or not, is likewise difficult to assess. An accurate answer is not in harmony with either of the two possibilities. On the one hand, the two prefixes are not

semantically different because they share all six submeanings in the radial category. On the other hand, the two prefixes are not semantically identical because they are partly specialized for particular submeanings (especially SO- for the meaning CONCOMITANT ACTION). Overall, in terms of datapoints, the zone of semantic overlap between S- and SO- is larger than the zone of their semantic divergence. In this complex system of correspondences, the specialization of one prefix (SO-) for a particular submeaning is a part of the dynamic development of the relationship of related forms. In this development, it is semiotically natural for the elements that differ in form to also develop partial semantic differences. In this study, I have argued that all submeanings including CONCOMITANT ACTION are closely related to each other both conceptually, via the cognitive mechanisms of metaphor and metonymy, and by means of verbs that simultaneously represent more than one submeaning of the prefix. In view of this, the coherent model of polysemy that I advocate in this chapter prevents me from analyzing S- and SO- as distinct morphemes.

What do we gain from this study? First of all, this empirical case study has theoretical implications. It enriches our knowledge about what types of relationships are to be found between linguistic elements. Moreover, I propose that this case study exemplifies how data can deviate from the mainstream model of allomorphy. Contrary to the conventional line of categorical assessment of definitional criteria, I suggest that the relationship of S- and SO- in Modern Russian can be best captured by extending the traditional concept of allomorphy. In particular, I propose that this case should be viewed as an exemplar of Non-Standard Allomorphy rather than Non-Allomorphy (recall the discussion of these terms in Chapters 1 and 2).

In future research it would be fruitful to explore the relationship between S-/SO- and another historically related prefix SU- found in nouns like *suglinok* 'clayish soil' (< *glina* 'clay') and *sumrak* 'twilight' (< *mrak* 'darkness'). Another attractive line of research would undertake a close analysis of marginal verbs formed by S- and SO- that largely remain beyond the scope of the present study. Lastly, it is reasonable to take a diachronic perspective and compare these prefixes in Modern Russian to the Church Slavonic prefix СЪ- in order to track the changes to this morpheme that shaped the non-trivial allomorphy it exhibits today.

Chapter 5

At the crossroads of phonology and semantics: The prefixes O- and OB- ‘around’

5.1 Introduction: The puzzle of the prefixes O- and OB-

This chapter⁶⁸ examines an interesting phenomenon in Russian morphology: the behavior of the prefixes O- and OB-. Do these prefixes constitute two distinct morphemes or are they allomorphs of a single morpheme? Previous discussion in scholarly works has been limited to lexical data provided in dictionaries. A number of controversial uses of these prefixes were discovered and led to the hypothesis that the original single morpheme has split into two distinct morphemes (Markov 1970; Alekseeva 1978; Andrews 1984; Krongauz 1998). This chapter summarizes the results of two empirical tests of this hypothesis: first, against a comprehensive dataset extracted from the Russian National Corpus; secondly, via a psycholinguistic experiment. The results of both studies do not give compelling support for the predictions of the Split Hypothesis. I offer a detailed analysis of the radial network of submeanings expressed by the two prefixes, an independent account for the subject responses to experimental stimuli, as well as statistically robust analyses of both studies. On this basis I propose an alternative hypothesis of a single morpheme with Non-Standard Allomorphy and argue for a unified account of O- and OB-. The study of this particular problem of Russian morphology encourages us to reevaluate the traditional understanding of allomorphy coined in absolute terms.

The question of whether the prefixes O- and OB- represent a single morpheme or two distinct morphemes in Modern Russian is not as simple as it might look at first glance. A study of the literature reveals a long-standing debate that engaged not only many Russian linguists (Markov 1970; Aver'janova 1974; Alekseeva 1978; Nefed'ev 1994, 1995), but also Slavists worldwide (Roberts 1976, 1981; Hougard 1973; Andrews 1984). A growing number of recent studies devoted to Russian O- and OB- (Krongauz 1998, Itkin 2007; Dobrušina 2009; Shull 2003; Košev 2004a,b, Baydimirova 2010; Baayen et al. 2013) show that interest in the status of these two prefixes has grown immensely in the last decade. There are new parallel studies of corresponding related prefixes in other Slavic languages as well (Będkowska-Kopczyk 2012; Będkowska-Kopczyk & Lewandowski 2012; Šarić et al. 2013).

This profound interest in these two prefixes is motivated by the non-trivial interaction of phonological and semantic factors which condition the use of O- and OB-. On the one hand, as Krongauz (1998: 132, 141) fairly points out, the two prefixes regularly appear in verbal forms that belong to the same paradigms:

⁶⁸ The results of this study were presented at a number of conferences: the study on prefixes was reported on at the 15th International Philological Conference at St. Petersburg State University in 2011, in the Summer seminar of the Norwegian Cognitive Linguistic Association (NORKOG) in Oslo in 2012, at the 4th Cognitive Linguistics Conference (UK-CLC) at State College London in 2012. I am grateful to all audiences for fruitful discussions. I am especially grateful to the anonymous reviewers of the journal “Voprosy Jazykoznanija” for their comments (cf. Endresen 2013).

(1)	<i>obo-gnat'</i> _{INF.PF}	– <i>ob-gonju</i> _{SG1.PF}	'leave behind, outstrip'
	<i>obo-brat'</i> _{INF.PF}	– <i>o-beru</i> _{SG1.PF}	'pick, gather, rob'
	<i>o(b)-bit'</i> _{INF.PF}	– <i>obo-b'ju</i> _{SG1.PF}	'upholster, plank, beat off'
	<i>o-peret'sja</i> _{INF.PF}	– <i>obo-prus'</i> _{SG1.PF}	'lean on'
	<i>ob-žeč'</i> _{INF.PF}	– <i>obo-žgu</i> _{SG1.PF}	'burn'

Note that in these examples OBO- alternates with both O- and OB-. The use of these prefixes in forms of the same verbal lexemes suggests that they carry identical meaning. The choice of the prefix depends on the morphophonemic context, namely the initial phonemes of the simplex base that a prefix is attached to and the vowel/zero alternation in the root morpheme (recall Chapter 3). Moreover, there exist pairs of verbs like *ob-strič'* and *o-strič'* 'cut off', *ob-strogat'* and *o-strogat'* 'plane a log', *ob-kleit'* and *o-kleit'* 'glue'. Within each of these pairs the semantic contribution of O- and OB- is essentially identical. Given these examples, it is reasonable to conclude that O- and OB-, as well as the vocalized⁶⁹ variant OBO-, are positional variants of a single morpheme.

On the other hand, as we see from the verbs *o-krásit'*_{PF} 'color', *ob-krádyvat'*_{IPF} 'rob', and *obo-krást'*_{PF} 'rob', all three prefixes can attach to simplex bases with the same initial consonant cluster and the same place of stress. As another piece of evidence that the phonological conditioning tolerates some variation, consider the verbs *o-bežat'* and *ob-bežat'* 'run around smth'. They represent two pronunciation variants of a single verbal lexeme. The two forms are used in free variation and in many contexts lack any semantic contrast. In fact, the variant *ob-bežat'* with the geminate labial consonant is usually characterized as more colloquial. The consonant-final shape of the prefix in this environment is considered non-standard, but it occurs quite frequently in spoken Modern Russian. These instances of an overlap in the use of O- and OB- indicate that these prefixes do not meet the requirement for perfect complementary distribution, contrary to what we expect for proper allomorphs. In addition, Russian has minimal pairs of verbs which show that O- and OB-/OBO- can modify the meaning of the simplex base in distinctly different ways:

(2)	<i>o-sudit'</i> _{PF} 'sentence'	vs. <i>ob-sudit'</i> _{PF} 'discuss'	< <i>sudit'</i> _{IPF} 'judge'
	<i>o-delit'</i> _{PF} 'present with'	vs. <i>ob-delit'</i> _{PF} 'deprive'	< <i>delit'</i> _{IPF} 'divide'
	<i>o-gret'</i> _{PF} 'hit smb hard'	vs. <i>obo-gret'</i> _{PF} 'heat'	< <i>gret'</i> _{IPF} 'heat'

The semantic differences between these prefixed verbs are often interpreted as evidence that O- and OB- are not allomorphs but rather two distinct morphemes that carry different meanings.

From the perspective of the two traditionally cited criteria for allomorphy, identical meaning and complementary distribution (Matthews 1974: 116; Bauer 2001: 14; Haspelmath 2002: 27; cf. the discussion in Chapter 2), the two Russian prefixes present us with a puzzle: their semantics can be absolutely identical but sometimes turns out to be strikingly different, while their distribution, though partly influenced by phonology, can tolerate some overlap and thus does not fit the narrow bounds of perfect complementation. We can conclude that neither of the two traditional criteria of Standard Allomorphy is satisfied in this case. Note that the data leads us to contradictory

⁶⁹ Recall from Chapter 3 that I use the term *vocalized* in order to refer to the longest form of the prefix, that is the form that ends in a vowel, as in RAZ- ~ RAZO-.

conclusions. The examples in (1) suggest that O- and OB- alternate within the paradigm and must represent a single morpheme, whereas the verbs in (2) demonstrate that O- and OB- exhibit semantic contrast and should be interpreted as different morphemes. Yet, intuitively the two prefixes are associated as similar and closely related. However, one cannot deny their differences. A question arises as to what predominates and has more explanatory power – the differences or the similarities between the two affixes?

It is no surprise that the issue of the morphological status of O- and OB- has generated a long-standing controversy in the literature. Some scholars argue that O- and OB- are phonological variants of a single morpheme (Zaliznjak & Šmelev 2000: 83; Vinogradov et al. 1952: 589 – 592; Isačenko 2003/1965: 148; Roberts 1976, 1981; Townsend 1968: 127; Timberlake 2004: 404; Barykina et al. 1989), while others insist on distinguishing them as two separate morphemes: O- and OB-/OBO- (Švedova et al. 1980: §851). The two opposite views have been partially combined within a Hypothesis of Morphological Split. Henceforth I will refer to it as the Split Hypothesis. This hypothesis explains the contradictions in the synchronic behavior of O- and OB- in terms of a diachronic process of divergence of the two allomorphs that originally belonged to a single morpheme *OB-. As a consequence of this divergence, or split, Modern Russian ends up with two distinct morphemes O- and OB-, which differ both in terms of semantics and phonology.

The Split Hypothesis is very elegant, and it was backed up by a number of thorough studies. Markov 1970 and Alekseeva 1978 provided diachronic arguments for the morphological split. Andrews 1984 conducted a detailed analysis of minimal pairs of verbs prefixed in O- and OB- in Modern Russian (e.g. *o-sudit* ‘condemn’ vs. *ob-sudit* ‘discuss’, *oledenet* ‘freeze to ice’ vs. *obledenet* ‘become covered with ice’). Krongauz (1998: 131-148) described the semantic characteristics of the two distinct morphemes as well as the hierarchy and phonological conditioning of their allomorphs. However, despite all these studies that elaborated the Split Hypothesis, the issue of the mutual relationship between O- and OB- still remains debatable, as pointed out by I.B. Itkin (2007: 84). Indeed, there are new studies that argue for or build upon the idea of semantic unity and coherence of the two prefixes (Košelev 2004b; Dobrušina 2009). Yet, on the other hand, there are also new proponents of the two-morphemes-account (Itkin 2007: 85). This means that the controversy about the synchronic relationship between O- and OB- is not resolved yet, and the problem remains. The likely reason is that the Split Hypothesis has not been rigorously tested against a large dataset. This is precisely the objective of the present study. I address the question of whether O- and OB- are positional variants or independent morphemes and investigate this in the light of new data from the corpus and a psycholinguistic experiment.

In this chapter, I report on the results of tests of the Split Hypothesis from two different perspectives. First, I examine comprehensive data from the Modern Russian lexicon, namely all perfective verbs prefixed in O-, OB-, and OBO- attested in the Russian National Corpus (www.ruscorpora.ru) and *the Grammatical Dictionary of the Russian Language* (Zaliznjak 1980). Second, I test the Split Hypothesis experimentally against the active mechanisms of word production present in native speakers’ mental grammar. I demonstrate that the findings of both case studies give some partial support for the predictions of the Split Hypothesis, but at the same time contradict it to an even greater degree. My results make it possible to propose an alternative solution – the Hypothesis of Non-Standard Allomorphy, which implies that the traditional theoretical understanding of what allomorphy is should be revisited.

This chapter is organized as follows. In 5.2, I present the Split Hypothesis with focus on its drawbacks and predictions. Section 5.3 is devoted to the corpus study. I detail my methodology of data collection, address the conflicting phonological and semantic factors that affect the distribution of prefixes, describe my model of prefix polysemy for O-, OB-, and OBO-, and present a statistical analysis of their distribution across the recognized submeanings. In 5.4, I describe how the experiment was designed and administered in order to address the predictions of the Split Hypothesis. The discussion of experimental results focuses on major trends and variation across subjects. In 5.5, the findings from both case studies are compared. I outline the alternative account for the prefixes' relationship in terms of Non-Standard allomorphy and argue that it is more accurate in capturing their complex behaviour.

All materials of both corpus and experimental studies are available at <http://hdl.handle.net/10037.1/10078>. There one can find the tagged database of the analyzed prefixed verbs, the database of collected subjects' responses, and the R codes for the statistical analyses. Stimuli and nonce words used in the experiment are presented in Appendices 3 and 4.

Before I turn to the next section I should comment on my use of terminology. In this study I adopt an agnostic view concerning the morphological status of the prefixes in question. Therefore, I use the term *prefix* neutrally, without assigning the formants O-, OB-, and OBO- any anticipatory status as variants or distinct morphemes. In order to refer to their morphemic status I use the terms *morpheme* and *allomorph*.

5.2 The Split Hypothesis

5.2.1 Proponents and their arguments

The Split Hypothesis for the prefixes O- and OB- was proposed in Markov 1970, Alekseeva 1978, Andrews 1984, and Krongauz 1998. According to this hypothesis, in the history of Russian the allomorphs of the initially single morpheme *OB- underwent a process of divergence and eventually split into two morphemes. As a result, in Modern Russian O- and OB- became semantically disjoined and dissociated. The prefix OB- became specialized for the spatial meaning of circular motion: e.g. *ob-letet* 'fly around' < *letet* 'fly'. Meanwhile, the prefix O- became a part of morphological constructions⁷⁰ O-...-IT' and O-...-ET' employed for formation of factitive and inchoative verbs respectively which denote an imposition or acquisition of a new quality by an object: *o-žestočit* 'make cruel' < *žestokij* 'cruel'; *okamenet* 'petrify' < *kamen* 'stone'.

Markov 1970 was the first to comment on the unequal synchronic productivity of O- and OB- in two different derivational patterns. They can be illustrated with the following examples. The verb *ob-glodat* 'gnaw around' (*ob-* 'round' + *glodat* 'gnaw') is modelled after the verbs of spatial motion like *ob-letet* 'fly around' (*ob-* 'round' + *letet* 'fly') with the prefix OB-, contrary to the factitive verb *o-vinovat-i-t'* 'make one guilty' (*o-* + adjective *vinovat(yj)* 'guilty' + *i + t'*) formed analogously after factitives like *o-svetl-i-t'*

⁷⁰ A morpheme that combines a prefix and a suffix into a construction can be referred to in a number of different ways. In Russian linguistic literature such morphemes appear in "*prefixal-suffixal pattern of derivation*" ("pristavočno-suffiksál'nyj sposob slovoobrazovanija") (Zemskaja 2006: 85; Nemčenko 1979: 39) and are called *konfiks* in the Kazan' linguistic school (cf. bibliography in Markov 1970: 7) and *circumfix* according to the Moscow linguistic school (Uluxanov 1996: 132). I will refer to a two-componential morpheme of this type as a *morphological construction* following Booij 2010.

'lighten' (*o-* + adj. *svetl(yj)* 'light' + *i + t*'), choosing *O-*, but not *OB-*. Since Markov's article (1970), the complementary productivity of *O-* and *OB-* in these different derivational patterns has become the central argument of the Split Hypothesis. Markov himself views this trend as the major criterion for morphological distinction and independence of modern *O-* and *OB-*. Along these lines, he suggests that the parallel transitional intermediary forms and meanings ("perežitocnye formy i značeniya")⁷¹ that coexist in Modern Russian with the outlined pattern can be neglected. In making this statement, Markov in fact extrapolates his argument about the two productive patterns to the entire available data, and this step is very problematic. Indeed, the so-called transitional intermediary forms, as I show further, constitute a large part of speakers' linguistic input and competence. While a linguist can perhaps be selective in terms of which data to account for, it is very unlikely that the speakers can likewise abstract themselves away from the "transitional" linguistic input when developing their linguistic competence. However, the selective approach to data has become standard in subsequent inquiries of this issue.

The idea that two phonological variants of a single affix can develop into autonomous morphemes was first proposed in Markov's paper (1970) and further elaborated in a dissertation by Alekseeva (1978) written under his supervision. Alekseeva 1978 investigates the use of the two prefixes in deverbal and deadjectival derivation starting from the oldest Russian and Church Slavonic texts all the way up to Soviet-era fiction and journalism of 1970s. This study does not provide a detailed chronology of semantic and morphological split of the two prefixes. Nevertheless, it lends support to Markov's point that the phonological conditioning of *O-* and *OB-* was inconsistent already in the early Old Russian period (Markov 1970: 5, Alekseeva 1978: 6). In addition, Alekseeva describes the following key tendency: a number of prefixed verbs motivated in Old Russian by simplex verbal bases (e.g. *obagriti* < *bagriti* 'empurple', *obogatiti* < *bogatiti* 'enrich', *odobriti* < *dobriti* 'approve', *ožestočiti* < *žestočiti* 'obdurate', *oživiti* < *živiti* 'enliven', *ozariti* < *zariti* 'enlighten', *o(b)rositi* < *rositi* 'moisten', *osvoboditi* < *svoboditi* 'liberate', *osvoiti* < *svoiti* 'familiarize') later became additionally associated with nominal and adjectival bases (e.g. *obogatiti* 'enrich' < *bogatyj* 'rich'). And so they came to represent a new derivational type, especially after the initial motivating verbal bases had gone out of use or become marginal (ibid: 6). Alekseeva (1978: 10) says that *O-* and *OB-* developed as distinct morphemes already in Old Russian. At the same time, as she points out, many verbs like *orumjanit'* 'redden' and *ozelenit'* 'make green' still preserve multiple derivational links in Contemporary Russian, being associated both with a verbal simplex (*rumjanit'* 'redden', *zelenit'* 'make green') and a nominal or adjectival stem (*rumjanyj* 'ruddy', *rumjana* 'rouge', *zelenyj* 'green') (ibid: 12). However, the existence of such transitional cases does not prevent Alekseeva from positing the clear contrast, opposition and distinctiveness for the two morphemes *O-* and *OB-* synchronically.

It is worth pointing out that in Alekseeva 1978 the number of such transitional cases where the verbs with factitive ('make X be Y') and inchoative ('become Y') semantics are synchronically motivated by both nominal and verbal bases is severely underestimated. This shortcoming is exacerbated by the fact that Alekseeva deliberately

⁷¹ Markov does not provide any specific examples of intermediary forms and meanings. However, it is clear from his discussion that he refers to those verbs where the use of *O-* and *OB-* does not conform to the two patterns (spatial and factitive) that he outlined.

leaves out all verbs⁷² where the function of the prefix is assumed to be “purely aspectual” (ibid: 12), in other words bearing no other meaning. A vast majority of these excluded verbs are factitives and inchoatives. Thus, the prefixes in *ogloxnut* ‘deafen’ and *obednet* ‘become poor’ are assumed to lack any meaning beyond marking perfective aspect, because they form Natural Perfectives (following terminology of Janda 2007) from simplexes *gloxnut* ‘deafen’ and *bednet* ‘become poor’. The idea that the semantic emptiness of prefixes in Natural Perfectives is an illusion has been persuasively argued in a number of recent studies on Russian aspectual prefixes (cf. Janda 2012, Endresen et al. 2012; Janda et al. 2013 and their bibliography). Regarding O- and OB- in particular, the problem of semantic emptiness was addressed in Roberts 1976: 65 and disputed in Baydimirova 2010: 56-58. So, in this light, it appears that the conclusions in Alekseeva 1978 are based on a limited and biased dataset. Moreover, although tracking the partial opposition of O- and OB- is certainly valuable, it is overgeneralized to the entire system, with focus only on their contrast, whereas intermediate and less clear cases of overlap receive minimal attention and in fact are not accounted for within the proposed model.

Andrews 1984 argues for the Split Hypothesis independently from Russian scholars and in different terms. She explicitly motivates the search for semantic differences between Russian O- and OB- by means of semiotic theory: the crucial theoretical assumption here is that “a distinction in form alone eventually, if it is maintained, becomes a distinction in form *and* meaning” (italics is used in the original text, cf. Andrews 1984: 485).⁷³ In addition, what arguably drives this insightful study is the structuralist assumption that a human language is a system of oppositions. Therefore the focus is drawn to contrastive uses of O- and OB- in minimal verbal pairs like *odelit* ‘endow’ – *obdelit* ‘do someone out of his fair share’ and *osudit* ‘condemn’ – *obsudit* ‘discuss’. Andrews adopts the approach articulated by Roman Jakobson and aims her analysis at describing the general invariant meaning of O- and OB- with regard to markedness/unmarkedness and the semantic difference that would be present in each of their contextual occurrences (ibid: 480). Andrews thoroughly examines a set of sixty minimal pairs. Based on this deliberately selected data, Andrews concludes that the two prefixes represent two distinct morphemes characterized by two different invariant meanings. As a result, the fine semantic nuances of contrastive uses of O- and OB- in verbs are generalized into an abstract definition of two invariants. According to Andrews, both morphemes express the total involvement of an object in the verbal process. In this light, the morpheme OB- refers to “a more concrete involvement”, while the morpheme O- is additionally “marked for the fact that there are ADDITIONAL CONSEQUENCES which extend beyond the process itself” (ibid: 482). Compare the pairs like *obledenet* ‘become covered with ice’ – *oledenet* ‘become frozen all through, both externally and internally’; *obsudit* ‘discuss, talk around, express judgements about smth’ – *osudit* ‘condemn, change one’s status by judgement’; *občistit* ‘peel, clean the surface,

⁷² Note that the number of such perfective verbs prefixed in O-, OB-, and OBO- is very large. In Modern Russian the three prefixes taken together form 226 aspectual pair correlates (Natural Perfectives) according to the Exploring Emptiness database developed at the University of Tromsø (Endresen et al. 2012: 243). Therefore, among nineteen Russian prefixes, O-, OB-, and OBO- constitute the fourth most frequently attested aspectual marker in Modern Russian, after PO-, S-, and ZA-. In Old Church Slavonic these prefixes dominated even more: taken together, O- and OB- are attested in 191 perfectives, overriding any other prefix including the highly frequent PO- (176 verbs), U- (163 verbs), and S’b- (146 verbs), according to data from Słoński (1937: 66-100). On aspectual categories in OCS, see Eckhoff & Janda 2014.

⁷³ The importance of this idea can hardly be overestimated. See Chapter 2 for discussion.

rob all over’ – *očistit’* ‘purify (e.g. the soul), change an object externally or internally’. Andrews convincingly argues that in these pairs the morpheme OB- can only refer to the outer and often detrimental modification of the surface, while the counterparts prefixed in O- imply the internal change of an object. Chronologically, Andrews attributes the development of O- and OB- as distinct morphemes to “at least the time of Puškin” (first thirty years of the 19th c.) claiming that “their distinction expanded over the years” (ibid: 487).

It is noteworthy to mention that a similar point about the contrast between O- and OB- is made in two Russian accounts – the Academy Grammar of the Russian Language (Vinogradov et al. 1952: § 911) and a study by Aver’janova (1974: 31). Both accounts concern the higher degree of abstractness and the figurative (metaphorical) nature of the prefix O- as opposed to more spatial, concrete, often colloquial and even substandard (vernacular) parallel uses of OB-. Again, the difference between O- and OB- is examined in contextual uses that involve immediate nominal arguments: *osypat’ laskami, uprekami, podarkami* ‘lit. strew someone all over with endearment, reproaches, gifts’ vs. *obsypat’ pudroj, peskom* ‘strew something all over with powder, sand’; *obežat’ (*obbežat’) glazami prisuststvujuščix* ‘look around the attendees, lit. run with eyes around the attendees’ vs. *obežat’/obbežat’ vokrug doma* ‘run around the house’. However, the establishment of these differences between O- and OB- does not motivate these authors to posit that the two prefixes belong to two separate morphological units.

Krongauz (1998: 131-148) contributes to the Split Hypothesis by elaborating its phonological part. According to his version of the split, the semantic and phonological factors which condition the choice between the two morphemes interact in the following way. Semantics plays the major role: OB- is specialized for spatial semantics (*exat’* ‘drive’ – *ob’exat’* ‘drive around’), while O- refers to a qualitative change of an object (*žestokij* ‘cruel’ – *o-žestočit’* ‘make someone cruel’). Within these two semantic domains, the choice of the prefix is further determined by phonology. According to Krongauz, both morphemes share the same set of three allomorphs – O-, OB-, and OBO-. However, these allomorphs are organized in different hierarchies, and the principles of their distribution are different too (ibid: 141-143). I summarize this account visually in Table 1. This table presents both the semantic distinction between the two morphemes and the phonological distribution of the allomorphs⁷⁴.

The choice of the allomorph depends on the initial phonemes of the base that a prefix attaches to. For both morphemes, the phonological contexts differ in terms of how problematic they are for the realization of the default allomorph.

The default (the least contextually determined) allomorph of the spatial morpheme OB- is *ob-*. OB- is realized as *obo-* if the root’s onset consonant cluster is not compatible with the coda consonant *b* of the prefix (*obo-žrat’sja* ‘overeat’, *obo-gnut’* ‘bend around’)⁷⁵ or if the root contains the vowel/zero alternation described in Chapter

⁷⁴ Note that CC stands for an initial consonant cluster of the simplex base (root morpheme), V is the initial vowel phoneme of the base, and ɤ is the position that refers to the vowel/zero alternation in the root: e.g. *obžeč’*_{INF} – *obožgu*_{1.SG} ‘burn’. Note that in Table 1 and its description I use capital letters to indicate the distinct morphemes O- and OB- and small letters to refer to the allomorphs *o-*, *ob-*, and *obo-*.

⁷⁵ These two examples provided by Krongauz are problematic. I suggest that the presence of OBO- in *obo-žrat’sja* and *obo-gnut’* rather should be explained by the vowel/zero alternation in the root morpheme: compare the corresponding imperfective forms *ob-žirat’sja* and *o-gibat’*. The cluster *bgn* claimed by Krongauz to be ill-formed is nevertheless attested in the verbs *obgnit’* ‘decompose’ and *obgnobit’* ‘insult someone a lot’ which potentially undermine

3 (*obo-drat'*_{INF} – *ob-deru*_{1.SG} ‘tear off all over’). Also, *ob-* is replaced by the allomorph *o-* if the root’s initial consonant is the labial *b* or *p*.

Single morpheme *OB-	
Morpheme 1 OB-	Morpheme 2 O-
Meaning	
Spatial: ‘move around’	Factitive / Inchoative: ‘change an object’
Example	
<i>ob''exat'</i> ‘drive around’ < <i>exat'</i> ‘drive.VERB’	<i>o-žestočit'</i> ‘make X cruel’ < <i>žestokij</i> ‘cruel.ADJ’
Hierarchy of allomorphs	
<i>ob-</i> >> <i>obo-</i> >> <i>o-</i>	<i>o-</i> >> <i>ob-</i> >> <i>obo-</i>
Allomorph selection rules:	
/OB-/ → [ob] / always except: /OB-/ → [obo] / _ CC, if *bCC /OB-/ → [obo] / _ C $\bar{\tau}$ C, where $\bar{\tau}$ → \emptyset /OB-/ → [o] / _ b, p	/O-/ → [o] / always except: /O-/ → [ob] / _ V, v, j, l, n, r, m /O-/ → [obo] / _ CC, if *bCC /O-/ → [obo] / _ C $\bar{\tau}$ C, where $\bar{\tau}$ → \emptyset

Table 1: Hypothesis of Morphological Split: The morphemes O- and OB- and their allomorphs.

The default allomorph of the factitive morpheme O- is *o-*. It always appears in the output unless the root starts in a vowel (*o-žestočit'* ‘make someone cruel’ but *ob-ostrit'* ‘sharpen’). The root initial sonorants *j*, *l*, *n*, *r*, *m* and the labiodental *v* normally trigger the allomorph *ob-*, but not always: e.g. *o-jagnit'sja* ‘give birth to a lamb’, *o-licetvorit'* ‘animate’, *o-nemet'* ‘become numb’, *o-rosit'* ‘irrigate’, *o-molodit'* ‘rejuvenate’, *o-veščestvit'* ‘materialize’. If one of the consonants listed above belongs to a cluster not compatible with preceding *b* or containing the vowel/zero alternation, theoretically the allomorph *obo-* should occur instead of *o-* or *ob-*. According to Krongauz (1998: 147), the allomorph *obo-* of the morpheme O- can be found in the verb *obo-lgat'* ‘believe, defame someone’ due to the vowel/zero alternation in the root (*lgat'* ‘to lie’ – *lož'* ‘a lie’) and, additionally, the consonant cluster *blg* which is problematic for Russian phonotactics.

5.2.2 Drawbacks of the Split Hypothesis

The Split Hypothesis is elegant and appealing, yet it is not without drawbacks. First of all, one can easily come up with plentiful counterexamples in the Modern Russian lexicon where the choice of the prefix does not follow the predictions of the Hypothesis. For instance, the verbs *okol'cevat'* ‘place a ring around smth’, *okružit'* ‘encircle’, *ogorodit'* ‘fence around’ contain the prefix O- instead of the expected OB-, despite the spatial

Krongauz’s argument. Furthermore, in examples *obokrast'* ‘rob’ and *oboznat'sja* ‘take someone for someone else’ the prefix OBO- cannot be explained by incompatibility of *b* with *kr* and *zn* because of attested counterexamples like *obkradyvat'* ‘rob’ and *obznakomit'* ‘introduce to many people’.

semantics of circular movement and the fact that the phonological context is clearly unproblematic for OB- (the root initial *k*, *kr*, *g*).

Moreover, the Split Hypothesis does not predict minimal pairs where the verbs prefixed in O- and OB- are close synonyms, like *ob-kleit'* and *o-kleit'* 'glue all over', *ob-strič'* and *o-strič'* 'cut off', *ob-strogat'* and *o-strogat'* 'plane from all sides', *ob-kopat'* and *o-kopat'* 'dig around'.

Note that in some verbs neither the spatial semantics 'move around' nor the factitive / inchoative semantics of a qualitative change is easily detectable: *obdelat'* (*dela*) 'manage, succeed' (< *delat'* 'do'), *obmerit'* 'cheat in measuring' (< *merit'* 'measure'), *obmanut'* 'deceive' (deetymologized), *opolzti'* 'slip' (< *polzti'* 'crawl').

Overall, the rich polysemy of the prefixes O-, OB-, and OBO- is rather simplified in the Split Hypothesis. The wide scope of various submeanings that these prefixes can express is reduced to the opposition of circular movement and qualitative change, both of which are quite broadly understood. The selective approach to data and the focus on contrastive uses thus avoids a detailed semantic analysis of the entire body of relevant vocabulary. The Split Hypothesis should be tested and verified against a comprehensive list of prefixed verbs.

Last but not least, one cannot but notice contradictory statements that undermine the argument. On the one hand, the Split Hypothesis suggests that the split has already taken place so that its results can be observed in Modern Russian (Krongauz 1998: 147). On the other hand, the split is described as an on-going process which is not yet completed (Krongauz 1998: 149), and in which the two morphemes still preserve a strong connection (Krongauz 1998: 138). This stipulation leaves room for some counterexamples in the lexicon: it is claimed that some verbs (without further clarification) preserve their prefix from the time before the split, whereas other verbs might be not affected by the split yet. This explanation lacks specific examples and undermines the argument.

I test the Split Hypothesis against corpus and experimental data without any theoretical bias, adopting an agnostic view on the status of O- and OB-. I abandon the selective approach to data and instead examine the most exhaustive dataset available. The objective of this study is to provide an account that can capture both the diversity of the data as well as its significant patterns and profiles.

5.2.3 Predictions of the Split Hypothesis

Verification of the idea about distinct morphological identities of the prefixes O- and OB- requires that we approach it as a linguistic hypothesis. In order to test it, we need to formulate its claims in terms of predictions that are empirically testable.

The Split Hypothesis has two parts: diachronic and synchronic. In both case studies, I am interested only in the synchronic part: do the prefixes O- and OB- represent two distinct morphemes in Contemporary Standard Russian?

According to the Split Hypothesis, the central mechanism that determines the distribution of O- and OB- in existing and possible verbs of Modern Russian is the proposed distinction: spatial vs. factitive/inchoative semantics and the corresponding deverbal vs. deadjectival/denominal patterns of derivation. It is expected that all kinds of transitional cases or remnants inherited from the time before the split should constitute a much smaller part of data than the verbs affected by the split.

Also, it is expected that the meaning 'move around' is not related to the meaning 'qualitative change of an object'. These two meanings are assumed to be so different that

they cannot be described within the polysemy of a single morpheme. They can only belong to two separate prefixal morphemes.

Taking into account the hierarchical allomorphic distribution of the two morphemes proposed by Krongauz (recall Table 1), one can distinguish between three relevant phonological conditions:

- **Stem-initial vowels** allow only the prefix OB-, therefore the choice between prefixes here is purely phonological, while the impact of the semantic distinction is neutralized;
- **Stem-initial obstruents (except *p, b, v*)** allow both O- and OB-, and the choice between them depends entirely on the target semantics;
- **Stem-initial sonorants and *p, b, v*** lie between the first two groups: they allow variation, the semantic and the phonological factors compete, and the choice between O- and OB- depends on the impact of the winner. By means of example, consider the verb *o(b)-bežat* 'run around smth', which allows variation in the shape of the prefix even within the same context *vokrug doma* 'around the house'. In *o-bežat*, the phonological factor presumably overrides semantics, because the surface shape of the prefix demonstrates that it is more important to avoid the geminate *bb* than to preserve *ob-* as a default shape bearing spatial meaning. By contrast, in *ob-bežat*, it is arguably the semantic factor that drives the choice of the prefix, while the phonotactic requirement to avoid the geminate *bb* is less important.

To conclude, the phonological contexts most diagnostic for testing the impact of the semantic distinction are the contexts described in condition (2): they are neutral and unproblematic for both O- and OB-. According to the Split Hypothesis, in front of stem-initial obstruents other than *p, b, v*, the prefixes O- and OB- should be distributed complementarily according to the proposed semantic distinction. Under this neutral phonological condition we expect the default allomorphs of the two morphemes, and the choice between them should be determined exclusively by target semantics. Exactly these contexts are the key focus of the experimental study presented in section 5.4.

Note that in works that propose the Split Hypothesis the allomorph *obo-* is usually assumed to belong exclusively to the morpheme OB- (except Krongauz (1998: 147) whose arguments are disproved by Itkin (2007: 85)). Indeed, *obo-* can be explained in terms of Russian phonotactics and the vowel/zero alternation typical for consonant-final prefixes (cf. Chapter 3). As an allomorph of the morpheme O-, *obo-* becomes problematic because there is no active phonological rule that explains why in some contexts *o-* turns into *obo-*. This is taken into account in both case studies.

In the next section I test the Split Hypothesis against the Modern Russian lexicon, while in 5.4 I test it experimentally by eliciting speakers' active mechanisms of word production.

5.3 Corpus study

5.3.1 Goal and justification for a new semantic model

The goal of my corpus study is to test the Split Hypothesis against the most extensive dataset possible. Why was it necessary to carry out a new semantic analysis of verbs

prefixed in O-, OB-, and OBO- after all the work done in previous studies (Roberts 1976, Krongauz 1998, Košlev 2004b)? How is my analysis different and novel?

First, earlier studies of lexical data on the three prefixes in question were limited to verbs attested in dictionaries. I provide the first corpus-based study of the issue.

Second, in order to compare the three prefixes I need a rather fine-grained classification of prefix submeanings as well as the analysis of their interconnections that would justify assigning them to a single polysemous network. The crucial questions for my analysis are: Is it possible to trace a connection between the spatial and factitive meanings or are they unrelated? Can each of the two prefixes express both submeanings or not? How are other submeanings of these prefixes motivated and distributed across verbs?

The question concerning the relationship of different submeanings of O- and OB- remains open. Roberts 1976 offers a detailed account of submeanings for all three prefixes. However, he draws distinctions between submeanings that belong to prefixed verbs rather than to the prefixes alone. Consider, for instance, the meaning B^{vt3} 'Fear, caution' proposed for the prefix in verbs *opasat'sja* 'fear' and *osteregat'sja* 'beware of, avoid', or the meaning B^{s10} 'Birth of animals' for the prefix in the verb *oporosit'sja* 'farrow' (Roberts 1976: 69). Nevertheless, Roberts' account shows that each of the submeanings that he distinguishes is attested for each of the three prefixes. However, Roberts does not explicate how "spatial", "modal", and "affective" subgroups of meanings are related to each other, though this could become a key argument in favor of a single model of polysemy and a single-morpheme-account.

Similarly, Košlev 2004b identifies three abstract conceptual meanings, each of which can be expressed by all the three prefixes. However, this study does not discuss how these conceptual meanings relate to each other. Krongauz 1998 refers to the semantic account provided in dictionaries and Academy grammars which list all submeanings without any mention of their semantic links.

Contrary to previous accounts, my model of prefix polysemy aims to reveal and motivate possible semantic links between the submeanings of O- and OB-.

What makes my analysis different is that I propose a semantic model that makes use of the notions and terminology of Cognitive Linguistics. I suggest that a cognitive account makes it possible to look at the semantics of O- and OB- from a different perspective. Instead of narrowing a rich network of polysemy down to an abstract semantic invariant or an incoherent list of submeanings, a cognitive approach⁷⁶ models polysemy as a radial network of related semantic subcategories. The subcategories are organized hierarchically around a central prototypical meaning that is based on basic human bodily experience of space and movement. The prototypical meaning motivates most of the other submeanings of the network by means of cognitive mechanisms of

⁷⁶ It should be mentioned that the approach to semantics taken by Cognitive Linguistics is in many aspects close to that of the Moscow semantic school described in Apresjan 1995, 2005, 2009. The similarity between the two frameworks has been recognized earlier in Raxilina 1998: 278 and Janda 2010: 4. First of all, both schools treat semantics seriously and claim that meaning plays a major role in human languages. Both schools value comprehensive datasets and the idea of structural organization of semantics and lexicon. Moreover, both approaches share interest in polysemy and hierarchical organization of linguistic meaning. Yet their key notions and metalanguages are different. On the other hand, a *radial category*, which is a cornerstone of Cognitive Linguistics, can be compared to the *polevaja struktura* 'field-type structure' in studies of Russian grammatical categories in the Functional Grammar School of A. Bondarko (Bondarko et al. 1996).

metaphor and metonymy. Crucially, a cognitive model of prefix semantics makes it possible not only to posit cognitive motivations for the various submeanings, but also to compare polysemous units within a single radial semantic network. In this study, I also adopt the Radial Category Profiling methodology proposed in Nessel et al. 2011. I employ this methodology in order to test and measure the degree of semantic similarity and divergence of O- and OB- with regard to the fifteen submeanings that I recognize and observe in each of them.

5.3.2 Data collection

I examined all perfective verbs prefixed in O-, OB-, and OBO- and attested in two authoritative sources of Contemporary Standard Russian – the Russian National Corpus and the Grammatical Dictionary of the Russian Language (Zaliznjak 1980). In order to account for their semantics, I consulted five dictionaries: Ožegov & Švedova 2001, Efremova 2000, Ušakov 2008, Evgen'eva 1999, and Kuznecov 2000. The verbs were culled from the database of the frequency dictionary Lyashevskaya & Šaroff 2009, based on the Modern Subcorpus of the RNC, which includes authentic Russian texts created in 1950-2007 (=92 million words).

In structuring the data I applied the following methodology. First, in order to avoid duplication of data, reflexive verbs were merged with their non-reflexive counterparts if the postfix -SJA marks intransitivity and does not alter the lexical meaning of the base verb. This strategy is different from the one used in Krongauz 1998, where reflexives and non-reflexives are counted as separate verbal lexemes. In my database, verbs like *obespokoit'* 'disturb' and *obespokoit'sja* 'become disturbed' are represented as a single verbal lexeme *obespokoit'(sja)*. In case the postfix -SJA alters the lexical meaning of the base verb, the reflexives and non-reflexives are represented in the database as different lexemes (e.g. *ogovorit'* 'make a stipulation; slander' and *ogovorit'sja* 'make a mistake in speaking').

I interpret verbs like *obezumit'* 'deprive of mind' as derivatives formed from a prepositional phrase *bez uma* 'without mind' and the prefix O-, following Townsend (1968: 143) and contrary to Pacjukova (2009) who postulates a combined prefix OBEZ-.

The number of entries occupied by a polysemous verb in my database corresponds not to the number of its lexical submeanings recognized in the dictionaries but to the number of subcategories that its prefix represents according to the model of polysemy proposed in Figure 1.

The database has 1,037 verbal lexemes and is available in Excel spreadsheet format at <http://hdl.handle.net/10037.1/10078>.

5.3.3 The cognitive model of polysemy:

Radial network of submeanings and their interconnections

The semantic analysis of verbs prefixed in O-, OB-, and OBO- shows that all three prefixes are highly polysemous. I propose a semantic model given in Figure 1 that accounts for all the submeanings. This model is a radial network which includes fifteen related submeanings organized around a spatial prototype. Most importantly, it shows that the complex semantics of all three prefixes can be described in a single model. This model incorporates all submeanings into one network, bringing together spatial and non-spatial uses of prefixes and showing how they are interrelated.

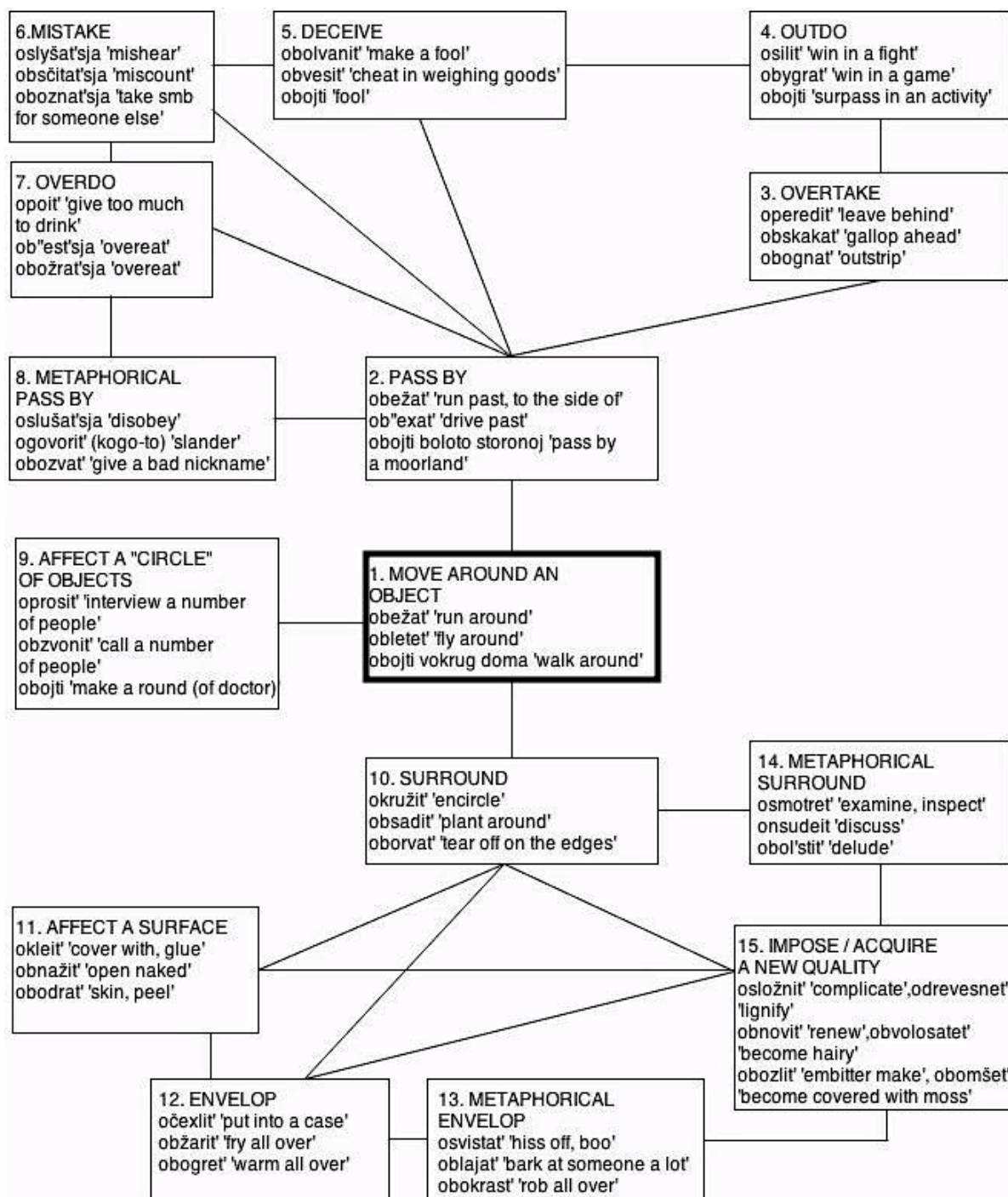


Figure 1: Radial category network of submeanings expressed by O-, OB-, and OBO-.

As I demonstrate in Figure 1, each of the submeanings distinguished in this model is attested for all three prefixes. This is a good piece of evidence for the close semantic connection between O-, OB-, and OBO-. In support of this claim I provide illustrative examples both in the model and in the text below. Note that this model of polysemy represents a network of submeanings of the prefixes rather than a semantic classification of prefixed verbs. Therefore, each submeaning corresponds to a different semantic contribution of a prefix to a simplex stem⁷⁷.

⁷⁷ In order to identify the semantic contribution of a prefix I apply a standard procedure, comparing the meaning of the prefixed verb with the meaning of its simplex base: e.g. *obletet* 'fly

In the center of Figure 1 there is the prototypical meaning shared by O-, OB- and OBO-: 1.MOVE AROUND AN OBJECT. This submeaning of the prefix can be observed in verbs that denote different kinds of movement in space: *obletet'* 'fly around something', *o(b)bežat'* 'run around something', *obožiti (vokrug doma)* 'walk (around the house)' that are formed from unidirectional motion verbs *letet'* 'fly', *bežat'* 'run', and *idti* 'walk'. This submeaning corresponds to image schema 1 (see Figure 2) which depicts a mobile Trajector that moves around a stable Landmark so that its trajectory represents a prototypical full circle.

The submeaning 1.MOVE AROUND AN OBJECT has a special status in this network because it serves to motivate all the remaining submeanings, directly or indirectly. The submeanings are depicted as rectangles. The lines between the rectangles indicate semantic links between the submeanings. There are two kinds of relations: 1) submeaning A can be a source and motivation for submeaning B (e.g. 2.PASS BY motivates 3.OVERTAKE, and 10.SURROUND motivates 14.METAPHORICAL SURROUND); or 2) there are verbs that simultaneously represent the submeanings A and B of the prefix (e.g. 11.AFFECT A SURFACE and 15.IMPOSE/ACQUIRE A NEW QUALITY are related because in the verbs *opalubit'* 'cover with a deck', *operit'(sja)* 'cover with feathers', and *oblyset'* 'grow bald' the prefix expresses both submeanings). The extensions of the prototypical meaning come in two types according to two cognitive mechanisms – metaphor and metonymy. Metonymic extensions occur when an image schema is reduced to its part or is re-interpreted with a different focus. Metaphorical extensions are those that apply an image schema to a different domain. This accounts for the shift from the domain of concrete spatial relations to the more abstract conceptual domains of time, human relations, emotions, qualities, etc.

I first outline how the model in Figure 1 is structured, and then present each submeaning in turn in more detail.

Although the graphic representation is a matter of convention, this radial network of meanings can be subdivided into two parts: the top part includes submeanings 2-8, whereas the bottom part includes submeanings 10-15. The submeanings are grouped in the top and the bottom parts because they interpret the spatial prototype (the circular movement) in two different ways, as I explain below. In a sense, submeaning 9 does not belong to either of these subgroups of meanings, because it implies a complex multiple Landmark.

Submeanings 2-8 focus on keeping distance from the Landmark: consider 2.PASS BY where the Landmark is an obstacle that remains outside of Trajector's path (*obožiti boloto storonoj* 'pass by a moorland'). 3.OVERTAKE (*operedit'* 'leave behind') and 4.OUTDO (*obygrat'* 'win in a game') refer to outscoring a competitor (Landmark). Furthermore, leaving out a Landmark beside one's path is metaphorically transferred into avoiding or missing a crucial point (a Landmark) in a situation. Making a mistake, accidentally (5.MISTAKE: *ogovorit'sja* 'mispronounce') or deliberately (6.DECEIVE: *obsčitat'* 'short-change'), also involves passing by the proper performance. Similarly, overdoing something (7.OVERDO: *ob''est'sja* 'overeat') involves passing by the proper limit of an activity.

Submeanings 10-15 take up the opposite facet of the prototype: when moving around the Landmark, the Trajector can enclose it, contacting and affecting its boundaries (10.SURROUND: *okružiti'* 'encircle'), surface (11.AFFECT A SURFACE: *okleit'* 'cover

around' < *letet'* 'fly' + OB- 'MOVE AROUND AN OBJECT'; *obzvonit'* 'give a call to a number of people' < *zvonit'* 'call' + OB- 'AFFECT A "CIRCLE" OF OBJECTS'.

with', *obodrat* 'skin, peel'), or a three-dimensional volume (12.ENVELOP: *obžarit* 'fry on both sides, all over', *očexlit* 'put into a case', *obovšivet* 'become lice-ridden'). Moreover, one can surround and envelop an object not only spatially but also metaphorically with an activity (14.METAPHORICAL SURROUND: *obsudit* 'discuss', *opisat* 'describe', *obol'stit* 'delude'; 13.METAPHORICAL ENVELOP: *oblajat* 'bark at someone a lot', *osvistat* 'hiss off, boo', *obokrast* 'rob all over'), and, further, with a new quality – concrete (15.IMPOSE/ACQUIRE A NEW QUALITY: *okruglit* 'round', *obramit* 'frame', *obliscevat* 'face', *obnažit* 'nude') or abstract (*opečalit* 'sadden', *obvinit* 'accuse'). Thus, affecting and changing the Landmark shifts from the spatial domain (*obledenit* 'cover with ice') to the domain of abstract characteristics (*obednit* 'impoverish').

Now I turn to how all these submeanings are systematically related.

Submeaning 2.PASS BY is a metonymical extension of the the prototype 1.MOVE AROUND AN OBJECT. The full-circle trajectory is reduced here to a semi-circle, as shown in image schema 2 (Figure 2) or can even be completely neutralized as in image schema 3 (Figure 2). The Trajector bypasses the Landmark by leaving it outside its path. So, the key idea in both cases is the Trajector's spatial disjunction from the Landmark, in other words leaving the Landmark outside the trajectory, and in a sense, missing the "target". The submeaning 2.PASS BY should be assigned to the prefix in most of the same verbs of spatial motion that I listed for the prototypical meaning, if they are put in a special context: e.g. *ob"exat* (*boloto, opasnoe mesto storonoj*) 'bypass a moorland, an unsafe place', *obletet* 'bypass flying', *obnesti* 'leave out in serving'. Independently of context the submeaning 2.PASS BY can be found in the prefix of the verbs *obognut* 'bypass' and *obteč* 'flow past'. This submeaning plays an important role in the semantic network because the reduction of the full-circle trajectory motivates the remaining submeanings in the upper part of the model (Figure 1).

Note that the configuration of 2.PASS BY implies that the Trajector moves, while the Landmark remains stable.

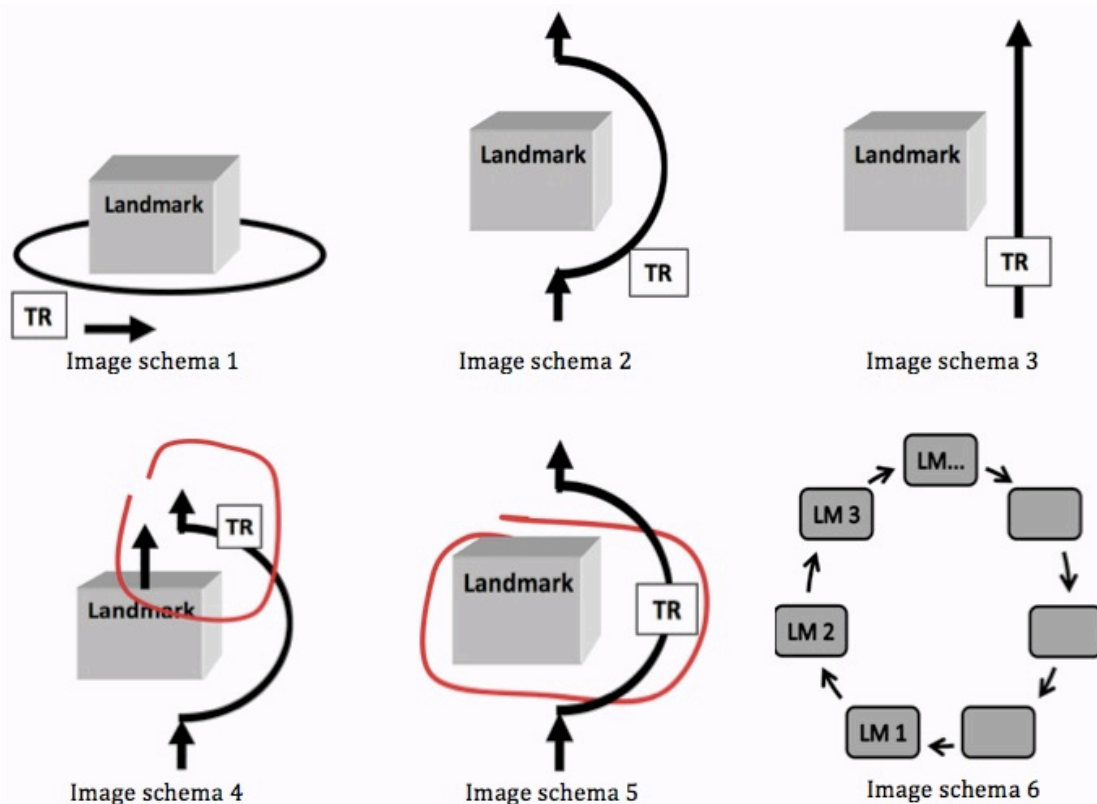


Figure 2: Image schemas that motivate the submeanings of O-, OB-, and OBO-.

Submeaning 3.OVERTAKE resembles image schema 2, but refers to a situation where both the Trajector and the Landmark are moving objects. Compare the verbs *obognat* ‘outstrip’, *obskakat* ‘gallop ahead’, *operedit* ‘leave behind’ and *obožiti* ‘outstrip’. As an illustration consider the following corpus example:

- (3) – *Ee [Isèk-Kyrgan] ne mogla obožiti na skačkax ni odna lošad’ v stepi.* [D.N. Mamin-Sibirjak. Ak-Bozat (1895)].

‘No other horse on the steppe could **overtake** her [Isèk-Kyrgan – name of a horse] in a horse-race.’

Submeaning 3.OVERTAKE brings the focus to that part of image schema 2 that corresponds to outdistancing (or overtaking) the Trajector. Image schema 4 shows how this configuration is reinterpreted by means of zooming in to a specific part, the *construal* (Langacker 1999: 206) that this submeaning entails.

Submeaning 4.OUTDO represents a metaphoric extension of the spatial submeaning 3.OVERTAKE into the domain of human activities, which involve various kinds of competition: *obygrat*’ and *obščelkat*’ both glossed ‘win over someone in a game’, *oblovit*’ ‘outdo others in catching fish’, *oborot*’ ‘overcome, fight down’ (compare *neoborimoe želanie* ‘invincible desire’), and *osilit*’ ‘win in a fight’.

By contrast with image schema 4, the image schema 5 (Figure 2) focuses on the “avoided” part of the configuration marked in red, where the Trajector abandons the original route and takes a deviating semi-circular path. The idea of missing the target (original Landmark) motivates several submeanings (5-8) at the same time.

Submeaning 5.DECEIVE refers to deviations made on purpose for personal gain. This meaning of the prefix can be found in many verbs like *obsčitat*’ (*kogo-to*) ‘short-change (someone)’, *obdelit*’ ‘leave one out of one’s fair share’, *obxitrit*’ ‘deceive, trick’, *obvesit*’ ‘cheat in weighing goods’, *obmerit*’ ‘cheat in measuring’, *ob’egorit*’, *o(b)lapošit*’, *obdurit*’ all denoting ‘swindle’. The umbrella verb in this list is the deetymologized and synchronically opaque *obmanut*’ ‘deceive’ (even though the base of this verb is unclear, the prefix is still analyzable).

Submeaning 6.MISTAKE includes careless accidental blunders: *oslyšat’sja* ‘mishear’, *obsčitat’sja* ‘miscalculate’, *ogovorit’sja* ‘mispronounce’, *oboznat’sja* ‘take someone for someone else’, *opečatat’sja* ‘misprint’, *opisat’sja* ‘make a mistake in writing’, *oseč’sja* ‘miss the target in shooting’, *ošibit’sja* ‘make a mistake’, and *oprostofilit’sja* ‘goof’.

Submeaning 7.OVERDO implies a mistake made to the detriment of oneself by exceeding the proper limit to an activity. This meaning employs the prefixes O-, OB-, and OBO- in combination with the postfix -SJA. The pattern of this morphological construction is very productive and can be found in many verbs that denote basic human activities like eating and drinking: *ob’est’sja*, *obožrat’sja*, *oblopat’sja*, *obtreskat’sja* all glossed as ‘overeat’ as well as *o(b)pit’sja* ‘drink too much’. Moreover, the same pattern with the same semantics can be observed in verbs referring to other types of human activities: e.g. *obkurit’sja* ‘smoke too much, get over-intoxicated’, *obxoxotat’sja* and *oboržat’sja* ‘laugh too much’, *obyskat’sja* ‘spend too much time searching in vain’, *oblenit’sja* ‘become too lazy’, *obuzit*’ ‘make too narrow’, and *obkornat*’ ‘cut too short and uneven’. One can form such a verb from practically any verbal simplex that denotes an activity, as evidenced by numerous marginal verbs attested in the corpus: e.g. *občitat’sja* ‘read too much’, *obrevet’sja* ‘cry too much’, *obšeptat’sja* ‘whisper too much’, *obmečtat’sja* ‘spend too much time dreaming’, and *obvorovat’sja* ‘do too much robbing’.

Submeaning 8.METAPHORICAL PASS BY of the prefix can be observed in verbs that describe metaphorical deviations from the Landmark. Several verbs denote ‘slander, defame’: *ogovorit’ kogo-to, obolgat’, oklevetat’*. Similarly, *obozvat’* refers to giving someone a bad nickname instead of using a proper name. The verb *oslušat’sja* conceptualizes disobedience in terms of deviation from the proper expected behaviour. In *oboždat’* ‘wait for a while’, the Landmark is a short period of time that one has to spend waiting for someone. While waiting, one “bypasses” this time slot temporally without experiencing it otherwise. This interpretation explains the use of the prefix in this verb as bearing the meaning 8.METAPHORICAL PASS BY.

Submeaning 9.AFFECT A “CIRCLE” OF OBJECTS is directly motivated by the prototype (Figure 1) and is based on image schema 6 (Figure 2). This submeaning implies multiple Landmarks that become subjected to an activity⁷⁸. All of them have the same status and are conceptualized as a single set, or a “circle”. The Trajector moves along this imaginary circular trajectory from one Landmark to another and affects each of them in turn. However, the prefixed verbs of this subcategory do not single out individual Landmarks but rather refer to them taken together as a whole set, or a “circle”. This submeaning is very productive and is well-described in the literature (cf. “*total’noe dejstvie*” ‘total action’ in Košelev 2004b: 68). The prefix can contribute this meaning to verbs of various kinds: unidirectional motion base verbs that form *oboiti* ‘make a round (of doctor, sentry)’, *obletet’* ‘flying visit many places’, *obnesti (vsax)* ‘serve round to all or several people’, as well as non-directional motion base verbs that motivate the derivatives *obbégat’* ‘running visit many places’, *oblazit’* ‘climb everywhere’, and *ob’ezdit’* ‘traveling visit many places’. In addition, the prefix can carry this semantics in verbs denoting activities other than motion through space: e.g. *oprosit’* ‘interview a number of people’, *obstirat’* ‘wash clothes for many people or all the clothes for one person’, *obštopat’* ‘darn all the clothes or clothes of many people’, *obzvonit’* ‘give a call to a number of people’, and *obzanimat’* ‘borrow from many people’.

The remaining submeanings are located in the bottom part of the model (Figure 1) and represent image schema 1 (Figure 2). These submeanings are motivated by the prototypical circular motion, understood through the medium of submeaning 10.SURROUND, which introduces the idea of close contact with the Landmark – contact that affects its qualities.

Submeaning 10.SURROUND is spatial. Its configuration implies that the Trajector surrounds the Landmark and affects its edges. The focus is on the outer spatial boundaries of the Landmark, where something can be attached, as in (4), or detached and removed, as in (5):

- (4) *obsadit’ dom cvetami* ‘plant flowers around the house’; *obvjazat’* ‘tie around; edge in chain-stitch (knitting)’; *o(b)gorodit’* ‘fence around’; *obmurovat’* ‘encircle with a stone wall’; *obstroit’* ‘surround with buildings or parts of a building’; *obložit’* ‘put around’; *obstavit’* ‘surround; furnish’; *obvit’* ‘wind around, entwine’; *obšit’* ‘edge, border around’; *obmotat’* ‘wind around’;

⁷⁸ Multiplication of the Landmark is a phenomenon found in the semantics of other Russian prefixes as well, for example in distributive use of PERE-: *perečitat’ (mnogo knig)* ‘read through (many books)’ (cf. Janda 1986: 161).

- (5) *obkusat* 'bite all over', *obglodat kusok* 'gnaw all over', *o(b)kopat* 'dig around', *obryt* 'dig around', *ob"est po krajam* 'eat at the edges', *obkrošit* 'crumble, break off into pieces from the outside, on the edges', *oblomat kraja* 'break off (edges)'.

The directionality of the impact thus depends on the lexical meaning of the simplex verbal base. Interestingly, some verbs can imply impact in both directions: *ogresti* 'rake around, towards the object or away from it' (Ožegov & Švedova 2001).

It is important that submeaning 10.SURROUND is relevant for a large group of verbs that are doubly motivated by both a verbal and a nominal/adjectival base: *opojasat* 'gird', *okružiti* 'encircle', *ograničiti* 'surround with borders, limit', *obnajtoviti* 'tie around with a rope', *obmeževati* 'surround with boundaries', *o(b)čertiti* 'draw around', *obuzdat* 'put a bridle on', *okol'cevat* 'ring', *okantovat* 'mount all around'. I argue that the prefix in these verbs simultaneously bears the spatial meaning 10.SURROUND and the factitive meaning 15.IMPOSE A NEW QUALITY.

A number of similar verbs like *okajmiti* 'decorate with edging', *obramiti* 'put into a frame', *ocepiti* 'surround, cordon off', and *okučiti* 'make a pile of soil around a plant' are associated with only a nominal base (*kajma* 'edging', *rama* 'frame', *cep* 'chain', *kuča* 'pile'). I suggest that despite the lack of any verbal base (**kajmiti*, **ramiti*, **cepiti*, **kučiti*), the spatial meaning 10.SURROUND of the prefix in these verbs is nevertheless undeniable.

Whereas submeaning 10.SURROUND of the prefix refers to impact on the outer edges of an object, submeaning 11.AFFECT A SURFACE enlarges the area of impact to the entire surface of an object, while submeaning 12.ENVELOP goes even further and implies a three-dimensional impact on the entire body of an object. In some verbs the meaning of the prefix can be interpreted in terms of both submeaning 11 and submeaning 12, depending on the immediate context of the verb. For instance, the verbs *obryzgat* 'splash, besprinkle' and *ostrogati* 'plane' can imply affecting only one surface of an object or affecting it all over from all sides. However, although such ambiguous cases exist, it is possible to distinguish between submeanings 11 and 12, assigning many verbs to the former meaning of the prefix or the latter. I suggest that the prefix carries submeaning 11.AFFECT A SURFACE in *okleit* 'cover with, glue', *oplevat* 'spit on the surface all over', *ozerkaliti* 'cover with mirrors', *okrasiti* 'paint', *okropiti (svjatoj vodoj)* 'spray (with holy water)', *okroviti* 'cover with blood', *ošpariti* 'scald by pouring hot water on', *oteret* 'rub all over the surface', *obrit* 'shave off', *oškuriti* 'rub a wooden surface with sandpaper to make it smooth', *ošlifovati* 'polish the surface', *otesati* 'hew the surface (wood, stone)', *obodrat* 'tear off skin or cover, peel', *obtoptati* 'trample all over the place', *obšelušiti* 'make flake off', *ogravirovati* 'engrave', *ocarapat* 'scratch off', and *opaliti* 'singe, burn the surface'. Note that again, as observed in the 10.SURROUND group, some of the verbs in the 11.AFFECT A SURFACE subcategory describe attaching something to a surface (*okleit* 'cover with, glue'), while others imply removal of some substance from the surface (*obrit* 'shave off') depending on the lexical meaning of the base. The verbs *obiti* and *obbit* (< *bit* 'beat') can refer to both covering the surface with a new layer or removing it from the surface, depending not solely on the base verb but on the larger context. The directionality of the impact is not specified in the prefix itself.

Submeaning 12.ENVELOP is assigned to the prefix in verbs denoting:

- eating (*obževati* 'chew all over', *obklevati* 'peck all around', *obsosati* 'suck all over', *oblizati* 'lick all over', *obgryzti* 'gnaw from all sides', *obkusati* 'bite all over');
- cooking (*obžariti* 'fry on both sides, all over', *obvariti* 'pour boiling water over; scald', *obvjaliti* 'jerk all over, from all sides');

- attachment of objects (*obkleit* 'glue all over', *obvešat* 'hang objects and cover with them all over', *oblepit* 'stick all over', *oblipnut* 'stick all over');
- removal of objects (*obščipat* 'pluck all over', *obobrat* 'pick all over; rob');
- actions involving tactile contact with an object (*o(b)ščupat* 'touch all over', *obcelovat* 'kiss all over', *obmusolit* 'beslobber, soil all over');
- affecting an object from all sides (*oblit* 'pour all over', *obmerit* 'measure all directions', *obvaljat* 'roll (in) all over', *obvertet* 'wrap up in something', *oputat* 'entangle all over', *obvoloč* 'envelop, cover all over', *okutat* 'wrap up all over', *objuxat* 'sniff all over', *obogret* 'heat, warm all over', *obrasti* 'grow all over');
- damaging an object from all sides (*obstreljat* 'fire at, bombard, shoot all over', *obxlestat* 'whip from all sides', *obcarapat* 'cover all over with scratches').

Note that submeaning 11.AFFECT A SURFACE and submeaning 12.ENVELOP can often be contrasted in minimal pairs of verbs like *okleit* 'cover with, glue' vs. *obkleit* 'glue all over'; *ocarapat* 'scratch the surface' vs. *obcarapat* 'scratch all over', where the verbs prefixed in O- describe impact made on one surface of an object (that is a *two-dimensional spatial effect*), while their counterparts prefixed in OB- imply an impact on all the surfaces of an object (in other words a *three-dimensional spatial effect*). In these examples the two prefixes oppose each other not in terms of concreteness vs. abstractness or spatial vs. non-spatial domain (because 11.AFFECT A SURFACE and 12.ENVELOP are both spatial and concrete), but in terms of the affected dimensions (surface vs. volume).

There are many verbs that lack a verbal base and are motivated by a noun or an adjective instead. Notably, in such verbs the prefix can likewise express the spatial meaning 11.AFFECT A SURFACE (*opalubit* 'cover with a deck', *obdernit* 'cover an area with turf', *operit'(sja)* 'cover with feathers, plumage', *obomšet* 'get covered with moss', *oblučit* 'irradiate', *opautinit'sja* 'get covered with a spider web', *obnažit* 'open naked', *obvetrit* 'make rough by exposure to wind', *o(b)salit'(sja)* 'cover with fat', *okorit* 'peel the bark') or the spatial meaning 12.ENVELOP (*obmundirovat* 'provide uniforms', *obmoxnatet* 'become very hairy').

There are also intermediate cases – the verbs with double motivations (due to existence of both a verbal and a nominal base) and therefore two submeanings of the prefix – spatial and factitive/inchoative at the same time. As an example, consider the verb *ogrjaznit* 'make dirty', where the prefix has both the submeaning 11.AFFECT A SURFACE (having the base verb *grjaznit* 'make dirty') and the submeaning 15.IMPOSE A NEW QUALITY (being additionally motivated by the adjective *grjaznyj* 'dirty'). The prefixes have this double derivational and semantic motivation in many verbs: *obagrit* 'make crimson', *obgladit* 'make smooth', *openit* 'cover with foam', *osnežit* 'cover with snow', *obmaslit* 'cover with butter', *obelit* 'whitewash, vindicate', *obzelenit* 'make green, cover with trees, flowers', *opušit* 'cover, powder (of hoar-frost or snow)', *odrjablet* 'become flabby', *oserebrit* 'cover with silver', *obvolosatet* 'become hairy, get lots of hair everywhere', *obgnit* 'decompose on the surface', *ogravirovat* 'engrave', *oštukaturit* 'plaster', *oblicevat* 'face', *oblyset* 'grow bald', *o(b)smolit* 'cover or saturate with resin', *oplešivet* 'become bold', *opryščavet* 'become covered with pimples', *oščetinit'(sja)* 'bristle up' (11.AFFECT A SURFACE); and *oblapat* 'paw all over', *ogranit* 'cover with facets', *okruglit* 'round off', *otumanit* 'make foggy, cover with fog', *ocinkovat* 'cover with zink all over' (12.ENVELOP).

Submeaning 13.METAPHORICAL ENVELOP is expressed by the prefix in the verbs *oblaskat* 'treat with tenderness', *obrugat* 'swear a lot at someone or something', *obučit*

'teach', *ograbit'* 'rob', *obvorovat'* 'rob all over', *oblajat'* 'bark at someone a lot', *obmaterit'* 'curse someone a lot', *obxajat'* 'criticize all over', *obxamit'* 'insult by being rude', *obletat' (samolet)* '(tech.) test a plane by flying', and *obkurit' (trubku)* 'adapt, make the smoking device more convenient to smoke with'. These verbs indicate that impact of various kinds of actions can be conceptualized in terms of the spatial subcategory 12.ENVELOP, which "wraps" a person (Landmark) with an activity (Trajector). As a result, the amount of activity is understood in terms of exhaustiveness and surplus, that is why the prefix makes the verbs expressive. The expressive effect is obtained because an activity "envelops" entirely and from all sides: *oblajat'* does not simply imply 'bark at someone', but rather 'bark at someone a lot', and can be literally decomposed into the description 'wrap one into an extensive amount of barking'. Similarly, *obvorovat'* 'rob all over' assumes that all or many valuables are stolen, not just one or two. Example in (6) illustrates how the spatial and metaphorical impact can combine:

(6) *V detstve ja byl i oblejan, i oblizan...* [V. Astaf'ev. Zrjačij posox. (1978-1982)].
'When I was a child, I was **cherished** and **treated with care (lit. licked).**'

Submeaning 14.METAPHORICAL SURROUND is a metaphorical extension of submeaning 10.SURROUND from the spatial domain to the domain of human activities. The impact on the object (Trajector) is conceptualized in terms of spatial enclosure in verbs like *obdumat'* 'think over', *obrisovat'* 'outline, depict', *obgovorit'* 'discuss', *obol'stit'* 'delude', *obsudit'* 'discuss', *opisat'* 'describe', *obygrat' (čto-to)* 'use when creating something', *obžit' (novuju kvartiru)* 'render habitable, assimilate a new place as home', *o(b)smotret'* 'look at something from different sides, look around'.

Verbs with double motivational links are found in this group also: *očarovat'* 'charm, fascinate', *odurmanit'* 'stupefy', *obmyslit'* 'think over', *obmozgovat'* 'think over'. These verbs are associated with both verbal (*očarovat'* 'charm' < *čarovat'* 'charm', *obmyslit'* 'think over' < *myslit'* 'think'), and nominal simplex base (*očarovat'* 'charm' < *čary* 'bewitchery, charms', *obmyslit'* 'think over' < *mysl'* 'thought'). Such verbs serve as interlink that connects the submeaning 14.METAPHORICAL SURROUND with the submeaning 15.IMPOSE/ACQUIRE A NEW QUALITY.

Submeaning 15.IMPOSE/ACQUIRE A NEW QUALITY, as shown above, is expressed by many verbs with double (verbal-nominal) motivations. In addition, this submeaning can be observed in many verbs that have only one, non-verbal, simplex base. Depending on derivational structure and semantics, the following subtypes can be distinguished:

- MAKE X BE Y deadjectival factitive verbs:

<i>ogolit'</i> 'bare'	< <i>golyj</i> 'nude'
<i>obnovit'</i> 'renew'	< <i>novyj</i> 'new'
<i>oblegčit'</i> 'lighten, relieve'	< <i>legkij</i> 'light'

- BECOME Y deadjectival inchoative verbs:

<i>oglupet'</i> 'become stupid'	< <i>glupyj</i> 'stupid'
---------------------------------	--------------------------

- GIVE X denominal factitive verbs:

<i>obvinit'</i> 'accuse'	< <i>vina</i> 'guilt'
<i>ocenit'</i> 'evaluate, assign value'	< <i>cena</i> 'price'
<i>okrylit'</i> 'inspire, lit. give wings'	< <i>krylo</i> 'wing'
<i>opojasat'</i> 'girdle'	< <i>pojas</i> 'belt'

- GET X denominal inchoative verbs:
 obomšet ‘become covered with moss’ < *mox* ‘moss’
 opamjatovat’sja ‘come to one’s senses’ < *pamjat* ‘memory’

On the one hand, the factitive and inchoative verbs in this group include lexemes with concrete spatial semantics like *opojasat* ‘girdle’, *ogolit* ‘bare’, and *obomšet* ‘become covered with moss’ which represent prefixal submeanings 10.SURROUND, 11.AFFECT A SURFACE, and 12.ENVELOP. On the other hand, the factitive and inchoative verbs comprise lexemes that describe changes in physical and psychological states like *op’janit* ‘make drunk, intoxicated’, *oglupet* ‘become stupid’, *ogorčit* ‘grieve, distress’, *opečalit* ‘sadden’, and *okamenet* ‘become petrified, turn to stone’, thus providing a conceptual link to the metaphorical submeanings 14.METAPHORICAL SURROUND and 13.METAPHORICAL ENVELOP⁷⁹. In other words, the factitive / inchoative meaning 15.IMPOSE/ACQUIRE A NEW QUALITY is conceptually motivated by both spatial and metaphorical submeanings of O-, OB-, and OBO- and is thus well-incorporated into the proposed network of polysemy.

5.3.4 Radial category profiles of the prefixes O- and OB-

In the previous subsection I have shown that the semantics of the three prefixes in question can be described within a single model of polysemy, where the spatial and non-spatial submeanings are interrelated and motivated by a single prototype. However, across this network of submeanings the prefixes are not equally frequently attested. For the clarity of argument, I leave the use of OBO- aside and focus on the semantic distribution of O- vs. OB- alone.

While sharing the same semantic network, O- and OB- differ in terms of how frequently they express various submeanings. For example, submeaning 1.MOVE AROUND AN OBJECT is attested more frequently for the prefix OB- (in thirteen verbs) than for the prefix O- (in four verbs). There are differences between the two prefixes with regard to other submeanings as well. Those submeanings where a given prefix is dominant constitute its “center of gravity” – the pattern which I call a *radial category profile* of a prefix, following Nessel et al. 2011. Radial category profiles can be identified through the comparison of the relative frequency distribution of relevant data that represent O- and OB- in each of their submeanings. The relative frequency distribution of O- and OB- points to differences in their productivity for certain parts of the radial network. These differences can also be interpreted in terms of prominence: which submeanings are more prominent for O- and which ones for OB-? Is there always a contrast or is there any region of semantic overlap?

Table 2 summarizes the total numbers of individual verbal lexemes (type frequencies) which represent the fifteen submeanings of the three prefixes. In other words, Table 2 provides a detailed overview of the prefixes’ distribution across their submeanings and prefixed verbs. This quantitative information highlights those parts of the radial category where a given prefix is more prominent and where the largest contrast between O- and OB- can be observed.

⁷⁹ Będkowska-Kopczyk (2012: 26, 33) discusses corresponding related verbs with the prefix O(B)- in Slovene *občutiti* ‘feel’, *ogorčiti* ‘aggrieve’, *ostrāšiti* ‘frighten’, *odreveneti* ‘stiffen, turn numb (into wood)’, *okamneti* ‘petrify’ and argues that the use of O(B)- in verbs denoting human emotions is metaphorical and is motivated by the spatial meaning of encirclement.

Prefix Submeaning	O-	OB-	OBO-
1.MOVE AROUND AN OBJECT	4	13	1
2.PASS BY	2	6	1
3.OVERTAKE	1	1	2
4.OUTDO	3	4	1
5.DECEIVE	4	12	1
6.MISTAKE	8	1	1
7.OVERDO	5	23	2
8.METAPHORICAL PASS BY	4	2	3
9.AFFECT A "CIRCLE" OF OBJECTS	6	30	1
10.SURROUND	25	35	2
11.AFFECT A SURFACE	116	89	3
12.ENVELOP	24	79	4
13.METAPHORICAL ENVELOP	5	32	3
14.METAPHORICAL SURROUND	13	16	2
15.IMPOSE / ACQUIRE A NEW QUALITY	MAKE X BE Y	180	35
	BECOME Y	151	23
	GIVE X	103	33
	GET X	25	5

Table 2: Distribution of prefixed verbs across the prefixes and their submeanings.

Figure 3 visualizes the proportion of O- (light grey bars) vs. OB- (dark grey bars) in terms of relative frequency distribution shown on the horizontal axis. The total number of verbs that constitute 100% is different in each line of the diagram. For instance, 6.MISTAKE accounts for the total of 8+1=9 verbs, while 15.IMPOSE/ACQUIRE A NEW QUALITY yields 459+96=555 verbs. The raw numbers given in each bar are the type frequencies of the listed submeanings and correspond to the numbers of verbs provided in Table 2.

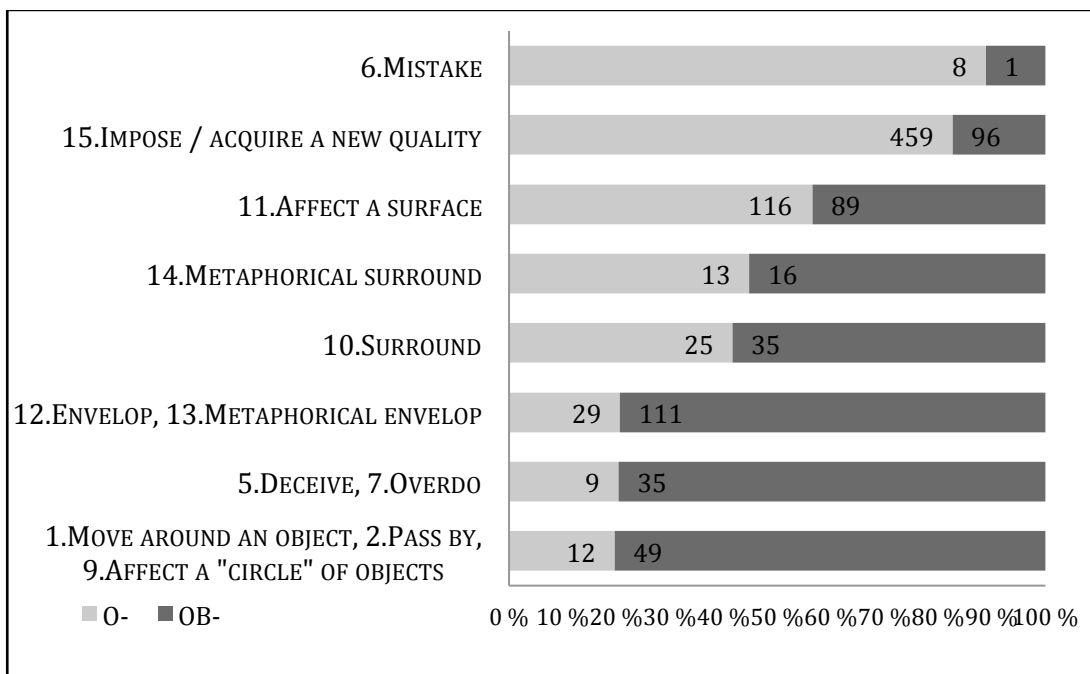


Figure 3: Distribution of the prefixes O- and OB- across semantic subcategories.

Along the vertical axis of Figure 3, there is a list of prefixal submeanings. Most submeanings are represented by a separate bar, but some submeanings are grouped together because they are semantically similar and otherwise they yield raw numbers that are too sparse for statistical analysis. I group together those submeanings that are closely related and profile the same dominant prefix. In particular, 1.MOVE AROUND AN OBJECT, 2.PASS BY, and 9.AFFECT A “CIRCLE” OF OBJECTS are represented by a single bar because all three of them belong to the concrete spatial domain. Grouping these submeanings together does not skew the overall quantitative pattern. Similarly, I group the submeanings 12.ENVELOP and 13.METAPHORICAL ENVELOP, because they are semantically close and share the predominance of the prefix OB-. Finally, the grouping of 5.DECEIVE and 7.OVERDO is justified, because they are both metaphorical extensions of 2.PASS BY and prioritize OB- over O-, as opposed to 6.MISTAKE which adopts the opposite strategy.

The distribution of verbs prefixed in O- and OB- across submeanings of these prefixes represents the correlation between the phonological shape of the prefix and its radial category profile. Statistical analysis shows that this correlation is statistically highly significant, meaning that the observed distribution is not random. Pearson’s Chi-square test shows that the distribution in Figure 3 is highly significant: $\chi^2=296$, degrees of freedom = 7, p-value < 2.2e-16. In addition, I calculate the Cramer’s V value in order to estimate the effect size of this distribution, that is the strength of association between the variables. Variable 1 is the prefix (O- or OB-); and variable 2 is the radial category profile, which is here the string of raw quantitative values from Figure 3. Given that Cramer’s V can vary from 0 to 1 – from absence of a relationship to a relationship of complementary distribution of variables, the effect is traditionally described as weak if $V=0.1$, medium if $V=0.3$, and strong if $V=0.5$ (King & Minium 2008: 329). In case of our distribution (cf. Figure 3), Cramer’s V value equals 0.5, proving that the difference between the two radial category profiles discovered in the data not only is statistically significant, but also displays a strong effect size.

What do these profiles look like? Figure 3 demonstrates that the prefix O- predominates in submeanings 15.IMPOSE/ACQUIRE A NEW QUALITY and 6.MISTAKE. At the other extreme, the prefix OB- prevails in the spatial submeanings 1.MOVE AROUND AN OBJECT, 2.PASS BY, 9.AFFECT A “CIRCLE” OF OBJECTS, and 12.ENVELOP, as well as in the metaphorical submeanings 5.DECEIVE, 7.OVERDO, and 14.METAPHORICAL ENVELOP. The two prefixes show a more balanced distribution for two spatial submeanings 10.SURROUND and 11.AFFECT A SURFACE and in one non-spatial meaning 14.METAPHORICAL SURROUND. We can conclude that in this area their dominances overlap. To a greater or lesser degree the overlap takes place in each submeaning, because both prefixes are attested in each node of the network. This observation provides evidence for the close connection of O- and OB-. Crucially, their distribution across submeanings forms a continuum, or a scale, as can be seen in Figure 3, rather than a clear-cut opposition.

We can also approach the distribution of the two prefixes in question from a slightly different perspective. Table 3 and Figure 4 present the distribution of O- and OB- across the same groups of submeanings as in Figure 3, but 100% equals here a total of individual verbal lexemes with a given prefix. Thus, there are 671 verbal lexemes prefixed with O-, and percentage ratio of semantic groups indicates quantitative proportions of these uses of O- within an entire set of verbs with this prefix.

Submeaning	O-	O- %	OB-	OB- %
1.MOVE AROUND AN OBJECT, 2.PASS BY, 9.AFFECT A "CIRCLE" OF OBJECTS	12	1.78%	49	11.34%
5.DECEIVE, 7.OVERDO	9	1.34%	35	8.10%
12.ENVELOP, 13.METAPHORICAL ENVELOP	29	4.32%	111	25.69%
10.SURROUND	25	3.72%	35	8.10%
14.METAPHORICAL SURROUND	13	1.93%	16	3.70%
11.AFFECT A SURFACE	116	17.28%	89	20.60%
15.IMPOSE/ACQUIRE A NEW QUALITY	459	68.40%	96	22.22%
6.MISTAKE	8	1.20%	1	0.23%
Total:	671	100%	432	100%

Table 3: Distribution of O- and OB- across semantic groups of submeanings.

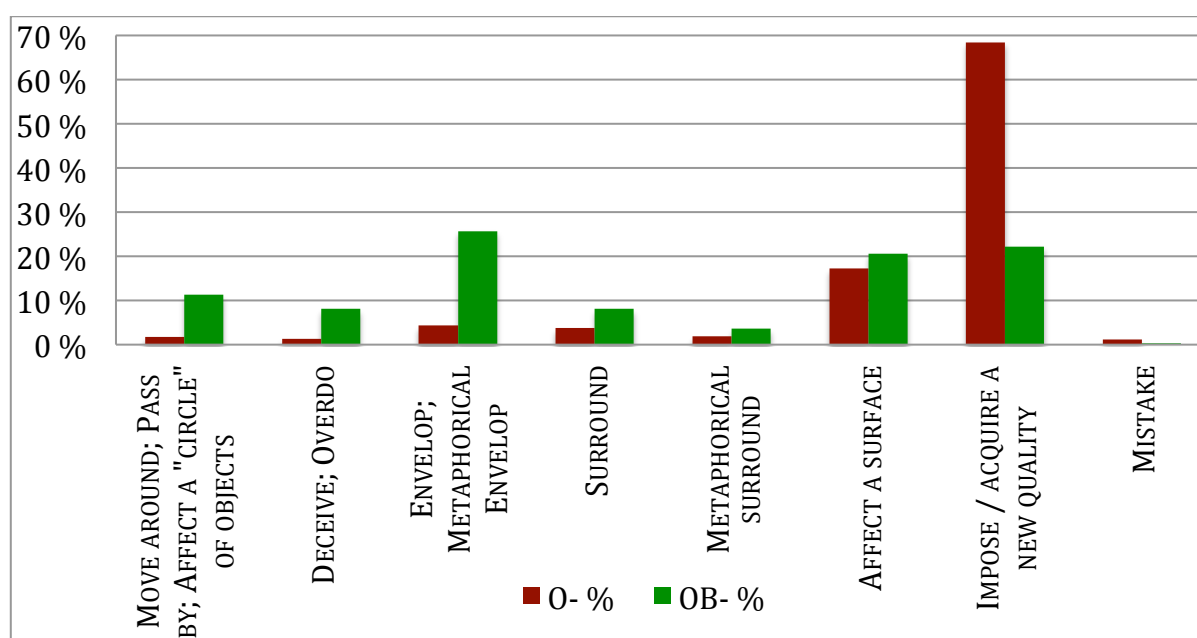


Figure 4: Radial Category Profiles of the prefixes O- and OB-.

Figure 4 visualizes relative frequency distribution of the two prefixes and clearly indicates the different centers of gravity that these prefixes have within the shared network of polysemy. Again, we can observe that most uses of the prefix O- correspond to the submeanings 11.AFFECT A SURFACE and 15.IMPOSE/ACQUIRE A NEW QUALITY, whereas the prefix OB- predominates in the three leftmost groups of submeanings: {1.MOVE AROUND AN OBJECT, 2.PASS BY, 9.AFFECT A "CIRCLE" OF OBJECTS}, {5.DECEIVE, 7.OVERDO}, and {12.ENVELOP, 13.METAPHORICAL ENVELOP}.

Along with these differences of radial category profiles, Figure 4 shows that there is a distributional overlap in each group of submeanings. Moreover, in the submeaning 11.AFFECT A SURFACE the two prefixes invest comparable amounts of their uses: 116 lexemes in O- and 89 lexemes in OB- which yield 17% and 20% of their overall uses respectively. This overlap concerns both semantics and distribution of O- and OB- and suggests that the distinction between the two prefixes is not clear-cut. Rather, there is a continuum with strong specialization.

However, a closer look at contrastive verbal pairs shows that the very semantic specialization of O- and OB- for different submeanings is gradual in nature.

In some minimal pairs like *obsudit* ‘discuss’ – *osudit* ‘condemn’ and *obžit* ‘render habitable’ – *ožit* ‘revive’ the semantic difference between the two prefixes is quite salient: OB- expresses submeaning 14.METAPHORICAL SURROUND, while O- bears the factitive submeaning 15.IMPOSE/ACQUIRE A NEW QUALITY, which is clearly more abstract and situated further away from the spatial prototype. Yet the two submeanings in question are closely related and are expressed simultaneously in verbs *odurmanit* ‘stupefy’ and *očarovat* ‘charm’.

In another minimal pair the verbs *obogret* ‘warm all over’ and *ogret* ‘swipe, hit someone hard (with a stick or other tool)’ look very different at first glance. Yet, the prefix OBO- realizes the submeaning 12.ENVELOP, whereas O- expresses a closely related submeaning 11.AFFECT A SURFACE.

In the minimal pairs discussed so far, the different semantic nuances of the prefixes result in quite substantial differences between the prefixed verbs. This is not the case in the minimal pair *obledenit* ‘cover with ice’ – *oledenit* ‘make ice-cold, turn into ice’, where the contrast between O- and OB- leads to a much less distinct difference of verbal derivatives. And in pairs like *obsypat* (*saxarom*) ‘strew (with sugar)’ – *osypat* (*zolotom*) ‘strew (with gold), make rich’, the distinction between the opposed counterparts shifts to register and collocations.

At last, the contrast between *obstrič* and *ostrič* ‘cut off’ or between *obbežat* and *obežat* ‘run around something’ is barely tangible, and the variants are interchangeable.

To sum up, when O- and OB- attach to the same verbal stems, the degrees of difference between the two prefixes are not the same, but form a scale. This scale includes both a strong opposition between the two prefixes on the one hand and a near-equivalence in the semantics of the two on the other hand.

The process that occurs when the two variants of a morpheme develop semantic differences seems very natural in the light of semiotics: the differences in form (shape) lead to differences in content (as discussed in Andrews 1984: 485). In this subsection, I defined these differences in terms of radial category profiles which highlight areas of near-complementary dominance and areas of overlap. Although O- and OB- show different patterns of attraction among submeanings, those submeanings are closely related and belong to a single radial network.

5.3.5 Statistical analysis:

Combined Classification Tree & Random Forests Model

The Radial Category Profiling analysis focuses only on the semantic distribution of O- and OB-. In order to take into account not only semantic, but also phonological and prosodic factors that can predict the choice of the prefix, the corpus data was additionally tagged according to the following variables:

Cluster Onset: yes (cluster onset), no (simple onset), vowel (no onset);

Onset place of articulation: labials, dentals, alveopalatals, velars;

Onset manner of articulation: sonorants, stops, fricatives, affricates;

Stress of target verb: stem, theme, other;

Base: verb, adjective, noun, ambiguous (for multiple motivations)

Both place and manner of articulation that characterize the initial consonant of the base (OnsetPlace and OnsetManner) were tagged according to Timberlake 2004. If a base begins with a vowel, it is tagged as ‘No onset’.

In order to account for a possible prosodic effect, I tagged the prefixed verbs for their stress pattern (StressTargetVerb): whether the stress is on the stem (e.g. *opénit* ‘cover with foam’, *osmýslit* ‘think over’, *ofórmít* ‘form, shape’) or on the thematic vowel suffix (theme) (e.g. *operít* ‘cover with feathers’, *osnastít* ‘equip’, *očexlít* ‘put into a case’). I classified as ‘other’ those verbs that represent neither of these patterns and carry stress on the infinitive suffix like *obojtí* ‘walk around’ and *oplestí* ‘weave around’.

All verbs that are synchronically analyzable were tagged for motivating base (Base). Most prefixed verbs are motivated by a verb (*obvesti* ‘lead around’ < *vesti* ‘lead’), an adjective (*obnovit* ‘renew’ < *novyj* ‘new’), or a noun (*obnulit* ‘turn into zero’ < *nul* ‘zero’). Those verbs that have multiple bases are tagged as ‘ambiguous’: e.g. *obvinit* ‘accuse’ < *vinit* ‘blame’, *vina* ‘guilt’. Verbs formed from a phrase like *olicetvorit* ‘personify’ (< *lico* ‘face, person’ + *tvorit* ‘create’) are tagged as if they have a nominal base. Verbs like *obezoružit* ‘disarm’ formed from a prepositional phrase (*bez oružija* ‘without weapon’) are tagged for nominal base, but the phonological characteristics of the onset are those that belong to the preposition *bez* ‘without’ which is adjacent to the prefix.

Among prefixed verbs in the database there are thirty-two verbs that are deetymologized: e.g. *oboronit* ‘defend’, *obresti* ‘obtain’, *odolet* ‘overcome’, *objazat* ‘oblige’, *oxmurit* ‘seduce’, *obidet* ‘offend’, etc. Such verbs pose problems in terms of defining the stem and its initial phoneme which is sometimes truncated or omitted (e.g. *oblačit* ‘put a garment over’ < *vlačit* ‘drag’). Since the etymological bases of such verbs can be unclear from a synchronic perspective, assigning clear semantics to the prefix is often problematic. Due to these reasons, I leave deetymologized verbs out of the statistical analysis.

Consonantal characteristics such as voiced vs. voiceless and hard vs. soft I leave out of this study because they were not discussed as prefix predictors in Table 1.

For statistical analysis I used the model of Classification Trees combined with Random Forests. A Classification Tree is able to capture complex interactions of multiple factors, while the Random Forests can rank the predictors with respect to their relative strength (Strobl et al. 2009). The properties of this model are discussed in greater detail in Chapter 9 (Section 9.5.3.6.3).

The dataset for this analysis includes 1,005 verbs; that is all analyzable verbs in the database. The dependent variable was Prefix (O-, OB-, or OBO-). The predictor variables were stress pattern of the prefixed verb (StressTargetVerb), onset place and manner of articulation (OnsetPlace, OnsetManner), type of the onset in the base (ClusterOnset), and type of the base (Base).

In order to avoid collinearity in the model, I excluded the prefix Submeaning from the inventory of tested predictors. Otherwise the factor Submeaning would account for many of the distinctions that are already captured by the factor Base. There are several reasons why this model is the optimal one. First, the assignment of structural motivations like a simplex verb or a simplex adjective are more objective than the assignment of submeanings to a prefix, no matter how comprehensive the proposed model is. On the one hand, submeanings of the prefixes often overlap with the lexical meaning of the base, especially in Natural Perfectives (Janda et al. 2013). Moreover, sometimes the semantics of a prefixed verb is not entirely compositional. The combination of the prefix with the base is often idiomatized and not clearly semantically decomposable. On the other hand, in many verbs the prefix can be simultaneously assigned more than one meaning, even with regard to the same lexical meaning of the prefixed verb. For example, the prefix of the verb *obsmolit* ‘cover or saturate with resin’

(< *smolit* ‘cover or saturate with resin’, *smola* ‘resin’) expresses the spatial submeaning 11.AFFECT A SURFACE and the factitive submeaning 15.IMPOSE/ACQUIRE A NEW QUALITY. However, such cases cannot be captured by the chosen statistical model, because it can only account for a single submeaning of the prefix in each prefixed verb (in other words, one value of each variable per time). In addition, the factor Base has four values (verb, adjective, noun, ambiguous), while the factor Submeaning has fifteen different values that are rather hard to group together without losing some of their distinctions. Finally, the factor Base corresponds best to the Split Hypothesis which posits two structural derivational patterns – deverbal verbs with morpheme OB- vs. deadjectival and denominal verbs with the morpheme O-. This analysis thus targets exactly the factors proposed by the Split Hypothesis supplemented by an additional factor StressTargetVerb.

Classification Tree analysis⁸⁰ yields the graph shown in Figure 5. The Tree shows how the data can be broken down to smaller subsets in the most optimal way, so that the factor variables could predict the choice of the dependent variable – the prefix. The Tree visualizes an algorithm of such recursive partitioning of data, where each level of daughter nodes is more pure than the parent nodes because they isolate parts of data with the majority of either variable value (Strobl et al. 2009: 326).

Figure 5 demonstrates that the distribution of data is governed by a complex interaction of factors. At the outset, the most statistically optimal split of data is made by the factor Base. The verbs are subdivided into two groups. To the left there are those that have a verbal base. To the right there are those that have an adjectival base, a nominal base or several bases, both verbal and non-verbal (tagged as ‘ambiguous’).

At the next level of split, nodes 2 and 13 introduce significant phonological predictors – ClusterOnset and OnsetManner (of articulation).

Deverbal prefixed verbs with a consonant cluster in the root morpheme (ClusterOnset: yes) form a group which is further partitioned by the predictors StressTargetVerb (node 3) and OnsetManner (node 5). Deverbal prefixed verbs with a single consonant in the onset of the root morpheme (ClusterOnset: no) or a root which starts in a vowel (ClusterOnset: vowel) form another group. This group is further subdivided by the factor OnsetManner in the nodes 8 and 10. Thus, within deverbal verbs the highest located predictor is the factor ClusterOnset, followed by StressTargetVerb and OnsetManner. Now we are in a position to compare this part of the tree with the subsplits of the other part.

In most verbs motivated by a noun, adjective, or multiple bases (ambiguous), the prefix shape can be predicted by the phonological factor OnsetManner (node 13), which sets off the verbs with a root that starts in a sonorant or a vowel. These are further motivated at a more granular level by another phonological factor – ClusterOnset (node 19). On the other hand, in those verbs where the root does not start in a sonorant or a vowel, the prefix further depends on an interaction of prosody (StressTargetVerb - node 14) with segmental phonology (OnsetManner - node 15). Summing up, in deadjectival and denominal verbs as well as transitional verbs with both verbal and non-verbal bases, the distribution of prefixes is governed by the same phonological and prosodic factors, but these factors interact with each other in a different way. In particular, here the factor OnsetManner is located higher than ClusterOnset and StressTargetVerb.

⁸⁰ The formula used for Ctree is `dat.ctree = ctree(Prefix ~ StressTargetVerb + OnsetPlace + OnsetManner + ClusterOnset + Base, data=dat)`.

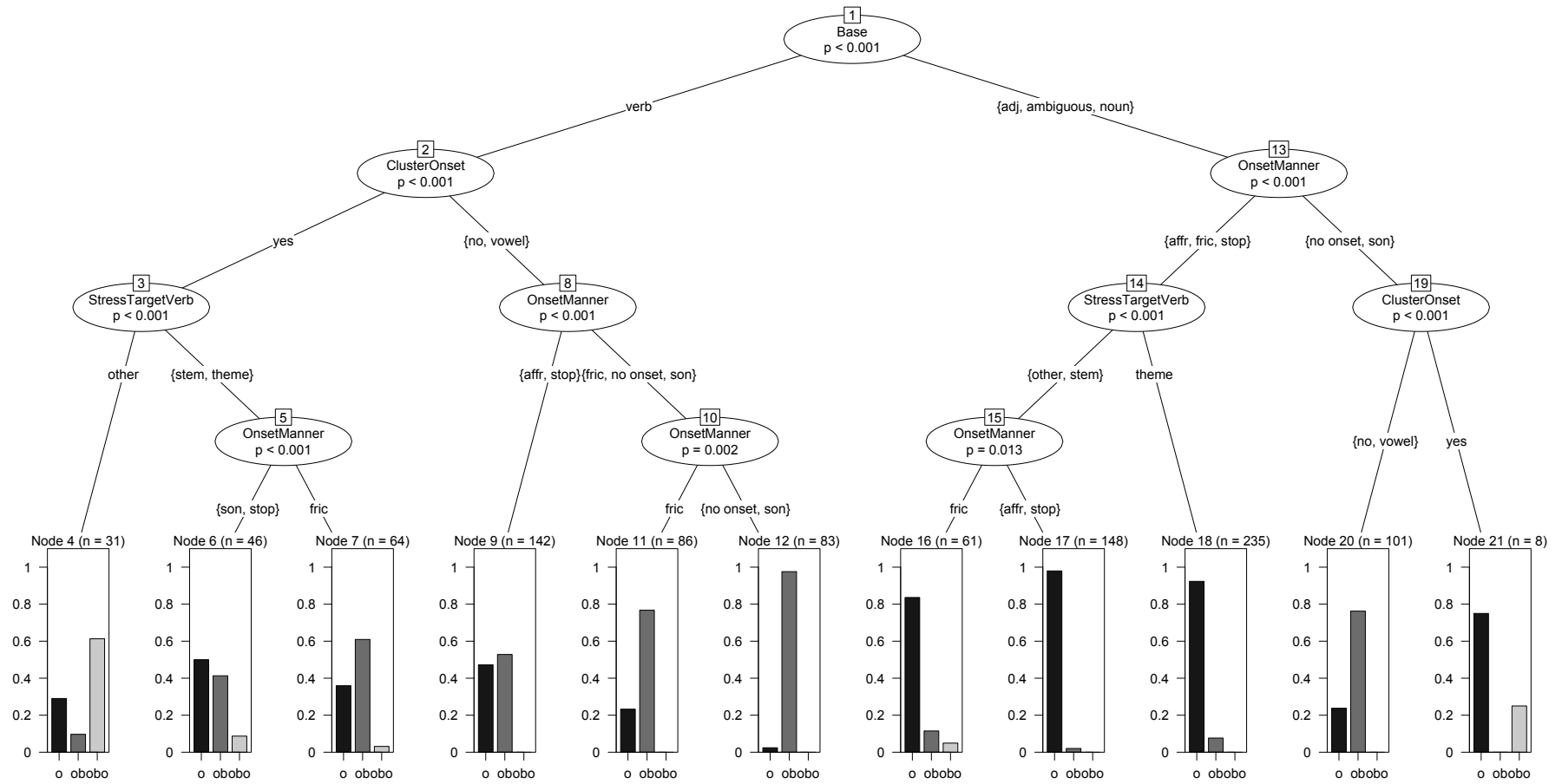


Figure 5: Classification tree for the corpus data on O- and OB-.

Crucially, phonological factors interact differently within two large domains set up by the type of the Base – deverbal verbs vs. verbs with adjectival, nominal and multiple motivating bases. This outcome supports the claim of the Split Hypothesis: indeed, complex clusters turn out to be more predictive within the group of deverbal verbs, while sonority and manner of articulation of the onset have more predicting power in the class of deadjectival, denominal, and multiply-motivated verbs.

However, the result of Classification Tree analysis suggests many insights that were not articulated in the Split Hypothesis. First of all, in each of eleven resulting subsets of data there is an overlap of two or all three prefixes. No complementary distribution is found. This means that the data partitioning accounts only for statistically robust tendencies in the distribution of data. In each subset there is always a minority of verbs, where the shape of the prefix cannot be explained by the series of most optimal splits.

Second, the Tree reveals a statistically significant effect of stress patterns. Stress turns out to have a significant impact in the domains of both verbal and non-verbal bases. This observation is new and was never discussed in the literature before.

Third, the model refines our knowledge about the phonological conditioning of the three prefixes in question. For example, nodes 20 and 21 demonstrate that in verbs motivated by an adjective, noun, or several bases that start in a vowel or a sonorant, the prefix OB- predominates only if the base does not have an onset consonant cluster (node 20). If a sonorant consonant is a part of a consonant cluster and thus forms a complex onset, the prefix OB- does not occur at all, but O- and OBO- appear, with high predominance of the former (node 21). Another observation comes from nodes 9, 11, and 12. These nodes demonstrate that in deverbal verbs with a root that starts in a single consonant or a vowel, the prefix OB- predominates in the majority of datapoints (see the grey middle bar in the square boxes of nodes 9, 11, and 12). However, the prefix O- appears under this phonological condition as well, and it is almost as strong a competing option as OB-, if the root starts in an affricate or a stop obstruent (node 9). This result refines the statement about the labials *b* and *p* that necessarily trigger the prefix O-.

While the Classification Tree visualizes the interactions of factors, the Random Forest analysis makes it possible to compare the predictive power of the factors in terms of different degrees of statistical significance. Figure 6 shows the result of the Random Forest analysis. The barplot in Figure 6 presents those factors that have a statistically significant impact on the distribution of verbs across the three prefixes. Here one can see the same four factors that appeared in the Classification Tree in Figure 5, plus the factor OnsetPlace (of articulation).

For each factor the model calculates and assigns a score of importance. According to these scores, the factors are located along a single scale. The length of the bars in Figure 6 corresponds to the relative importance scores of the factors. The barplot demonstrates that the phonological factor OnsetManner is the most powerful predictor in data distribution, followed by Base and less important ClusterOnset, OnsetPlace, and StressTargetVerb. Note that although the Base appears in the topmost split of data in Figure 5, the strongest factor in the overall distribution is nevertheless OnsetManner. Recall that OnsetManner determines five different data splits in the Classification Tree (nodes 5, 8, 10, 13, 15), whereas Base appears only once.

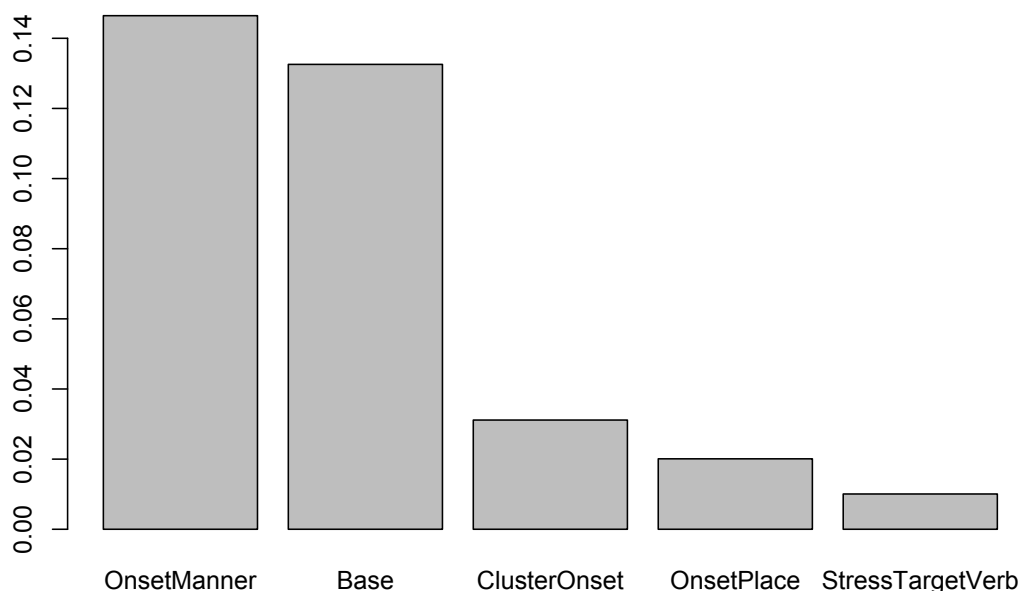


Figure 6: Barplot of corpus data: Variable importance scores for factorial predictors of O- vs. OB-.

In terms of statistics, this is a non-trivial methodological result: the factor that determines the topmost partition of data in the Classification Tree does not appear to be the strongest overall predictor according to the Random Forests model. The hierarchy of factors, set up according to their scores of importances, suggests that phonology (OnsetManner) overrides the morphological structure (Base) although their scores are rather close. Moreover, as I show in Section 5.4.3.2, this result is parallel to the outcome of the Random Forests analysis of experimental data.

To sum up, the Random Forests modelling of corpus data features the two factors proposed by the Split Hypothesis – derivational pattern (Base) and phonology (OnsetManner) – as the strongest predictors of the prefix. However, this model ranks these factors in the reverse order: phonology comes first, followed by the derivational pattern. This contradicts the prediction of the Split Hypothesis.

I will now address the issue of various derivational patterns in more detail.

5.3.6 Patterns of derivation and intermediate cases

According to the Split Hypothesis, the morphemes O- and OB- correspond to two derivational patterns. The morpheme OB- is employed for deverbal derivation⁸¹, while the morpheme O- belongs to the morphological constructions (circumfixes) O-...-IT' and O-...-ET' used in denominal and deadjectival derivation.

My findings on derivational patterns of corpus data make it possible to refine these statements. Table 4 demonstrates that the prefixed verbs associated with both

⁸¹ The term *derivation* accounts well for Specialized Perfectives. With regard to Natural Perfectives, there is a controversy in the literature on Russian aspect whether it is appropriate to discuss them in terms of derivational morphology. The relationship between members of aspectual pairs is ambiguous. Percov 2001 explicitly shows that the nature of this relationship cannot be described entirely by *derivation* or by *inflection* (cf. also Zaloznjak & Šmelev 2000: 14-16 for discussion). This issue lies beyond the scope of this study and does not affect the argument.

verbal and non-verbal simplex bases (like *obvinit* ‘accuse’ < *vinit* ‘blame’, *vina* ‘guilt’; *op’janit* ‘make drunk’ < *p’janit* ‘make drunk’, *p’janyj* ‘drunk’) are very frequent.

Perfective Type	Definition	Examples	# verbs
Natural Perfectives	Perfective verbs that form aspectual pairs with corresponding simplex imperfectives and are formed via a prefix that does not alter the lexical meaning of the simplex (so-called “ <i>čistovidovaja pristavka</i> ”).	<i>obstrič</i> ‘cut off’ <i>ocarapat</i> ‘scratch’ <i>ostynut</i> ‘cool down’	45
Natural Perfectives that are simultaneously Factitive Perfectives	A Natural Perfective that has both a verbal and an adjectival or nominal simplex base.	<i>okruglit</i> ‘make round’ <i>oblicevat</i> ‘facet’ <i>opečalit</i> ‘sadden’ <i>op’janit</i> ‘make drunk’	150
Specialized Perfectives	Perfective verbs formed via a prefix that alters the lexical meaning of the imperfective simplex.	<i>ožit</i> ‘revive’ <i>obryt</i> ‘dig around’ <i>obognat</i> ‘outstrip’	406
Specialized Perfectives that are simultaneously Factitive Perfectives	A Specialized Perfective that has both a verbal and an adjectival or nominal simplex base.	<i>obvinit</i> ‘accuse’ <i>obostrit</i> ‘sharpen’ <i>obmaslit</i> ‘cover with oil’ <i>obledenit</i> ‘cover with ice’	182
Factitive Perfectives (including factitives and inchoatives)	A perfective verb that has only nominal or adjectival simplex base and lacks a verbal simplex base.	<i>okrylit</i> ‘lit. give wings’ <i>obnovit</i> ‘renew’ <i>obomšet</i> ‘become covered with moss’	222
Deetymologized perfective	Simplex base is synchronically unclear.	<i>obmanut</i> ‘deceive’ <i>obnjat</i> ‘embrace’ <i>oblačit</i> ‘robe’	32
Total			1,037

Table 4: Types of perfective verbs formed by the prefixes O-, OB-, and OBO-.

Verbs that are doubly motivated constitute the vast majority of Natural Perfectives (aspectual pair counterparts of imperfective verbs, cf. Janda 2007) found with these prefixes – 150 (77%) verbs of a total of 195 Natural Perfectives (=45+150, cf. the second and the third lines of Table 4).

In addition, verbs that are doubly motivated constitute almost one-third of all Specialized Perfectives, that is 182 verbs (fifth line of Table 4; 31% of the total amount of Specialized Perfectives which equals 406+182=588 verbs).

In total, verbs with double motivations comprise 332 (150+182) lexemes (marked in grey for reader’s convenience), or 32% of all verbs prefixed in O-, OB-, and OBO- attested in the corpus-based database. These verbs are transitional between exclusively deverbal and exclusively denominal and deadjectival derivation. I suggest that the existence of such verbs and their abundance demonstrate that the two derivational patterns proposed within the Hypothesis of Morphological Split are in fact

not as categorical and unambiguous as they are assumed to be. Quite the opposite, the examples form a continuum rather than a binary opposition of two derivational patterns.

5.3.7 Hypothesis testing against lexicon: Discussion of results

In this section the Split Hypothesis was tested against the Modern Russian lexicon. Two factors are found to have a statistically significant impact on the shape of the prefix – type of simplex base (verbal vs. non-verbal) and manner of articulation characteristic of the initial consonant of the root. This indicates that the distribution of prefixes according to phonological variables and derivational patterns does take place. At the same time, this correlation should not be overgeneralized to the entire body of data.

I showed that the two semantic domains that were presented in the Split Hypothesis as distant and unrelated can be analyzed as parts of a single semantic network. I demonstrated that spatial and factitive submeanings can be incorporated into a single cognitive model. Factitive and inchoative uses of the prefixes can be viewed as metaphorical extensions of spatial semantics. The spatial change of an object via surrounding, coverage, or wrapping serves as a source domain for non-spatial qualitative changes in the target domain of human emotions and behavior. The link between the spatial and factitive uses is justified by a large number of transitional verbs with multiple motivations that belong to both domains. The verbs with double verbal-adjectival (or verbal-nominal) motivations comprise 32% of the database. This contradicts the clear-cut distinction between deverbal and deadjectival/denominal derivational patterns expected according to the Split Hypothesis.

Radial category profiling shows that OB- is frequently used in spatial submeanings, and O- in factitive and inchoative submeanings. On the other hand, I identify extensive semantic overlap in the use of O- and OB-, and also productivity of O- for some spatial submeanings. Moreover, in many denominal and deadjectival verbs the prefix O- expresses both factitive and spatial senses at the same time.

To sum up, corpus data confirms the tendency for distinctive use of O- and OB-, but also reveals their close relationship and overlap. The latter obstructs assignment of independent morphological status to these two prefixes.

5.4 Experiment

Corpus data presents challenges for the Split Hypothesis. However, one can argue that the prefixes in Russian verbs are lexicalized and idiomatized, so that no actual choice of a prefix takes place when an utterance is generated. Highly frequent prefixed verbs might be reproduced by speakers as precomposed units, together with their “ingrown” prefixes. Thus, one might doubt that the distribution of prefixes across standard verbs reflects any on-going processes conditioned by productivity. On the other hand, this does not dismiss the findings described in the previous section, because the corpus study represents the type of lexical data that shapes the linguistic competence of native speakers. Yet, such a study cannot verify predictions on what determines the choice of the prefix when novel words are generated. Recall that modern patterns of productivity are a key argument in favor of the Split Hypothesis. So far linguists tried to describe the mechanisms that govern contrastive productivity patterns of O- and OB- by analyzing sporadic novel words. Following up on this research, I conduct a psycholinguistic

experiment which specifically targets elicitation of the production of new words. Such production is governed by consistent mechanisms in native speakers' mental grammar.

5.4.1 Goal, design, subjects, administration

The goal of the experiment was to test the Split Hypothesis against active mechanisms of word production. The key idea was to collect native speakers' responses on the choice of the prefix under fixed conditions set up according to three factor variables – **semantics** (spatial vs. change-of-state semantics), **phonology** (initial phoneme of the simplex base), and **prosody** (place of stress in a simplex base). The effect of prosody has not been explored before, but it had to be taken into account in order to control for all variables introduced by the stimuli.

These three factors became major parameters that determined the experimental design. In order to address each of these factors separately and measure their impact on the choice of a prefix, the factors were "isolated" from each other by means of three types of questionnaires – A, B, and C. Each subject was exposed to only one type of questionnaire.

Each questionnaire contained sixty-two short narratives that were two or three sentences long. Each narrative was preceded by a real or nonce word accompanied with a brief definition. In each narrative, there was a gap in the place of a verb. The gap was marked with dots. The subjects were asked to fill the gap by generating a perfective verb prefixed with O-, OB-, or OBO-. More than one variant was allowed. In order to generate a verb, the subjects had to use a simplex base (a stimulus) given at the beginning of each narrative. In questionnaires A and B, the bases were verbs of motion like *letet'* 'fly' and *bežat'* 'run', whereas in questionnaire C the bases were adjectives like *rususkij* 'Russian' and *živoj* 'alive'. After going through a set of instructions and a few examples, a subject was asked to read all narratives out loud, generating the missing prefixed verbs spontaneously without giving the task a second thought. Below I provide examples of experimental tasks containing nonce words typical for each type of questionnaire (for the full list of experimental narratives see Appendix 3):

(7) Questionnaire type A / B

Gúzvit' (A) / Guzvít' (B) – (o ptice) prixramyvát', pritvorjajas', čto odno krylo slomano.

Ptica zametila lisu i stala manit' ee proč' ot gnezda. Pripodnjav odno krylo, kak budto ono bylo slomano, ptica otbežala čut' podal'se, vokrug kamnja, i, podoždav, kogda lisa posleduet za nej, vzmyla vverx.

'Gúzvit' (A) / Guzvít' (B) – (of a bird) to limp pretending that a wing is broken.

The bird noticed a fox and began luring it away from the nest. It raised a wing as if it was broken, and ran a little bit further away, around a rock, and after waiting for the fox to follow it, the bird took off.'

(8) Questionnaire type C

Guzvyj – krasočnyj.

Prišla osen' i prinesla s soboj svežest', veter i novye kraski. Osen' raskrasila list'ja v zoloto i purpur, zastelila tropy mjagkim kovrom, prevratila les v bogato ukrašennyj, guzvyj terem. Osen'-masterica postaralas' na slavu, preobrazila les, ego.

'Gúzvyj – colorful.

The fall came bringing along freshness, wind, and new colors. The fall painted leaves with gold and purple, covered the paths with soft carpets, turned the forest into a palace, rich in ornament and colorful. The mistress-fall did its best, it transformed the forest, it.'

The three prefixes available for creating a prefixed verb were always the same (O-, OB-, and OBO-), but the verbs and adjectives offered as stimuli were different. In questionnaires A and B, real Russian verbs like *bežat'* 'run', *vesti* 'lead', *gnat'* 'chase', *katit'* 'roll', *nesti* 'carry' occurred alongside nonce verbs⁸² like *guzvit'*, *loprit'*, *čavit'*, *znupit'*, *žgavit'*, etc. Similarly, in questionnaire C, real adjectives like *svetlyj* 'light', *grubyyj* 'rude', *složnyj* 'complex', *nemeckij* 'German' were interspersed with nonce adjectives like *guzvyj*, *lopryj*, *čavyj*, *znupyj*, and *žgavyj*. In this way the experiment brought into play the phonological variable, namely the initial phonemes of the motivating base. Note that I made use of the same nonce stems *guzv-*, *lopr-*, *znup-*, shaping them as verbs (e.g. *guzvit'*) for the first group of subjects who received questionnaires A and B, and shaping them as adjectives (e.g. *guzvyj*) for the second group of subjects who worked on questionnaire C. The questionnaire types A and B differ only in place of stress on stimulus verbs – on the stem, as in *gúzvit'* (type A), or on the thematic vowel, as in *guzvít'* (type B).

Table 5 summarizes the distribution of sixty subjects across questionnaire types, stimuli (verbs with stress on the stem, verbs with stress on the theme, and adjectives), and target responses (verbs of motion and factitive verbs).

Subjects	Target verbs	Expected response	Stimuli	Stimulus example	Q	Subjects number
Group 1	Verbs of motion	<i>ob-guzvit'</i>	Verbs	<i>gúzvit'</i>	A	15
				<i>guzvít'</i>	B	15
Group 2	Factitive verbs	<i>o-guzvit'</i>	Adjectives	<i>gúzvyj</i>	C	30

Table 5: Distribution of the subjects across types of questionnaires (Q).

Table 5 shows that a stimulus verb like *gúzvit'* unambiguously conditions the spatial semantics of the target verb and consequently triggers the prefix's meaning 1.MOVE AROUND AN OBJECT. Therefore, the expected response to this stimulus in a non-problematic phonological context (i.e. initial obstruent consonant of the base) was OB-: *obguzvit'*. Following the same logic, a stimulus-adjective like *guzvyj* necessarily implies a factitive meaning of the target verb and the corresponding meaning of the prefix – 15.IMPOSE A NEW QUALITY. Given that the stem *guzv-* introduces an unproblematic phonological

⁸² The use of nonce words in this survey is justified as a common fruitful practice employed in experimental studies of this kind (cf. Gor & Chernigovskaya 2003; Rodina 2007; Makarova 2009).

context (obstruent consonant in the onset of the base), the response expected for this stimulus should contain the prefix O-: *oguzvit'*. Summing up, the semantics of the prefix was prescribed by the morphological type of the stimulus base (verb vs. adjective) and by the semantics of syntactic context provided in the narrative. In addition, the initial phonemes of the stimulus stem capture the phonological and prosodic⁸³ factor variables.

The use of nonce words made it possible to test forty-six stems that begin with twenty different single consonants and eight diverse consonant clusters. Because the presence of the vowel/zero alternation (or the so-called “fleeting” or “mobile” vowel, cf. Chapter 3) in the root is a historically conditioned property of individual morphemes (Gouskova 2012), in the experimental setting it was impossible to provide nonce words with such an alternation. Also, in order to maintain an optimal length for an experimental trial, it was necessary to limit the scope of tested phonemes to hard paired consonants like *b, l, z*, thus excluding their palatalized (soft) counterparts *b', l', z'*. The consonants' hardness was determined by the use of the Cyrillic letters *a, o, y* for adjacent vowels in the nonce stems, whereas the letters *я, ё, ю, у, е* which normally mark palatalization were excluded. However, the experiment includes the soft pairless consonants *j, č* and *šč* which lack a hard counterpart. On the other hand, the phoneme *f* is excluded because it is never preceded by *b* (Roberts 1981: 72; Andrews 1984: 478). Otherwise, the initial phonemes of the nonce stems represent the entire inventory of Russian consonant phonemes, listed in Table 6.

Table 6 provides a list of all stimuli and corresponding expected responses (target verbs) predicted on the basis of 1) prefix semantics and 2) the onset of the base. Grey shading highlights those single initial consonants which are evaluated in Table 1 as unproblematic for both O- and OB-, no matter what semantics they bear. Therefore, under the unproblematic phonological condition the choice of the prefix should depend entirely on the target semantics of the prefixed verb (factitive vs. spatial), and a maximal contrast between the two morphemes should take place in this phonological domain. Crucially, the onsets marked in grey favor neither O- nor OB-, so they leave the choice of the prefix to the impact of semantic factors. The spatial vs. factitive semantics of the prefix and the target verb are triggered by the morphological class of the stimulus (verbs vs. adjectives) and the pattern of derivation (prefix vs. circumfix accordingly). The nonce onsets in grey are therefore the key environments where semantic contrast is tested, while the remaining onsets of nonce words as well as all real words serve as controls and distractors.

⁸³ The impact of stress pattern on the choice between O-, OB- and OBO-, has not been explored before. However, the study of corpus data showed that the place of stress has a statistically significant effect on the choice of the prefix (Section 5.3.5). Also, subjects who participated in the pilot version of the experiment commented on the role of stress in stimuli. Thus, stress patterns gained special attention as an additional variable which should be controlled. I examined the correlation of stress with the prefix in 809 attested verbs which preserve the stress on the same syllable as in the simplex base. The findings are described in detail in Baydimirova 2010: 70-74. Here I limit myself to the observation that the prefix OBO- can be attached only to stems that carry stress on the initial syllable. This suggests that in the experiment nonce adjectives should be of the type *tkábyj* (with the stress on the initial syllable), but not *tkabój* (with the stress on the ending). This condition creates an optimal context for each of the three prefixes to be realized, whereas stimuli like *tkabój* eliminate the possibility of OBO-.

C	Adjectival stimuli	Expected target verbs	Verbal stimuli	Expected target verbs
b	bostyj	o-bostit'	bostit'	o-bostit' (ob-bostit')
	buklyj	o-buklit'	buklit'	o-buklit' (ob-buklit')
v	vurlyj	ob-vurlit' (o-vurlit')	vurlit'	ob-vurlit'
	važdyj	ob-važdit' (o-važdit')	važdit'	ob-važdit'
g	guzvyj	o-guzvit'	guzvit'	ob-guzvit'
	gabyj	o-gabit'	gabit'	ob-gabit'
d	duktyj	o-duktit'	duktit'	ob-duktit'
	damlyj	o-damlit'	damlit'	ob-damlit'
ž	žaxlyj	o-žaxlit'	žaxlit'	ob-žaxlit'
	žusklyj	o-žusklit'	žusklit'	ob-žusklit'
z	zopryj	o-zoprit'	zoprit'	ob-zoprit'
	zupyj	o-zupit'	zupit'	ob-zupit'
j	jupyj	ob-jupit' (o-jupit')	jupit'	ob-jupit'
	jalyj	ob-jalit' (o-jalit')	jalit'	ob-jalit'
k	kočlyj	o-kočlit'	kočlit'	ob-kočlit'
	kampyj	o-kampit'	kampit'	ob-kampit'
l	lusyj	ob-lusit' (o-lusit')	lusit'	ob-lusit'
	loprij	ob-loprit' (o-loprit')	loprit'	ob-loprit'
m	murlyj	ob-murlit' (o-murlit')	murlit'	ob-murlit'
	momlyj	ob-momlit' (o-momlit')	momlit'	ob-momlit'
n	nadyj	ob-nadit' (o-nadit')	nadit'	ob-nadit'
	nokryj	ob-nokrit' (o-nokrit')	nokrit'	ob-nokrit'
p	puryj	o-purit'	purit'	o-purit' (ob-purit')
	patlyj	o-patlit'	patlit'	o-patlit' (ob-patlit')
r	roglyj	ob-roglit' (o-roglit')	roglit'	ob-roglit'
	ražnyj	ob-ražnit' (o-ražnit')	ražnit'	ob-ražnit'
s	saglyj	o-saglit'	saglit'	ob-saglit'
	suryj	o-surit'	surit'	ob-surit'
t	tulyj	o-tulit'	tulit'	ob-tulit'
	tovyj	o-tovit'	tovit'	ob-tovit'
x	xopyj	o-xopit'	xopit'	ob-xopit'
	xušnyj	o-xušnit'	xušnit'	ob-xušnit'
c	cavyj	o-cavit'	cavit'	ob-cavit'
č	čupyj	o-čupit'	čupit'	ob-čupit'
	čavyj	o-čavit'	čavit'	ob-čavit'
š	šadryj	o-šadrit'	šadrit'	ob-šadrit'
	šaklyj	o-šaklit'	šaklit'	ob-šaklit'
šč	ščulyj	o-ščulit'	ščulit'	ob-ščulit'
gn	gnoryj	o-gnorit'	gnorit'	ob-gnorit' (obo-gnorit')
žr	žrapyj	o-žrapit'	žrapit'	obo-žrapit'
zn	znupyj	o-znupit'	znupit'	ob-znupit' (obo-znupit')
čt	čtusyj	o-čtusit'	čtusit'	obo-čtusit'
žg	žgavyj	o-žgavit'	žgavit'	obo-žgavit'
sp	spulyj	o-spulit'	spulit'	ob-spulit'
sk	skolyj	o-skolit'	skolit'	ob-skolit'
tk	tkabyj	o-tkabit'	tkabit'	obo-tkabit'

Table 6: Stimuli and responses expected according to the Split Hypothesis.

The initial sonorants *l, m, n, r, j* and the consonants *b, p* and *v* are biased towards either one or another prefix because of their phonological preferences. At the same time, these onsets tolerate both O- and OB-, although not on equal terms: recall *o(b)bežat'* 'run around' and *obnovit'* 'renew' but *onemet'* 'become dumb'; *obmirščit'* 'secularize' but *omolodit'* 'rejuvenate'. For such problematic phonological environments Table 6 provides two possible target verbs – the primary and the parenthesized – additionally possible but less expected response.

The bottom part of Table 6 contains nonce words with complex onsets. Underlining marks those initial consonant clusters that normally do not appear with a preceding *b* in Modern Russian: **bžr*, **bčt*, **bžg*, and **btk* (McGranahan 1975: 14-15). By contrast, the clusters *bgn, gzn, bsp*, and *bsk* are well-formed and attested in such verbs as *obgnit'* 'decompose', *obznakomit'sja* 'become familiar with many people' (Efremova 2000), *obsprašivat'* 'ask many people', and *obskakat'* 'gallop'. If a nonce stem has a complex onset, the expected target verbs should be prefixed in O- for the factitive meaning, but with OB- or OBO- for the spatial meaning. Under the spatial semantic condition with incompatibility of the initial consonant cluster with *b* (like **bžr*), the subjects are expected to prefer OBO-, but not O-, according to the hierarchy of allomorphs proposed for the morpheme OB-: *ob-* > *obo-* > *o-* (recall Table 1).

The issue of authenticity and well-formedness of nonce words demands an additional explanation. All nonce stems for this experiment were created manually, taking into account major well-formedness constraints of Russian phonotactics. I adopted a methodology well-described in the literature: each nonce word is created on the basis of a real Russian word via minimal modification of one or two phonemes (cf. Makarova 2009: 32). For example, the nonce adjective *lustyj* is created from the real adjective *lysyj* 'bald' by replacement of the root vowel. Likewise, the nonce verb *purit'* is based on the real verbs *burit'* 'drill' and *durit'* 'play tricks'. However, to follow this principle consistently was a hard task because each nonce root, once constructed on the basis of a real verb, has to resemble a real adjective, and vice versa. On the other hand, it was important to avoid creating nonce words that are too similar to real words, in order to escape unwanted associations that could bias subjects' responses. Therefore, in order to succeed in making the nonce words sound natural and conform Russian phonotactics, I consulted the reverse dictionary of Russian by Greve & Kroesche (1958), the study of McGranahan (1975), and my own native speaker intuitions. As many subjects pointed out, nonce words were often considered to be unfamiliar dialectal or archaic Russian words. This suggests that the nonce words meet the requirement of being native-like in their shape.

Although the spatial and change-of-state semantics of the prefixes in question are related by many real perfective verbs of "transitional" type associated with both verbal and nominal bases (e.g. *okol'cevat'* 'encircle with a ring' < *kol'cevat'* 'encircle with a ring', *kol'co* 'ring'), in the experiment these two semantic types had to be clearly contrasted and narrowed down to specific submeanings. As a result, the two broad semantic domains proposed in the Split Hypothesis were limited to the submeanings central in each domain. All the transitional types were avoided. The spatial semantics of O- and OB- was represented by the prototypical sense 1.MOVE AROUND AN OBJECT, clearly perceivable in motion verbs: e.g. *obletet' vokrug gnezda* 'fly around the nest'. The semantic domain of qualitative change was represented by the submeaning 15.IMPOSE/ACQUIRE A NEW QUALITY narrowed down to the factitive deadjectival pattern MAKE X BE Y: e.g. *osložnit'* 'complicate' < *složnyj* 'complex', not **složnit'*. In order to avoid overlap of the spatial and the factitive sense in the experiment, the narratives and the

definitions of nonce words introduce situations which unambiguously describe not spatial but qualitative change of an object with regard to human emotions, personal qualities, and food preferences. The change of an object, usually a person, described in the narratives, is triggered by a specific situation or a new experience. On the one hand, these situations have to be realistic enough to be easily perceived and understood by the subjects. On the other hand, the definitions of nonce adjectives must describe qualities that otherwise lack a separate Russian term: e.g. *bostyj* 'able to make beautiful dishes out of clay', *spulyj* 'unable to work with one's right hand', *žgavyj* 'having an obsessive desire to wash one's hands all the time'. With regard to verbal stimuli this was even more important. Given that Russian verbs of motion form a close and rather small class, the experiment made it possible to create nonce verbs for a variety of motion types that lack Russian separate terms: e.g. *jupit* '*peredvigat'sja na lyžax bez lyžnyx palok*' (move on skis without poles), *nokrit* '*peredvigat'sja na kolenkax*' (move on one's knees), *xopit* '*stremitel'no peremeščat'sja, prygaja pri ètom na skakalke*' (move fast jumping over a jump rope), *saglit* '*peredvigat'sja zadom napered, izredka ogljadyvajas', čtoby ne upast*' (walk backwards, looking over one's shoulder occasionally so as not to fall). The full list of Russian definitions of created nonce words and their translations into English are available in Appendix 4.

Note that the definitions of nonce verbs given above contain real verbs like *peredvigat'sja* and *peremeščat'sja* 'move' instead of verbs of motion like *idti* 'walk', *bežat* 'run', *exat* 'drive', *letet* 'fly'. This is a deliberate strategy of the experimental design. By contrast, in the preliminary pilot version I used definitions that contained real motion verbs: *šaklit* '*exat' verxom na verbljude*' (ride a camel), *loprit* '*idti, gromko topaja nogami*' (walk while loudly stamping one's feet). Piloting revealed a priming effect: no matter which nonce verb the subjects looked at, they would choose the prefix attested in Russian for the real motion verb used in the definition. Moreover, some subjects explicitly reported that they adopted this strategy. In other words, while choosing a prefix for the stimulus *šaklit*, they would look at *exat* 'drive, ride' given in the definition and produced *ob-šaklit* by analogy with *ob-exat*'. Similarly, if a definition contained the verb *idti* 'walk', the subjects would vote for the prefix OBO- in the nonce verb by analogy with *obo-jti*. In order to avoid this priming effect, I excluded from the definitions all motion verbs that have standard prefixed derivatives with O-, OB-, or OBO-. As a result, all verbs of motion were replaced with neutral verbs *peredvigat'sja* and *peremeščat'sja* 'move', which lack the counterparts **o(b)peredvigat'sja*, **o(b)peremeščat'sja*. Verbs like *val'sirovat* 'waltz' and *prixramyvat* 'limp slightly' were preserved in definitions because they lack derivatives in O- and OB-.

A priming effect was potentially possible for adjectival nonce stimuli too: *čupyj – sil'no pjanyj* 'very drunk' > O- in *opjanit* 'make drunk' > O- in *očupit*'. For this reason, the definitions for adjectival stimuli were reformulated, so that the wording would not bias the prefix choice via analogy. A number of strategies were employed: the use of a participle (e.g. *suryj – pogružennyj v unynie* 'dejected'), prepositional phrases (*čavyj – s xorošimi manerami* 'with good manners'), particles of negation in combination with *sposobnyj* 'able' (no **o(b)sposobit*) (*lusyj – ne sposobnyj est' rybu* 'not able to eat fish'), constructions with *polnyj* 'full of' (*kampyj – polnyj nadežd i novyx planov* 'full of hopes and new plans'), and combinations of several strategies (*čupyj – nemnogo vypivšij* 'a little drunk'). Real adjectives that lack factitive derivatives in O- or OB- were also used in definitions: *jupyj – suevernyj* 'superstitious', *šaklyj – razočarovannyj* 'disappointed', given that there does not exist **o(b)sueverit*' and **o(b)razočarovat*'.

Each questionnaire contained sixty-two experimental tasks including sixteen narratives with real word stimuli and forty-six narratives with nonce word stimuli. The narratives were ordered differently in each individual questionnaire. Items containing real words and nonce words were randomized separately. After three or four narratives with nonce words there was always a narrative with a real word.

Real word stimuli performed several functions. They were used as both distractors and controls. Correct responses to real word stimuli demonstrated that the task was properly understood and that the remaining responses were also given accordingly. In addition, the real word stimuli were selected in such a way that their derivatives contained different prefixes and thus prevented the subjects from developing a uniform strategy of prefixation for all tasks: e.g. compare *zloj* 'angry' > *obo-zlit* 'make angry', *nagoj* 'naked' > *ob-nažit* 'open naked', and *živoj* 'alive' > *o-živit* 'revive' for factitive verbs, and e.g. *idti* 'walk' > *obo-jti* 'walk around', *nesti* 'carry' > *ob-nesti* 'carry around', and *plesti* 'weave' > *o-plesti* 'weave around' for circular motion⁸⁴. Beyond that, some real word stimuli allowed more than one prefix option: e.g. *obežat* and *obbežat* for 'run around', or *ozlit* and *obozlit* for 'make angry'). Such stimuli were meant to activate subjects' awareness of their own individual preferences. A few stimuli involved production of prefixed verbs that are rarely used or of marginal status in Modern Russian. For instance, some subjects found it difficult to generate straight off factitives such as *onemečit* 'Germanize', *obrusit* 'russify', and *obamerikanit* 'Americanize' and explicitly expressed doubts concerning the marginal flavor of these words. Such stimuli were useful because they drew the subjects' attention to the choice of the prefix, making them face the full set of options they had in the tasks with nonce stimuli.

In order to be consistent in experimental design, the real word stimuli were presented in the same manner as nonce words: they were given short definitions that would fit the narrative. For example, the stimulus *golyj* was accompanied with the definition *bez ubranstva* 'without furniture and decorations' which corresponded to its submeaning realized in the narrative:

- (9) *Vse kartiny, tareločki i fotografii so sten snjali i upakovali v korobki. Mebel' perezvazili postepenno. Kogda sovsem steny, v komnate poselilos' èxo.*
 'All paintings, plates, and photographs were taken down and packed into boxes. The furniture was gradually removed. When they completely the walls, an echo appeared in the room.'

The subjects were allowed to choose more than one prefix. While a subject was reading the sentences out loud, a researcher was writing down the subject's responses on a separate questionnaire form. In addition, each interview was recorded acoustically

⁸⁴ Not all real word stimuli are verbs of motion in the strict sense. The class of motion verbs in Russian is closed and small, and their derivatives cannot represent all three prefixes O-, OB-, and OBO- equally well. For this reason, the experiment also involves other real spatial verbal stimuli like *gnut* 'bend', *plesti* 'weave', *kopat* 'dig', and *čertit* 'draw'. Target derivatives of these simplex bases contain O-, OB-, and OBO- that express the submeaning 1.MOVE AROUND AN OBJECT and the closely related submeaning 10.SURROUND: *obognut* 'bend around', *oplesti* 'weave around', *obkopat* 'dig around', *o(b)čertit* 'draw a line around'. In order to secure the spatial circular motion semantics in target prefixes, I use adverbial arguments *vokrug* 'around', *po krugu* 'around a circle', *po perimetru* 'along the perimeter of', and *neskol'ko raz* 'several times' to explicate the full-circle trajectory in the closest context of these verbs.

using the software Praat. Contrary to a self-administered questionnaire, this procedure had a number of crucial advantages. It accounted for place of stress, shortened the trial time, and revealed multiple cases of subjects' hesitation when they had a hard time choosing between competing options. Moreover, reading mode helped subjects to have better control over their own responses, especially in the production of nonce words.

The experiment was conducted in April, 2010. I interviewed sixty native Russian speakers who grew up and received primary, secondary, and higher education in Russia⁸⁵. In order to minimize effects of sociolinguistic factors, each subgroup of subjects was balanced according to gender, age, education level and field of professional expertise (cf. Baydimirova 2010: 83-87). Speakers with non-linguistic educational background were given priority. I tried to minimize the impact of possible dialectal differences by selecting subjects from different geographical regions of Russia. The majority of subjects came from St. Petersburg, Moscow, Iževsk, Murmansk, Arxangel'sk, Ul'janovsk, and Dubna. I now turn to the results of the experiment.

5.4.2 Results

Table 7 aggregates the number of stimuli and collected responses. It demonstrates that 3,720 stimuli triggered 3,878 responses. These numbers reveal variation in the choice of the prefix for both verbal and adjectival stimuli and with regard to both real and nonce bases. Many subjects found it challenging to choose a single prefix, finding more than one option appropriate. Table 7 shows that 480 real verbal stimuli triggered 497 responses, and 480 real adjectival stimuli triggered 490 responses. Regarding nonce words, 1,380 nonce verbs triggered 1,444 responses, and 1,380 nonce adjectives triggered 1,447 responses. It is worth mentioning that the variation in case of verbal and adjectival stimuli is relatively the same: compare 1,941 (=497+1,444) responses given to all 1,860 verbal stimuli (nonce and real) and 1,937 (=490+1,447) responses given to all 1,860 adjectival stimuli.

Type of stimuli		Number of stimuli		Number of responses		
Verbal	real	16*30=480	1,860	497	1,941	3,878
	nonce	46*30=1,380		1,444		
Adjectival	real	16*30=480	1,860	490	1,937	
	nonce	46*30=1,380		1,447		

Table 7: Numbers of stimuli and collected responses.

Of key interest are those responses that are triggered by nonce word stimuli. Their distribution across prefixes depending on phonological and semantic conditions is presented in Figures 7-10.

Figures 7 and 8 show that under the condition of nonce verbal stimuli (and spatial target semantics) the subjects prioritized the prefix OB- more often than O-. In the case of nonce adjectival stimuli (and factitive target semantics) the picture is quite the opposite and thus supports the central prediction of the Split Hypothesis.

⁸⁵ A preliminary version of the experiment was tested in a pilot study. Twelve subjects participated, both linguists and non-linguists. Piloting made it possible to correct the instructions, presentation of choices, phrasing of experimental tasks, and administration procedure (cf. details in Baydimirova 2010: 78-83).

However, there is also a large number of responses that do not comply with the predicted contrast. In particular, in case of initial consonant clusters the prefix O- is favored no matter what type of morphology (verbal or adjectival) was the trigger (cf. Figures 9 and 10).

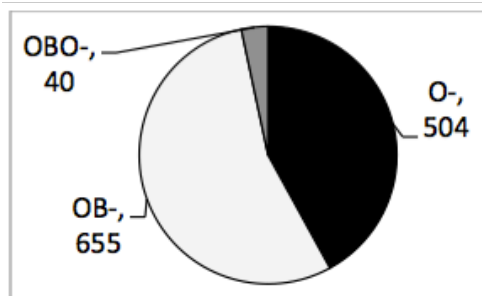


Figure 7: Prefix choices triggered by nonce verbal stimuli with simple onsets.

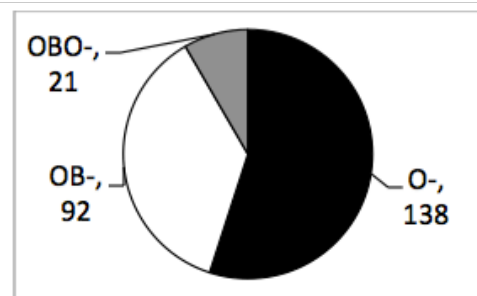


Figure 9: Prefix choices triggered by nonce verbal stimuli with complex onsets.

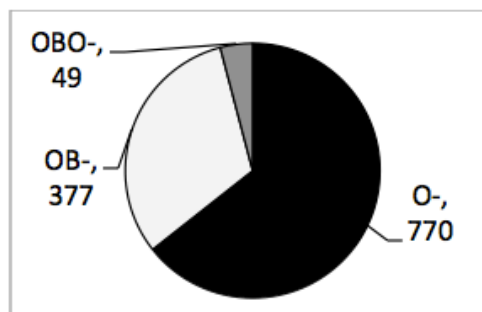


Figure 8: Prefix choices triggered by nonce adjectival stimuli with simple onsets.

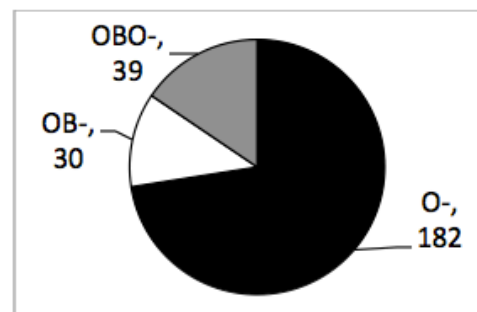


Figure 10: Prefix choices triggered by nonce adjectival stimuli with complex onsets.

In order to gain a more accurate picture, I examine separately those responses that are triggered by nonce word stimuli with phonological conditions unproblematic for both O- and OB- – single obstruent consonants other than *p*, *b*, *v* in the onset of the base. Recall that under this condition the choice of the prefix highlights the impact of target verb semantics. These responses are represented in Table 8:

Chosen prefix	O-		OB-, OBO- ⁸⁶	
	verb	adjective	verb	adjective
Stimulus type	verb	adjective	verb	adjective
Target semantics	spatial	factive	spatial	factive
Number of responses	315	495	382	197
Total:	810		579	

Table 8: Distribution of the prefixes in non-problematic phonological contexts.

In Table 8, shading marks those subjects' responses that comply with the predictions of the Split Hypothesis. They account for the majority of collected responses and reflect the

⁸⁶ Note that Table 8 aggregates the data on OB- and OBO- under a single rubric because the Split Hypothesis postulates them as allomorphs of the morpheme OB-.

statistically significant difference in the distribution of the two prefixes across the two types of stimuli – the verbal and the adjectival.

However, the number of responses that do not correspond to the predictions of the Split Hypothesis is larger than one would expect – 512 responses, which equals 36.8% of all collected responses triggered by the stimuli with unproblematic phonological onsets. This means that in more than a third of the experimental data, the tested semantic factor does not play any role in the choice of the prefix and does not determine the contrastive use of O- vs. OB-, contrary to what the Split Hypothesis predicts. In other words, the overlap (or non-contrastiveness) in the use of O- and OB- is much larger than what the morphological split can account for.

5.4.2.1 Variation in subjects' individual strategies

Figures 11 and 12 present individual subjects' response patterns in the choice of the prefix under the verbal stimuli condition (questionnaires A and B). The two diagrams summarize the responses of each subject and show how frequently each prefix was chosen throughout a questionnaire. Each bar demonstrates a response profile of one subject, labeled as A1, A2, etc. Figure 11 concerns the stress-on-stem condition, while Figure 12 corresponds to the stress-on-theme condition.

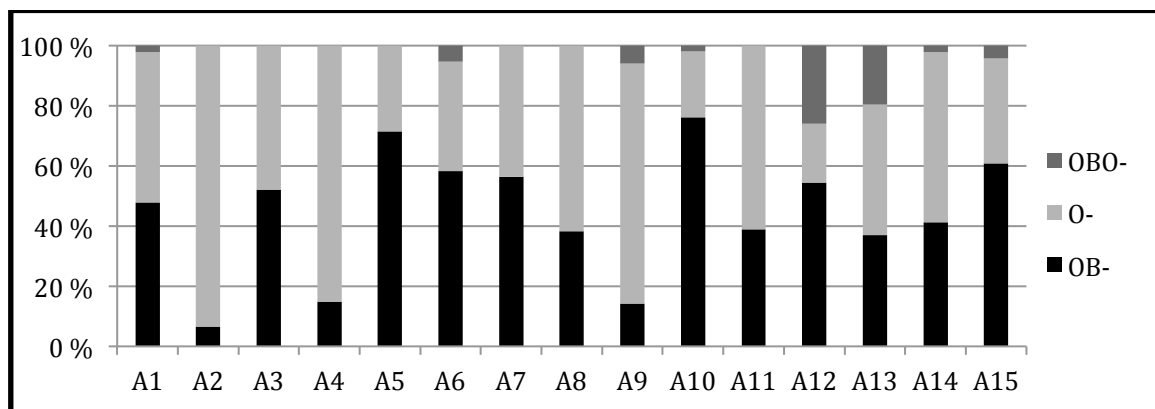


Figure 11: Individual strategies of subjects: questionnaire A (*gúzvit'*).

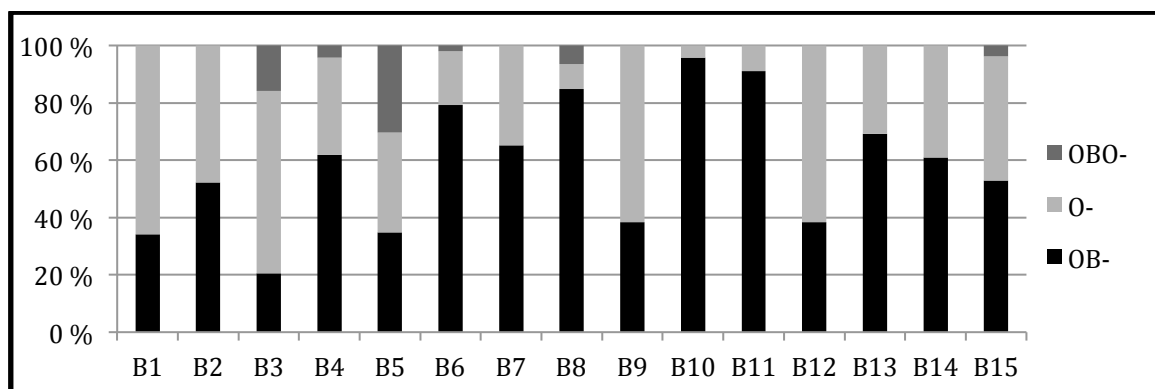


Figure 12: Individual strategies of subjects: questionnaire B (*guzvít'*).

As we can see from the Figures 11 and 12, not all subjects found the prefix OB- most preferable for the spatial meaning. On the contrary, individual patterns appear to be strikingly different. Some subjects preferred the prefix OB-, as expected according to the Split Hypothesis (cf. the subjects coded as A5, A10, B6, B8, B10, B11, where the letter

stands for the type of questionnaire, and the number is an anonymous code for a subject). However, for many subjects the dominant prefix for circular motion was O- (A2, A4, A9, A11, B1, B3, B9, B12), while for a third group of subjects O- and OB- were equally strong competing candidates for spatial semantics (A1, A3, A7, A13, B2, B5, B15). One should also take into account numerous decisions to respond with several prefix options to the same nonce stimulus, as well as a rather high degree of hesitation that many subjects experienced during the experiment. When a subject gave more than one response, the second and the third responses were recorded in the database alongside the first response. The order of responses reflects the real sequence in which the verbs were generated by a subject. The second and third responses are included in the numbers represented in Figures 7-10. However, the statistical models described below account only for the first response.

How can we interpret these findings on patterns of individual variation? They support the idea that different speakers might have different mental grammars, or versions of grammar, whereas a uniform model for grammar of a language is rather an abstraction, as has been argued previously in the experimental studies by Dąbrowska & Street (2006) and Dąbrowska (2012). Figures 11 and 12 suggest that while some speakers of Russian might have the contrast of O- and OB- in terms of specification of a particular prefix for spatial meaning in their mental grammar, other speakers might lack this contrast. Instead, for many speakers O- and OB- appear to be in free variation or competing forms, both able to express the meaning in question. In any case, the degree of variation in subjects' strategies observed in the experimental data cannot be captured within the Split Hypothesis. I conclude that these findings support a closer connection between the prefixes O- and OB- and contradict the Split Hypothesis rather than conform to it.

5.4.3 Statistical analysis

5.4.3.1 Linear Regression: Mixed-Effects Model

In order to diagnose a statistically significant tendency in the distribution of experimental data, we need a model that can distinguish between the subjects' individual preferences and the overall pattern that lies beyond those preferences. Therefore, I apply a Linear Regression Mixed-Effects Model, which can both capture the so-called random factors and calculate the impact of fixed factors (Baayen 2008: Ch.7). In our case the random factors are preferences of 1) individual subjects and 2) nonce stems for a particular prefix. The fixed factors are semantics and phonology.

The analysis in terms of a Mixed-Effects Model⁸⁷ yields a positive effect of random factors (Subject and Stem), as shown in Table 9, and then segregates the impact of fixed factors, summarized in Table 10.

Groups	Name	Variance	Std. Dev
Subject	(Intercept)	1.05651	1.02786
Stem	(Intercept)	0.12268	0.35026

Table 9: Coefficients for random effects in linear regression mixed-effects model (individual preferences of subjects and stems).

⁸⁷ The formula used for this analysis: $\text{FirstResponse} \sim \text{StimulusType} + \text{ClusterOnset} + \text{Manner} + \text{Place} + \text{PossibleWithB} + (1|\text{Stem}) + (1|\text{Subject})$

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.092701	0.447300	-2.443	0.01457 *
Stimulus Type verb	1.380137	0.282964	4.877	1.07e-06 ***
Cluster Onset yes	-0.801274	0.277100	-2.892	0.00383 **
Manner fricative	-0.055833	0.270803	-0.206	0.83665
Manner sonorant	1.297468	0.299502	4.332	1.48e-05 ***
Manner stop	-0.313799	0.309196	-1.015	0.31016
Place dental	0.074999	0.203353	0.369	0.71227
Place labial	-0.369741	0.252897	-1.462	0.14374
Place velar	-0.004172	0.278732	-0.015	0.98806
PossibleWithBTRUE	-0.095956	0.382561	-0.251	0.80195

Table 10: Coefficients for fixed effects in linear regression mixed-effects model.

Grey shading in Table 10 marks those factor variables that have a statistically significant impact on the choice of the prefix. The significance is marked with the symbol * in the rightmost column. The symbol *** indicates a *highly significant* p-value ($p < 0.001$), the symbol ** refers to a *very significant* p-value ($0.001 \leq p < 0.01$), and the symbol * marks a *significant* p-value ($0.01 \leq p < 0.05$) (Gries 2009: 32-34).

In addition I test the responses triggered by verbal stimuli for the impact of stress pattern (on the stem vs. on the theme)⁸⁸. The result is presented in Table 11.

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.02931	0.72403	-1.422	0.15513
Cluster Onset yes	-0.59640	0.23551	-2.532	0.01133 *
Stress Stimulus suffix	-5.12788	2.07949	-2.466	0.01367 *
Age	0.02366	0.02222	1.065	0.28693
Manner fricative	0.14928	0.31595	0.472	0.63659
Manner sonorant	1.07911	0.34768	3.104	0.00191 **
Manner stop	-0.12425	0.32493	-0.382	0.70216
Stress Stimulus suffix:Age	0.25547	0.08568	2.982	0.00287 **

Table 11: Coefficients for fixed effects in linear regression mixed-effects model (responses to verbal stimuli only).

The factors that have a statistically significant effect on the distribution of data are shown in Figures 13 and 14. Two factors – Cluster Onset in the base root (ClusterOnset:yes) and Onset Manner of articulation (Manner:sonorant) – have significant effects in the overall subjects' responses (Figure 13) and in the subset of responses to the verbal stimuli (Figure 14). The effect of these two factors is visualized in the leftmost and the central graphs respectively in both figures.

The vertical axis $p(ob)$ of each graph represents the scale of likelihood for the choice of the prefix OB-. The higher along this axis the graph rises, the more likely the prefix OB- is to be chosen under a particular condition. The horizontal axes stand for the tested variables (fixed factors): complex vs. simple onset of the base (Consonant Cluster: no vs. yes) and onset manner of articulation (Manner: affr = 'affricate', fric = 'fricative', son = 'sonorant', stop = 'stop').

⁸⁸ The formula is: FirstResponse ~ ClusterOnset + StressStimulus * Age + Manner + (1|Stem) + (1|Subject)

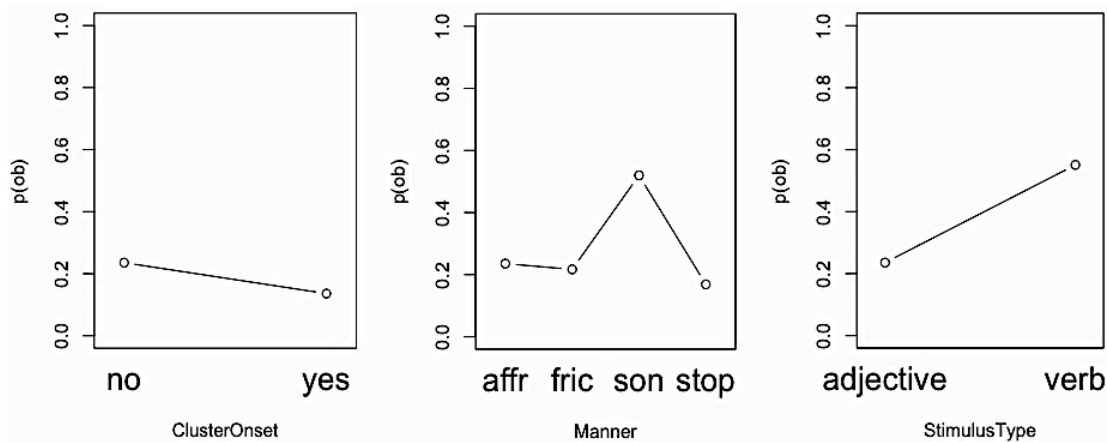


Figure 13: Subjects' responses on both verbal and adjectival stimuli (Questionnaires A, B, C)

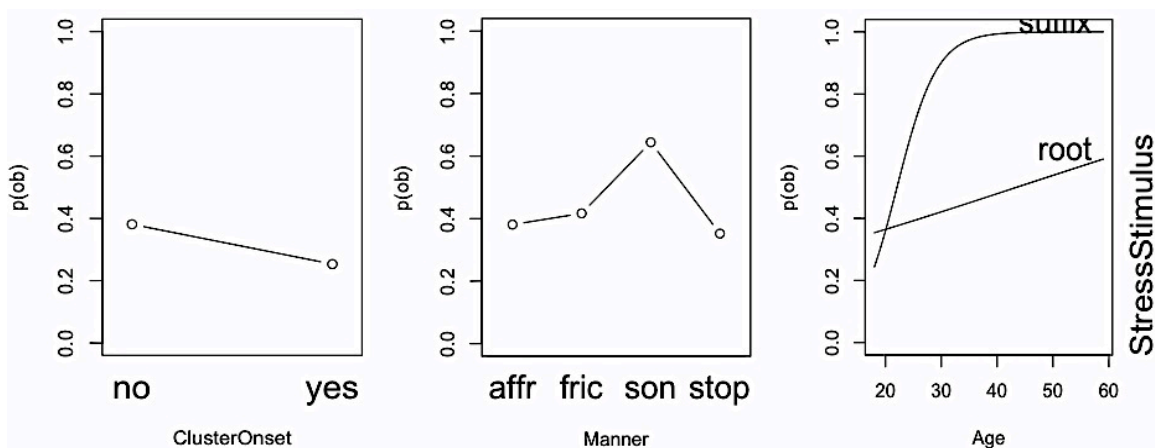


Figure 14: Subjects' responses on verbal stimuli (Questionnaires A, B)

On the leftmost graph of both figures there is a slanted line. It shows that the choice of the prefix OB- is less likely under the cluster onset condition. This means that for simplex bases that feature cluster onsets, statistically subjects prioritize O- over OB-. This holds for both verbal and adjectival stimuli, and that is why the graph is almost identical in both figures.

On the central graphs of both figures there is a zigzag line. The upward spike of the zigzag demonstrates that the likelihood of choosing the prefix OB- correlates with only one feature of the initial consonant of the base – its sonority. This is similar across both verbal and adjectival stimuli, as we see from the similarity of the zigzag lines in both figures. In other words, if a simplex starts in a sonorant consonant, the speakers give priority to the prefix OB-.

Both factors described above are phonological and can be interpreted in terms of Russian phonotactics.

In addition, the **rightmost graph of Figure 13** reveals a significant effect for the third factor – type of stimulus (verbal vs. adjectival trigger), which directly corresponds to the semantics of the prefix (spatial vs. factitive). The graph goes up along the vertical axis of likelihood, demonstrating that the prefix OB- is more likely to be chosen for verbal bases (Stimulus Type: verb). By contrast, the graph line is lower at the opposite edge, showing that the prefix OB- is less likely to combine with adjectival bases (Stimulus Type: adjective).

Last but not least, the **rightmost graph on Figure 14** indicates a statistically significant interaction between the place of stress on the stimulus verb (StressStimulus) and age of subjects (Age). The older a speaker, the more likely s/he is to choose the prefix OB- for verbal simplexes with stress on thematic vowel (suffix *-i-*): thus a stimulus like *guzvít'* triggers the response *ob-guzvit'*. This is an interesting finding that was not described in the literature before. We may speculate that this indicates an on-going change in Russian, however this issue is beyond the scope of my study.

5.4.3.2 Combined Classification Tree & Random Forests Model

In addition to Logistic Regression, I conducted a statistical analysis using a combined model of Classification Trees and Random Forests⁸⁹. Although this model does not account for the random effects (possible individual preferences of subjects and nonce stems), it provides a general picture of the complex interactions of factors in predicting the distribution of the data.

The Classification Tree shown in Figure 15 demonstrates that the interaction of factors is complex. Crucially, there is an overlap of all three prefixes in each subsplit of data (except the Node 13) along with significant preferences for O- vs. OB-. This proves that the distribution of prefixes is not random. Instead, there are robust tendencies.

One of the key advantages of this model is that it accounts for multifactorial conditioning of data distribution showing an overall ranking of factors along a single scale of variable importances (Strobl et al. 2009).

The factors tested by this model are ranked in Figure 16 with regard to each other and to their individual weight in data distribution measured in terms of statistical significance.

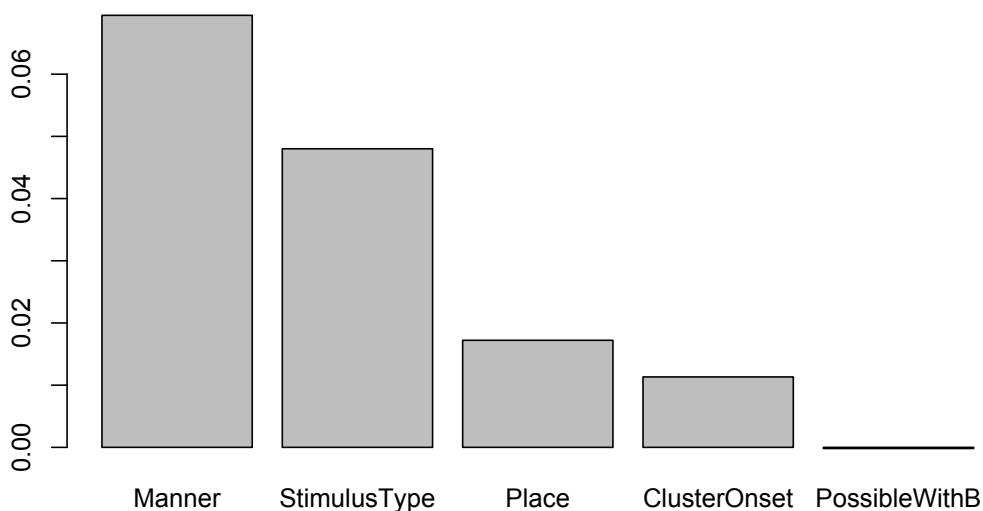


Figure 16: Barplot of variable importance scores for factorial predictors of O- vs. OB-.

Most importantly, the Random Forest model ranks the phonological factor Onset Manner of articulation (Manner) as the strongest predicting factor followed by the Stimulus type (Stimulus: verbal vs. adjectival). Recall that the two types of stimuli correspond to the semantic contrast between target prefixes.

⁸⁹ The formula used for Ctree is `dat.ctree = ctree(Prefix ~ StimulusType + ClusterOnset + PossibleWithB + Place + Manner, data=dat)`.

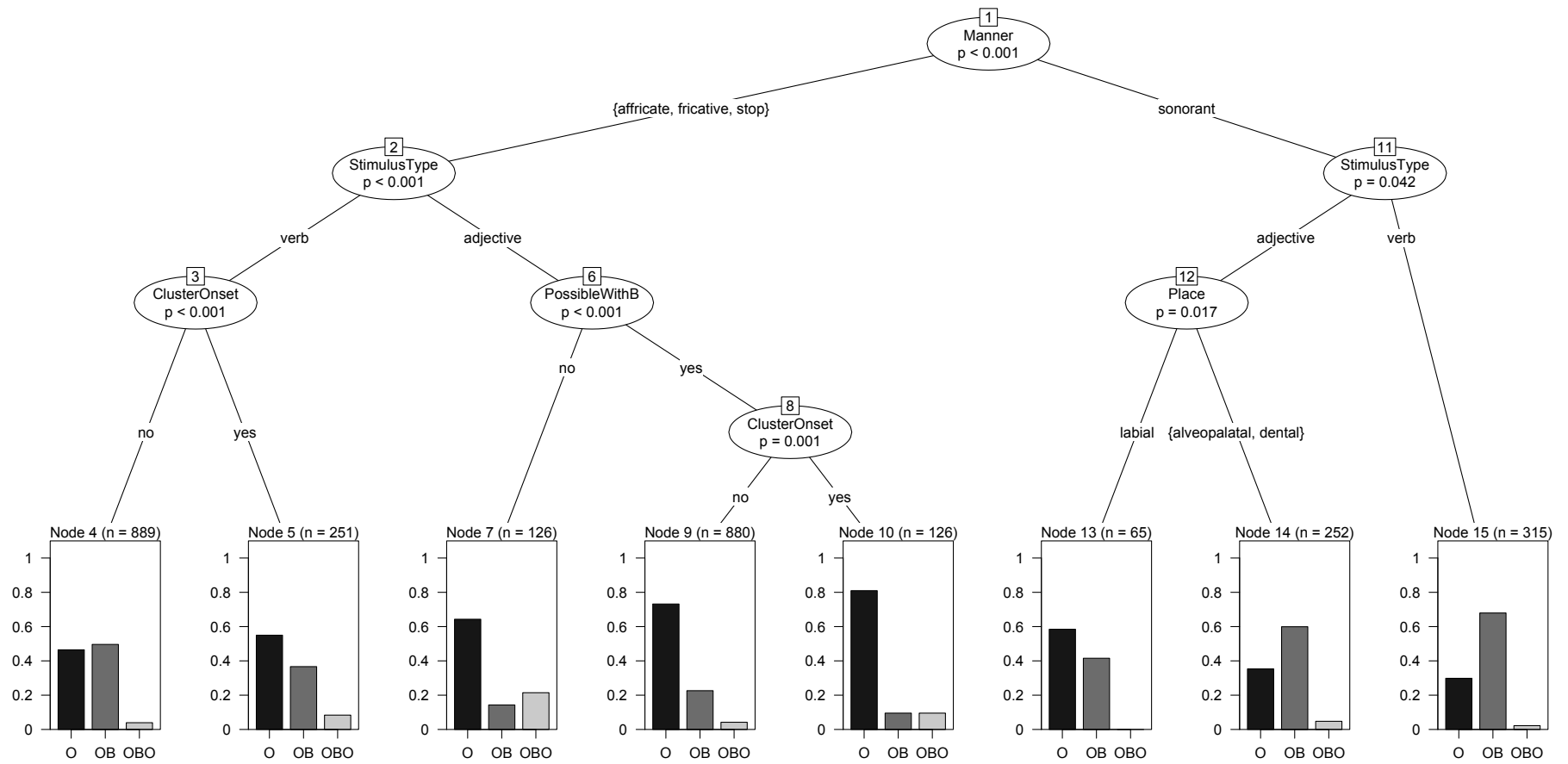


Figure 15: Classification tree for the experimental data on O- and OB-.

The remaining three factors are ranked as less important. Since the phonological factor predominates over the semantic factor, it is reasonable to conclude that O- and OB- represent allomorphs of a single morpheme with sub-morphemic differences.

The plot in Figure 16 presents two additional phonological factors that turn out to have a statistically significant effect on the choice of the prefix – place of articulation of the initial consonant of the base (Place) and the presence of a consonant cluster in the onset of the base root (Cluster Onset). The fifth phonological factor shown in the plot represents the compatibility of the stem's initial consonant cluster with the preceding coda of the prefix (Possible with B). The ranking of factors in Figure 16 suggests that the impact of this last factor is close to zero.

It is worth mentioning that the results of the Random Forests analysis of experimental data are very similar to that of corpus data (recall Section 5.3.5). This suggests that the experiment was conducted properly and that its results are reliable. In corpus data, onset manner of articulation and the type of the base were found to be the most important factors. Their importance scores calculated for corpus data demonstrate that the role of phonology overrides the role of derivational pattern, similarly to what we see in experimental data.

Interestingly, in corpus data the hierarchy of the two less important factors – Place and ClusterOnset – is reversed: the factor ClusterOnset is more predictive than Onset place of articulation, contrary to what was found in the experiment. Note that in the Tree model of corpus data the factor OnsetPlace (onset place of articulation) did not appear at all, whereas in the Ctree model of experimental data the role of this factor is more distinct. In particular, nodes 12, 13, and 14 of Figure 15 demonstrate that the presence of the labials *b* and *p* in the onset of the root makes a statistically significant difference in favor of the prefix O-. According to the Split Hypothesis (Table 1), we expect the impact of initial labials to appear in the domain of deverbal prefixed verbs. Contrary to this expectation, the role of labials in experimental data shows itself in the domain of factitive verbs (Node 13 in Figure 15).

5.4.4 Experimental results: Discussion

The results of the experiment presented above are ambiguous: they partly conform to the predictions of the Split Hypothesis but also contradict it to a large degree.

First, the experiment reveals a larger portion of overlap and variation in the use of O- and OB- than the Split Hypothesis predicts. Perhaps, the most crucial in this regard is the observed lack of clear semantic contrast between the two prefixes in phonologically unproblematic contexts, namely positions preceding non-labial obstruent onsets of the root morpheme.

Second, the collected responses evidence diverse individual strategies adopted by subjects as well as their hesitations and multiple responses with more than one prefix. These phenomena suggest that the two prefixes are closely related and often interchangeable under the same semantic condition.

Third, the statistical analysis shows that both phonological and semantic factors are statistically significant in the choice of the prefix. Yet between the two factors, the phonological factor has a larger overall effect and therefore more weight in prefix distribution.

Summing up, the experimental findings demonstrate that the categorical predictions of the Split Hypothesis about a clear-cut semantic contrast between O- and OB- are highly problematic. In terms of an alternative account I propose a Hypothesis of

Non-Standard Allomorphy which can capture both the similarity and the partial specification of the two prefixes within a cognitive model of a single morpheme.

5.5 Alternative account: O- and OB- represent a single morpheme with Non-Standard Allomorphy

Examination of both corpus and experimental data on the use of the two prefixes shows that the specialization of O- and OB- for two semantic and derivational patterns proposed in Markov 1970, Alekseeva 1978, Andrews 1984, and Krongauz 1998 does exist in Modern Russian. However, I argue that the linguistic status of this specialization should be regarded as a robust tendency rather than an absolute contrast, and the tendency should not be overgeneralized.

The results of both studies acknowledge the statistically significant impact of two factors conditioning the shape of the prefix – the sonorant onset of the base and the morphological type of the base (verbal vs. adjectival). In corpus data the scale of relative importance (Figure 6) supports the observation that the onset manner of articulation is a stronger predictor than the type of the base. Similarly, the experimental data on word-production mechanisms suggests that the phonological factor is more important and has more predictive power than the derivational (= semantic) factor (Figure 16). The sonorant onset in the base motivates the prefix choice for both verbal and nominal simplex bases and thus implies general phonological principles of prefix distribution which override the role of semantics. Thus, the higher ranking of phonology supports allomorphic status of the two prefixes within a single morpheme, while the lower ranking of the semantic factor corresponds to their sub-morphemic differences (cf. Figure 16).

The prefixes O- and OB- are specified with regard to factitive vs. spatial semantics and the adjectival/nominal vs. verbal derivational pattern respectively. Yet, this contrast accounts only for a part of the overall picture. Semantic and morphological analysis of verbal lexemes of Modern Russian reveals a large zone of overlap between O- and OB-. Recall that 32% of all verbs in the database are multiply motivated and are associated with both verbal and nominal simplex bases. The prefix O- in its turn can likewise express the spatial semantics of circular motion. Moreover, the submeaning IMPOSE/ACQUIRE A NEW QUALITY (change of state) expressed by a prefix in factitive and inchoative verbs is incorporated into a single semantic model of polysemy and is, in fact, indirectly motivated by the spatial image schema of circular motion. These observations contradict the contrastive account of the two prefixes and therefore advocate the alternative unified analysis of O- and OB- in terms of a single morphemic unit. This cognitive approach offers an explanation of various possible uses of the two prefixes within a single radial network of interrelated submeanings structured around the spatial prototype of circular motion.

Recall that in the experiment the lack of semantic contrast between O- and OB- was observed in a comparable degree to that found in corpus data: 36.8 % of all responses to unproblematic phonological stimuli do not differentiate between O- and OB- with respect to spatial vs. factitive semantics. The degree of overlap, lack of contrast, and subjects' individual strategies (which are often the opposite of tested predictions) indicate that the two prefixes maintain a close relationship, possible interchangeability, and morphemic cohesion.

The results of the two case studies conform to partial semantic divergence of O- and OB-, but at the same time they remain very close and this cannot be captured as a

morphological split. I argue that the relationship between these prefixes is far from the independence of distinct morphemes. Yet, if one accepts that they represent a single morpheme, the conclusion comes into conflict with how proper allomorphy is traditionally understood. Thus one has to admit that the allomorphy of O- and OB- is defective, because of the undeniable evidence of semantic differences, free variation, and deviations from the rules of phonological distribution. Neither of the two crucial criteria of Standard Allomorphy (identical meaning and complementary distribution) discussed in Chapter 2 is fully satisfied.

I suggest that the existing partial specialization of the two prefixes in question is overgeneralized within the Split Hypothesis as a result of a general terminological and theoretical problem with understanding allomorphy. The traditional theory of allomorphy has only two alternatives – the formants can be either assigned the status of allomorphs of a single morpheme or the status of distinct morphemes. This view itself represents an idealistic simplification of reality. I offer a third alternative which recognizes the possibility of Non-Standard Allomorphy under the cognitive unit of a single morpheme. I propose that the relationship of the prefixes O- and OB- in Modern Russian represents exactly this case. The Hypothesis of Non-Standard Allomorphy does not deny the fact of partial divergence and specialization of these two prefixes. On the contrary, it accounts for this fact by seeing it as a typical semiotic process: difference in form gives rise to difference in meaning. However, I propose that in the case of O- and OB- these differences do not go beyond the bounds of a single morpheme, as evidenced by the zone of overlap in both corpus and experimental data.

To interpret the morphological status of O- and OB- in terms of Non-Standard Allomorphy does not mean to simply return to the previous view which did not recognize their distinctions. The present study contributes a detailed description of the complex interaction of factor variables which condition the distribution of the two prefixes across real verbs and the choice of the prefix in the production of novel words. In this chapter, I have tested phonological, semantic, and prosodic factors, and calculated their relative impact. The present account detects the scope of overlap between the semantics of O- and OB- and employs quantitative methods for modeling and evaluation of the significance of factors.

Proponents of the Structuralist assumption about the binary nature of linguistic contrasts might find the idea of Non-Standard Allomorphy fraught with potential problems, such as loosening the bounds of linguistic categories and diffusing the concept of the morpheme. However, this might be exactly what takes place in languages, at least in Russian prefixes. Indeed, the bounds of the morpheme O- / OB- are apparently diffuse. Here we deal with a continuum of uses that do not fit into narrow clear-cut definitions and terms such as identical meaning and complementary distribution. Looking at this case study, one has to admit that allomorphy is a gradual phenomenon with some more standard instances like RAZ- and RAS- (cf. Chapter 3), and also some non-standard peripheral cases like O- and OB-.

Chapter 6

Grammatically conditioned allomorphy: The prefixes PERE- ‘over, across’ and PRE- ‘very’

6.1 Introduction

Allomorphy via borrowing? Can a borrowed prefix become an allomorph of a native morpheme? Addressing this theoretical question, I present the first corpus-based study of the native Russian prefix PERE- ‘over, across’ and a historically related prefix PRE- ‘very’ borrowed from Old Church Slavonic. As opposed to other Slavic languages, Russian preserves both prefixes and employs them in highly productive patterns of derivation, each in its own domain. In this chapter⁹⁰, I explore the non-trivial relation of these prefixes and investigate whether their status can be described as allomorphy. The goal of this study is to answer the following question: do these prefixes represent allomorphs of a single morpheme or two distinct morphemes in Modern Russian? The two prefixes have comparable but not identical semantics. Some simplex verbs can attach both PERE- and PRE-, showing that the distribution of these prefixes is not complementary. Yet, although neither of the two defining criteria of allomorphic relations is satisfied, I argue that this case can be best captured by extending the traditional understanding of allomorphy.

As opposed to most existing accounts that view PERE- and PRE- as different morphemes, I propose that in Modern Russian these prefixes represent a case of Non-Standard Allomorphy. I argue that PERE- and PRE- are variants of a single morpheme-intensifier that are distributed in terms of morphology: PERE- is a more “verbal” prefix, able to affect verbal categories of aspect and transitivity and extremely productive in the verbal domain as a morphological intensifier of activity (in submeanings REDO, OVERDO, DISTRIBUTE, SUPERIORITY, and THOROUGH). PRE-, on the contrary, is not productive with verbs and often does not affect verbal categories of the base. At the same time, PRE- functions as a productive marker of excess in adjectives, adverbs, some nouns and some words of other classes.

Overall, this case study contributes an empirical investigation of non-standard form-meaning asymmetry that might have far-reaching implications for defining the boundaries of morpheme identity.

6.1.1 Diachronic relation of PERE- and PRE-

It is well-known that Modern Russian has coexistent lexical strata of different Slavic origins: native words and forms of East Slavonic origin exist along with numerous

⁹⁰ The findings reported in this chapter were presented at the seminar on statistics at the Centre for Advanced Study at the Norwegian Academy of Science and Letters in Oslo in 2012, at the 7th Annual Meeting of the Slavic Linguistic Society at the University of Kansas in 2012, and at the International Cognitive Linguistic Conference in Edmonton in 2013. I am indebted to these audiences for helpful discussions and valuable feedback. All the shortcomings are my responsibility.

borrowings from Old Church Slavonic (Sussex & Cubberley 2006: 477). The latter was a liturgical language of the East Slavonic Orthodox Church and was based on a dialect of Old Bulgarian (Tolstoy 2002: 82). The widespread adoption of words and morphemes from Old Church Slavonic greatly influenced the lexical development of Russian and was one of the factors that made it lexically distant from the other East Slavic languages, Ukrainian and Belarusian⁹¹ (Pugh & Press 1999: 4).

The prefix PRE- belongs to the layer of Church Slavonic borrowings and corresponds to the native Russian prefix PERE- (Vasmer 1971: v. 3, 356; Lomtev 1961: 239). Thus, the two prefixes in question belong to two different Slavic layers which coexist in Contemporary Standard Russian.

As shown in Figure 1, the phonological shapes of the two prefixes exhibit different historical reflexes of the same Common Slavic root **per* (Vasmer 1971: v.3, 356): PERE- reflects pleophony, typical for East Slavic languages including Russian, while PRE- is the result of metathesis and vowel lengthening, typical for South Slavic languages including Old Bulgarian – the base of Old Church Slavonic (Tolstoy 2002: 82). Thus, the case of PERE- and PRE- refers to a non-trivial question of whether a borrowed prefix can become an allomorph of a native morpheme.

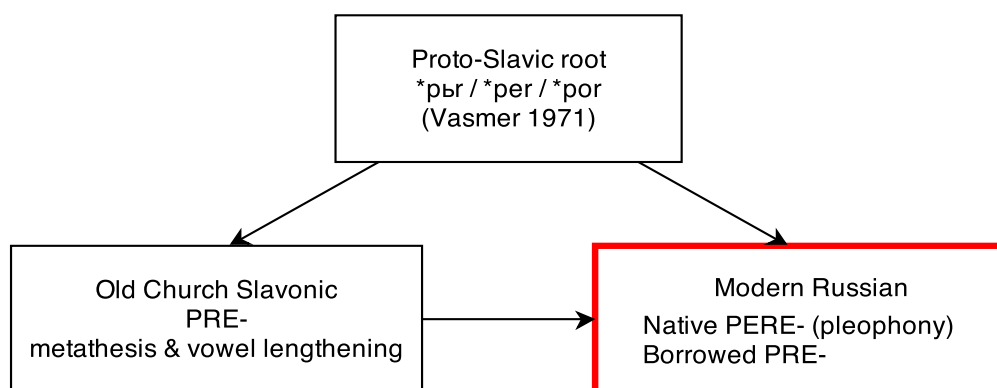


Figure 1: Historical origins of the prefixes PERE- and PRE-.

Interestingly, the same root with different quantities and qualities of the vowel (**pъr ~ *per ~ *por*) can be found not only in the prefixes PERE- and PRE- but also in the roots of the verb *sporit'* 'argue, dispute, debate' as well as in Slavonicisms *prekoslovit'* 'contradict', *pretit'* 'prohibit' and *zapretit'* 'ban', which are comparable with the native Russian verb *perečit'* 'contradict, go against' (Vasmer 1971: v.3, 358-359). Moreover, the same root **per* can be found in related nouns *prenije* 'dispute', *rasprja* 'quarrel', *spor* 'controversy, dispute', and *sopernik* 'rival', also borrowed from Church Slavonic into Russian (Vasmer 1971: v.3, 359, 362). The same root is also attested in the prepositions *vopreki* 'in spite of' and *poperjok* 'across' (Vasmer 1971: v.3, 238, 358). Although the semantic similarity of these related words is detectable from the modern perspective,

⁹¹ Like Russia, Ukraine and Belarus were converted to Orthodox Christianity in 988, and the language of the first manuscripts on these territories was Church Slavonic too (Comrie & Corbett 1993: 888-889). However, the annexation of Ukraine and Belarus to Poland and Lithuania and the long historical domination of Poland over centuries made a significant impact on the linguistic development of these East Slavic languages. As a result, today Ukrainian and Belarusian have fewer borrowings from Church Slavonic than Modern Russian has (de Bray 1969: 69). Also, while in Ukrainian and Belarusian some remaining terms from Church Slavonic were "naturalized", in Russian, on the contrary, they were preserved in their "original phonological form" (Sussex & Cubberley 2006: 478-9).

the correlations of the shapes of the root **per* clearly belong to diachrony. A question arises: are PERE- and PRE- associated synchronically? I address this issue in what follows.

6.1.2 Synchronic relation of PERE- and PRE-

In Modern Russian the distribution of PERE- and PRE- is not governed by any active automatic phonological rule. Yet, the two prefixes clearly preserve the formal similarity of their phonological shapes. It is not a trivial question whether PERE- and PRE- are related only diachronically or also synchronically. In other words, is the correlation between PERE- and PRE- synchronically perceptible for modern Russian speakers? This issue is complex and should preferably be addressed in a psycholinguistic experiment. Nevertheless, I suggest that the synchronic relation between PERE- and PRE- is present in Modern Russian. There are two arguments that I can propose in favor of this view.

First, one has to admit that a large part of the Modern Russian lexicon exhibits pairings of etymologically related doublets of East Slavic and South Slavic origin. In particular, there are many root morphemes that have both pleophonic (Russian, or East Slavic) and non-pleophonic (Slavonic, or South Slavic) variants: *bereg* ~ *breg* 'shore' (as in *pribrežnyj* 'near-shore'), *derevo* ~ *drevo* 'tree', *pered* ~ *pred* 'front, previous to' (as in *predok* 'ancestor'), *sered-* (as in *seredina* 'middle') ~ *sred-* (as in *srednij* 'average, middle'), etc. In addition, there are also similar alternating combinations of sounds in word pairs like *moloko* ~ *mleko* 'milk' (e.g. *mlečnyj put'* 'Milky Way') and *golova* 'head (body part)' ~ *glava* 'headman, chapter'. Itkin (2007: 270) lists thirty roots that have such alternating shapes and are motivated by the same historical process of Slavic liquid metathesis and pleophony⁹². Along with roots, Itkin also lists two prefixes: *čerez* ~ *črez* 'across, over' (as in *čezmernyj* 'excessive') and *pere* ~ *pre* 'over'. According to Itkin's description, all these root morphemes as well as the mentioned prefixes are variants paired with each other and related via a semi-regular morphophonological rule. This rule is a part of the Modern Russian grammar. It is worth mentioning that with regard to root morphemes like *derevo* ~ *drevo* 'tree' there is generally no doubt in the literature that the two variants belong to a single morpheme. In this light, the prefixes PERE- and PRE- can be grouped together with words that exhibit the same alternating phonological combinations, and therefore PERE- and PRE- should be associated with each other.

The second argument concerns the Russian orthographic rule that describes the spelling of the prefix PRE-. According to this rule, in case of doubt about the spelling of the prefix (PRE- or PRI-) in unstressed positions like *pr?-odolet'* 'overcome', *pr?-kratit'* 'stop', *pr?-kryt'* 'close slightly', one should try replacing the problematic prefix with PERE-: if PERE- works semantically (as in the two first cases: *pre-odolet'* 'overcome', *pre-kratit'* 'stop'), one should spell PRE- (and not PRI-). If the replacement of the problematic prefix with PERE- alters the meaning of the verb, as in *pr?-kryt'* (*perekryt'* 'block (e.g. a way)' ≠ *pri-kryt'* 'close slightly (e.g. a door)'), then one should spell the problematic unstressed prefix as PRI-. This orthographic rule is taught in Russian schools and is based on a semantic comparison between PERE- and PRE-. This spelling rule makes an explicit comparison of the two prefixes assuming that PERE- and PRE- share some semantic content and can replace each other in some contexts. In other words, Russian

⁹² This phenomenon is commonly referred to as the TOROT ~ TRAT alternation, where T stands for a consonant, R stands for a liquid consonant, and O stands for the vowels *o* and *e*.

orthography welcomes the association between PERE- and PRE- and assumes an explicit and conscious knowledge about their compatibility among speakers.

Summing up, there are two reasons to think that the two prefixes in question bear a synchronically real associative connection in the grammar of modern Russian speakers: 1) the opposition of PERE- and PRE- represents an alternation of pleophonic vs. non-pleophonic sound combinations which is otherwise widely attested in other morphemes; and 2) a common spelling rule makes a deliberate comparison of PRE- and PERE-.

6.1.3 On the similarity of formal shapes: PERE-, PRE-, PRED-, PRI-

In connection with the formal similarity of PERE- and PRE- discussed above, it is important to mention that there are at least two more prefixes in Modern Russian that have a similar phonological shape – PRED- and PRI-. What kind of relation do they have to PERE- and PRE-?

The prefix PRED- denotes ‘front’ and is clearly analyzable in the words like *predobedennyj* ‘before-dinner’ (< *obedennyj* ‘dinner’), *predystoričeskij* ‘prehistorical’ (< *istoričeskij* ‘historical’), *predšestvovat’* ‘forego, precede’ (< *šestvovat’* ‘walk’), and *predbannik* ‘dressing room in a bath-house, lobby’ (< *banja* ‘bath-house’). Historically, this prefix is related to the same root **per* that gave rise to PERE- and PRE- (Vasmer 1971, compare *pervyj* ‘first’ as originally meaning ‘walking in front of others’). Etymologically, PRED- contains the formant -d and is related to **per* in a similar way as the prepositions *po* and *pod*, *na* and *nad* are historically related to each other (Vasmer 1971). However, in Modern Russian the prefixes PRED- ‘front’ and PRE- ‘very’ can hardly be connected in terms of semantics. Rather, PRED- is comparable to another prefix PRE- which comes from Latin *prae-* ‘in front, in advance of’ and occurs in some loan words like *prefiks* ‘prefix’ and *prepozicija* ‘preposition’. I leave this second PRE- (< Lat. *prae-*) outside of my analysis and argue that the PRE- and PERE- that I examine are not synchronically related to PRED- despite the common diachronic origin and partial phonological resemblance with the latter.

The relation of PERE- and PRE- to the prefix PRI- is more complex. At first glance, PRI- ‘ARRIVE’ (Endresen et al. 2012: 262) is semantically not comparable with PERE- ‘over’ and PRE- ‘very’. Yet, at the periphery of their polysemies, it can be difficult to define how PRI- and PRE- differ in terms of their semantic contribution to a simplex verb. Compare the verbs *preumnožit’* and *priumnožit’*, both formed from the base *umnožit’* ‘multiply’. Corpus examples of their uses often seem almost identical:

- (1) *Ètot kapital nado ne tol’ko soxranit’, no i preumnožit’*. [B. Griščenko. *Postoronnij v Kremlje* (2004)]
‘This fund should be not only preserved but also **increased** (PRE-multiply).’
- (2) *Naš dolg – soxranit’ i priumnožit’ kul’turnoe nasledie*. [E. Savčenko. *S Rodinoj v serdce* (2004)]
‘Our duty is to preserve and **increase** (PRI-multiply) the cultural heritage.’

Native speakers often have difficulties choosing the proper spelling of the prefix in the verb *pr?-umnožit’*, and it is a challenge for professional linguists to describe the difference between the two possible alternatives. Indeed, dictionaries list both verbs *preumnožit’* and *priumnožit’* as equivalents with identical meaning (Rozental’ &

Telenkova 1999; Ožegov & Švedova 2001). The modern academic dictionary of Russian orthography (Lopatin 2007) suggests that these verbs are two unnecessary spelling variants of the same lexeme. Therefore, Lopatin 2007 eliminates *preumnožit'* and preserves only *priumnožit'*. Yet, Kuznecov (2000) points out a subtle distinction: PRE- 'multiply' denotes 'multiply a lot, several times', whereas PRI- 'multiply' means 'increase in addition to the existing amount'. In other words, we observe that *preumnožit'* features PRE- 'VERY', while the meaning of *priumnožit'* is supported by the PRI-'s sense 'ADD' (Endresen et al. 2012: 262). Because the base verb *umnožit'* 'multiply' itself implies a considerable increase in an original quantity, the difference between its PRE- and PRI- derivatives in this case is neutralized and seems elusive. Moreover, the subtle semantic difference between the two prefixed verbs is often not important in discourse (as in the sentences (1) and (2)), and leads to uncertainty in speakers' intuitions.

This observation makes a compelling case that even at the periphery of their polysemies, where PRE- and PRI- might be interchangeable, the two prefixes preserve their semantic differences and cannot be viewed as identical in their meanings despite their formal phonological resemblance⁹³.

The lack of semantic comparability between PRE- and PERE- on the one hand and PRED- and PRI- on the other hand suggests that the two latter prefixes constitute distinct morphemes synchronically unrelated to PERE- and PRE-. Summing up, an allomorphic relation with PRED- and PRI- is out of the question, despite the fact that PRED- and PRI- demonstrate very similar formal shapes to PRE- and PERE-.

6.1.4 Specificity of Russian among Slavic: Coexistence of PERE- and PRE-

Modern Russian preserves and makes use of both prefixes PERE- and PRE-, being different in this respect from Modern Ukrainian and Belarusian, where the prefix PRE- has almost completely disappeared from use (Lomtev 1961: 239; Žučenko 1969: 4). The corresponding prefix in Modern Belarusian is PERA- (compare *pera-plysci* 'swim across', *pera-spely* 'overripe'), while the Slavonic prefix PRE- is listed neither in the inventory of Belarusian verbal prefixes, nor in the inventory of Belarusian adjectival prefixes of 'high degree' like NAI-, ZA-, UL'TRA-, ARXI-, etc. (Adamovič et al. 1980: 15, 29, 94-96; de Bray 1969: 166-7). In Modern Ukrainian, the prefix PRE- is available only for adjectives and is marked as rare (e.g. *preveselij* 'very happy', *prečudovij* 'marvelous' (Pugh & Press 1999: 163; de Bray 1969: 102). This makes the Russian data especially interesting.

Modern Russian not only preserves both prefixes, but also uses them productively, each of them in its own domain. Since PERE- and PRE- coexist in Modern Russian, the question arises of what their status is.

6.1.5 State of the art: Previous accounts

Although both PERE- and PRE- play a significant role in word-formation of Modern Russian, they received dramatically different amounts of attention in the scholarly literature. The native prefix PERE- is well described (Flier 1985; Janda 1986: 134-173; Dobrušina et al. 2001: 76-80; Shull 2003: 113 – 119; Tatevosov 2008; Kagan 2011, 2013), whereas the borrowed PRE-, on the contrary, has been largely neglected (Soudakoff 1975: 231) and mentioned only as a Church Slavonic equivalent of PERE-

⁹³ This conclusion can be supported by the pair of verbs *preumen'shit'* (PRE-'belittle') 'belittle largely' and *priumnen'shit'* (PRI-'belittle') 'belittle slightly', where the two prefixes can be contrasted according to their distinct semantic contribution to the base *umen'shit'* 'belittle'.

(Townsend 1968: 59; 128). One reason for neglecting the prefix PRE- is its “vagueness of meaning”: Isačenko (2003/1965: v.2, 149-150) argues that most verbs prefixed with PRE- are deetymologized. Another reason is that most studies of the PERE-/PRE-correlation focus on the verbal domain of their use, where PRE- is indeed rare and unproductive (Spagis 1968: 6; Švedova et al. 1980: §873; Zemskaja 2006).

In this chapter I show that beyond verbs PRE- is a highly productive prefix that intensifies the semantics of adjectives and adverbs. As for the verbal domain, my analysis suggests that the semantics of PRE- is not always vague and is often compatible with that of the prefix PERE-.

The historical development of Russian and Church Slavonic verbs prefixed in PERE- and PRE- is addressed in Žučenko 1969 and Uluxanov 1974, 2004.

Regarding the status of these prefixes in Modern Russian, in the scholarly literature it is mostly assumed that PERE- and PRE- are two distinct morphemes, although the issue has been largely overlooked. To the best of my knowledge, Soudakoff 1975 is the only contrastive study that has addressed the question of the contemporary relationship between PERE- and PRE- in Modern Russian. Soudakoff argues that PERE- and PRE- are two distinct morphemes. This account is based on a limited amount of data: it is particularly focused on minimal pairs of verbs that differ in their prefix, like *pere-seč'* ‘cross’ vs. *pre-seč'* ‘stop’. This exclusive attention on the contrastive use of these prefixes in minimal pairs is arguably motivated by the Structuralist approach adopted by the scholar. Note that one of the core assumptions of the Structuralist framework is that the meaning of a linguistic sign is exactly the contrast that occurs in minimal pairs or similar oppositions of signs. Soudakoff aims to describe the semantics of prefixes in terms of contrastive features. The features assigned to PERE- and PRE- within the pairs of contrastive verbs are then extrapolated further to the non-contrastive uses of these prefixes, whereas a large amount of relevant data remains neglected. Crucially, Soudakoff 1975 reveals not only differences, but also strong similarities in the semantics of PERE- and PRE-, which undermines the central argument in favor of two distinct morphemes.

As opposed to Soudakoff 1975, I aim to account for a more comprehensive data set for these prefixes, including verbs that are not listed in reference works and dictionaries, but attested in the Russian National Corpus. In this chapter I present the first corpus-based study of PERE- and PRE- where the question of their morphological relations is revisited.

6.1.6 Two possible solutions and their problems

The goal of this chapter is to answer the question as to whether the two prefixes under consideration belong to a single morpheme or two distinct morphemes in Modern Russian. There are two possible solutions: the allomorphy hypothesis and the two morphemes hypothesis. Which of the two hypotheses can capture the non-trivial data best and misrepresent the data least? At first glance, it appears that both models encounter resistance in the data.

The two-morphemes-model fails to account for the strong semantic similarity of PERE- and PRE-. If the two prefixes are distinct morphemes, we fail to explain examples where these prefixes attach to the same base and provide a very similar semantic contribution. Compare the verbs *pere-dat'* (PERE-‘give’) ‘hand over’ and *pre-dat'* (PRE-‘give’) ‘commit to; betray’ both of which are formed from the same base verb *dat'* ‘give’:

- (3) *Peredaj mne sol', požalujsta, – skazala Marina.* [A. Gelasimov. God obmana. (2003)] '**Send** me the salt, please, Marina said'.
- (4) *Ja <...> toržestvenno predam tetrad' ognju.* [Je. Proškin. Mexanika večnosti (2001)] 'I will solemnly **commit** the exercise-book to flames'.

It is worth mentioning that although PERE- and PRE- contribute the same spatial sense 'TRANSFER' to the base *dat* 'give', they are not interchangeable: their distribution in the contexts (3) and (4) is not free, but conditioned by stylistic register and types of verbal arguments.

Another pair of verbs that demonstrates very similar semantic content for the two prefixes consists of the verbs *pere-gorodit'* (PERE-'fence') 'barrier' and *pre-gradit'* (PRE-'fence') 'obstruct, block', where the root *gorod* ~ *grad* 'town, enclosed place' appears in its Russian and Slavonic shapes accordingly. Corpus examples provided in (5) and (6) show that PERE- and PRE-derivatives can even be used in the same context describing a person who blocks someone's way:

- (5) *Rastopyriv ruki s kljuškoj i kostylem, on postaralsja peregorodit' mne dorogu.* [V. Zapašnyj. Risk. Bor'ba. Ljubov'. (1998-2004)]
'Having spread his arms with a hockey-stick and a crutch, he tried to **block** off my way.'
- (6) *Administrator dognal ego liš' na lestnice <...> i popytalsja pregradit' put' k dveri.* [N. Podol'skij. Kniga Legiona. // «Oktjabr'», 2001]
'The manager caught him at the staircase and tried to **block** his way to the door.'

On the other hand, such pairs of verbs like in (3)-(4) and in (5)-(6) demonstrate that there are simplex bases that can attach both prefixes, suggesting that the distribution of PERE- and PRE- across verbs is not complementary. This overlap in use violates the distributional criterion for a standard allomorphic relation. In addition, we can observe that in some contrastive pairs of verbs the semantic contribution of PERE- and PRE- to a simplex stem can be rather different. Compare the verbs *pere-terpet'* (PERE-'endure') 'endure, overcome suffering' and *pre-terpet'* (PRE-'endure') 'undergo suffering or a change', both formed from the base *terpet'* 'endure, suffer'. The non-identity in semantics violates the second, semantic, criterion for a standard allomorphic relation. Therefore, overall, we have a non-trivial case, where neither of the two criteria of allomorphy is perfectly satisfied. At the same time, the two prefixes are more than just competing synonyms: they are diachronically and synchronically related and preserve similar phonological shape as well.

Summing up, both accounts – in terms of distinct morphemes and in terms of standard allomorphic relations – for these prefixes are problematic. However, I suggest that if we overcome the traditional understanding of allomorphy as a clear-cut categorical phenomenon, we discover that these affixes are much more plausible as allomorphs or near-allomorphs than as two separate morphemes.

6.1.7 Goal of the present study and overview

In this study I explore whether PERE- and PRE- can be described as Non-Standard Allomorphs and thus enrich our knowledge about the variability of this phenomenon. I

focus on factors that condition the use of these prefixes in Modern Russian – their semantic and grammatical properties.

The remainder of this chapter is organized as follows. I compare the use of PERE- and PRE- in two domains – verbs (sections 6.2-6.5) and beyond verbs, that is in adjectives, adverbs, and nouns (section 6.6). In section 6.2, I introduce the verbal data and explain how it was extracted from the corpus and tagged. Then I adopt a cognitive linguistic approach and analyze the semantics of the two prefixes in terms of radial categories. Both PERE- and PRE- are very diverse in their semantics. So, I investigate whether they can both be described via a single semantic model of interrelated submeanings motivated by the same spatial prototype. In section 6.3, I show that such a model is possible: I present a network of fourteen senses, in eight of which PERE- and PRE- overlap. The large degree of semantic overlap suggests that these affixes might represent a single morpheme. Next, I apply the methodology of Radial Category Profiling and compare the two prefixes quantitatively, in terms of what submeanings they are most frequently attested in across verbs (section 6.4). I further subject their distributional differences to a statistical analysis. Then I turn to grammatical properties of PERE- and PRE- in section 6.5 and look at how frequently these prefixes affect the verbal categories of aspect and transitivity. A statistical analysis makes it possible to measure the degree of grammatical similarity and divergence in terms of percent of individual verbal lexemes where aspect and transitivity are affected by the prefix. In 6.6, I look at the data beyond verbs which demonstrates the prefix PRE- as a productive intensifier of quality. In 6.7, I formulate my proposal and in 6.8 I support my argument by comparing PERE- and PRE- to intensifiers of other kinds and in other languages than Russian. Conclusions are offered in 6.9.

6.2 Data

The data for this study was collected from the Russian National Corpus (www.ruscorpora.ru; henceforth RNC). I compiled a database that contains a total of 945 verbs, including 54 verbs prefixed in PRE- and 891 verbs prefixed in PERE-. In this subsection I explain the principles of how the data was culled and arranged.

I started out with a list of all verbs that contain initial sequences *pere* and *pre* regardless of their derivational structure. I extracted this data automatically by means of MySQL software from the word-list of the frequency dictionary Lyashevskaya & Šaroff 2009. This dictionary is based on the Modern Subcorpus of the RNC – a digital collection of texts created in 1950 – 2009 and containing 98 million words. The automatically extracted dataset was analyzed manually. When all typos and irrelevant items were weeded out, I obtained 1,836 verbs with analyzable prefixes PERE- (1,729 verbs) and PRE- (107 verbs).

However, this list contained duplicate datapoints – namely imperfective verbs replicating their perfective counterparts (e.g. *pereodet'*_{PF} – *pereodevat'*_{IPF} ‘change clothes on someone’), and reflexive verbs⁹⁴ with the intransitivizing postfix *-sja* (e.g. *pereodet'*_{PF} ‘change clothes on someone’ – *pereodet'sja*_{PF} ‘change clothes on oneself’). In order to avoid duplicate items, I focused on perfective verbal lexemes created by the attachment of a prefix. I merged paired perfective and secondary imperfective verbs into a single entry, assuming that the two aspectual forms represent a single verbal lexeme. However,

⁹⁴ I use the term *reflexive* in the broad sense to refer to all intransitive verbs in *-sja*, despite the fact that this postfix can carry different meanings.

some imperfective verbs occupy separate entries in the database: if they lack an attested perfective counterpart (e.g. *presledovat'* 'pursue', *perefutbolivat'* 'throw to a different place', *perešeptyvatsja* 'exchange whispers'). Similarly, reflexive verbs with *-sja* in a bare intransitivizing function were placed in the same entry with their non-reflexive counterparts: e.g. *pereodet'* 'change clothes on someone' – *pereodet'sja* 'change clothes on oneself' were merged into a single lexeme *pereodet'(sja)*. Exceptions include those cases when a) a non-reflexive counterpart is not attested (e.g. *pereserdit'sja* 'overcome one's anger') or b) *-sja* contributes a reciprocal semantics (e.g. *peregljanut'sja* 'exchange glances'). Separate entries represent imperfective reflexive verbs like *perepisyvat'sja* 'write to one another' if their perfective counterparts cannot express reciprocal semantics: compare *perepisat'(sja)* 're-write; copy'. Thus, a perfective verb with a parenthesized (*sja*) like *peregruppirovat'(sja)* 're-group' can result from a merger of four forms: *peregruppirovat'*_{PF.TR}, *peregruppirovyvat'*_{IPF.TR}, *peregruppirovat'sja*_{PF.INTR}, and *peregruppirovyvat'sja*_{IPF.INTR}.

Each verb in the database was manually assigned a simplex base and tagged for the submeaning of the prefix. In order to account for different meanings of polysemous verbs, I analyzed corpus examples and consulted relevant dictionaries.

If different meanings of a verb result from different submeanings of the prefix, they are represented in the database as separate datapoints. For example, the verb *perevarit'* 'PERE-cook' can be used in Modern Russian in three different senses that are motivated by different submeanings of PERE-: 'cook again' (submeaning 4.REDO), 'overcook' (submeaning 3.OVERDO), and 'digest' (Submeaning 5.TRANSFORM). Because all three senses of this verb exhibit different submeanings of PERE-, they are assigned different semantic tags and treated as different verbs, each assigned a single entry. Another example of polysemy is the marginal verb *pereljubit'* 'PERE-love'. Different contexts highlight different senses of PERE- contributed to the simplex base *ljubit'* 'love': 4.OVERDO - 'love too much' (7), 6.OVERCOME / DURATION - 'love during a period of time and stop' (8), and 12.DISTRIBUTE - 'love many people' (9).

- (7) *Možno bylo by skazat' – nedoljublennaja, no tut skoree drugoe – glupo pereljublennaja.* [E. Piščikova. Pjatietažnaja Rossija (2007)]
'One could say – she is under-loved, but this is another situation – **over-loved** in a silly way.'
- (8) *Ljublju ego – vot v čem vopros. Pereljublju – razvedus', čestnoe slovo.* [I. Grekova. Perelom. (1987)]
'I love him – this is my problem. When I **overcome love / stop loving** – I will divorce, I promise.'
- (9) *<...> ne vsë v žizni on ispytal, ne vse zapaxi perenjuxal, ne vsež ženščin pereljubil.* [N.N. Berberova. Kursiv moj (1960-1966)]
'<...> he has not experienced everything in life, not smelled all odors, not **loved** all women.'

In the database, I provide an illustrative example from the corpus like the ones in (7), (8), and (9) for each type of prefix use for each verb.

If different meanings of a prefixed derivative verb result from homonymous bases, I count them as separate lexemes: *peretopit'* 'make many objects sink' (< *topit'*

'sink') with PERE- 'DISTRIBUTE' vs. *peretopit* 'melt something into a different consistency or quality' (< *topit* 'melt') with PERE- 'TRANSFORM'.

Submeanings are assigned to PERE- and PRE- according to the model of polysemy described in Section 6.3. Each semantic tag is supported with an illustrative example from the corpus. This is particularly useful for polysemous verbs, where the context resolves potential semantic ambiguity and highlights a particular reading of a prefix-verb construction. Illustrative examples are also helpful in the case of many marginal low-frequency verbs represented in the database, because their semantics is often unclear without a context.

Each verb in the database that has a verbal simplex base is tagged for additional parameters: aspectual shift, shift in transitivity, and prefix stacking. Comparing verbal bases with prefixed derivatives in terms of aspect and transitivity, we can observe whether PERE- and PRE- differ in the ability to affect properties of a verb. Moreover, marking the cases when PERE- and PRE- are stacked over another prefix I expect that prefix stacking would correlate significantly with particular submeanings of these prefixes and with aspectual shift or transitivity shift.

Each verb is accompanied by its token frequency – the number of its attestations in the Modern Subcorpus of the RNC. I take into account texts created in 1950-2014. Token frequencies of verbs were obtained manually from searches conducted in March 2014. I use these token frequencies of prefixed verbs in order to distinguish between standard lexemes and marginal occasional coinages. Standard lexemes belong to the Modern Russian lexicon shared by most speakers, and I address them in my quantitative analysis. Marginal low-frequency verbs are analyzed separately, because they provide information of a different kind – they indicate productivity of specific prefixal submeanings and show the linguistic potential of the system. Based on examination of the data, I came to the decision that the threshold for standard vs. marginal lexemes that works for this dataset of verbs prefixed in PERE- and PRE- is 5 attestations in the Modern Subcorpus of the RNC.⁹⁵ This means that all verbs that are attested in the corpus more than 5 times are counted as standard lexemes, whereas those verbs that have 5 or fewer attestations in the corpus are considered marginal. The token frequency threshold is the same for verbs prefixed in PERE- and PRE-.

As a result, in my analysis I account for a total of 476 standard verbal lexemes, including 437 predicates prefixed in PERE- and 39 predicates prefixed in PRE-. The numbers are summarized in Table 1.

⁹⁵ Note that the threshold for standard and marginal verbs is different for different prefixes. This is because in this dissertation I choose to make the threshold prefix-specific rather than formal and arbitrary. The threshold for standard and marginal verbs is established anew for each particular dataset. In each case study, I choose a threshold that captures best the distinction between standard and marginal verbs containing a particular prefix. Recall that for verbs prefixed in S- and SO- the threshold is 10, meaning that verbs with 10 and more corpus attestations are counted as standard, whereas those that have less than 10 attestations are considered marginal. For PERE- and PRE-, this threshold would have been too high. Instead, in this study I take verbs that have over five corpus hits as standard verbs. The threshold of 10 attestations would have led to exclusion of 42 verbs in PERE- that are standard but for some reason are rarely attested in the corpus (arguably because some of them have colloquial flavor, as in *perešerstit* 'search for something item by item' and *perexotet* 'stop wanting'). Another possible explanation might be that verbs with different prefixes might differ in their overall frequency rates (for example, verbs in PRE- might be in general less frequent in use than verbs in S-). However, this issue is beyond the scope of this study.

Word Type	Token Frequency	Verbs in PERE-	Verbs in PRE-	Total
Standard	> 5	437 (92%)	39 (8%)	476 (100%)
Marginal	≤ 5	454 (97%)	15 (3%)	469 (100%)
Total		891 (94%)	54 (6%)	945 (100%)

Table 1: Standard and marginal verbs prefixed in PERE- and PRE-.

As shown in Table 1, marginal verbs prefixed in PERE- yield an even larger part of the dataset – a total of 454 lexemes. There are less verbs prefixed with PRE- among marginal lexemes (only 3% of all marginal verbs in the database) than among standard lexemes (8% of all standard verbs in the database).

This peculiarity of the distribution is motivated by the fact that the marginality of verbs in PERE- and the marginality of verbs in PRE- are of different kinds. Marginal verbs in PERE- are mostly occasional new coinages created on the basis of productive semantic uses of this prefix, such as ‘REDO’, ‘OVERDO’, ‘SUPERIORITY’, ‘DISTRIBUTE’, and ‘INTERCHANGE’: e.g. ‘REDO’ in *krasit’* ‘paint’ – *perekrasit’* ‘paint anew’ >> *malevat’* ‘paint’ – *peremalevat’* ‘paint anew’, *opublikovat’* ‘publish’ – *pereopublikovat’* ‘publish anew’, etc. By contrast, marginal verbs in PRE- are mostly obsolete lexemes that went out of Modern Russian use: *preukrasit’* ‘decorate a lot’ (< *ukrasit’* ‘decorate’), *prestaret’* ‘become old’ (< *staret’* ‘grow old’), *preslušat’* ‘disobey’ (< *slušat’* ‘listen to, obey’), *preselit’sja* ‘move to a different location’ (< *selit’sja* ‘settle’), *prepožalovat’* ‘arrive’ (< *požalovat’* ‘visit, grant’), *prepinat’sja* ‘argue with’ (< *pinat’* ‘kick’), *preosuščestvljat’* ‘transform into something new’ (< *osuščestvljat’* ‘carry out, arrange’), *preobratit’sja* ‘transform into’ (< *obratit’sja* ‘turn’), *premenit’* ‘change’ (< *menjat’* ‘change’), *prevyprovaživat’* ‘see off the premises towards another location’ (< *vyprovaživat’* ‘door’), and *prevodit’* ‘transfer to’ (< *vodit’* ‘lead’). Marginal verbs in PRE- point to old uses of this prefix, they are few and unproductive. Marginal verbs in PERE-, on the contrary, indicate the direction where the linguistic system is going: highly productive patterns of verbal prefixation supported by numerous corpus examples. The productivity of particular morphological and semantic patterns is the main reason for the abundance of marginal verbs prefixed in PERE- (97% of all marginal verbs) in the corpus-based dataset.

6.3 Semantic Analysis: Radial network of submeanings

In this section I present a contrastive semantic analysis of PERE- and PRE- in prefixed verbs. I adopt a Cognitive Linguistic approach and model polysemy as a network of interrelated submeanings organized around a central spatial prototype (Lakoff & Johnson 1980).

This study is inspired by the cognitive analysis of the Russian prefix PERE- proposed by Janda (1986: 134-173). The key advantage of Janda’s account of PERE- is that it goes beyond an incoherent inventorization of prefix sub-uses. Instead of making a list of submeanings, Janda establishes a semantic model of polysemy which is organized into a radial structure. In this semantic network, abstract senses of the prefix are motivated by a spatial prototype, and each submeaning entails a spatial configuration of the Trajector and the Landmark. Crucially, schematic graphical representations proposed by Janda for semantic subcategories make it possible to explain how various submeanings of PERE- are related to each other via cognitive mechanisms. With a few changes, the semantic model of PERE- proposed by Janda has been employed in modern corpus studies of Russian aspectual prefixes (Endresen et al. 2013) and is supported by the seminal analysis of PERE-’s semantics in terms of the Scalar Hypothesis (Kagan

2013). Because Janda's inventory of PERE-'s submeanings is very thorough and detailed, it suits perfectly our purpose to compare PERE- and PRE- in terms of semantics. Therefore, in this chapter I adopt this model with only a few changes and map the polysemy of PRE- to the network of PERE-.

The central question that I address in this analysis is how the loan prefix PRE- is similar to and different from its native East Slavic equivalent PERE- in the verbal domain. Is the semantic content of PRE- as vague and elusive as it is portrayed (Isačenko 2003/1965: 149), or is it detectable and analyzable? And if so, is PRE- semantically comparable with PERE- in terms of the same model of polysemy? I suggest that only two of 39 standard verbs prefixed in PRE- are synchronically deetymologized, i.e. their motivational link between the base and the prefixed verb is lost: *presmykat'sja* 'creep' and *preminut'* 'fail to'. Although it might be still possible to connect the former verb with its simplex *smykat'sja* 'interlock, come close', and the latter verb with *minut'* 'pass (of time)', the exact semantic contribution of the prefix PRE- is hard to define. I suggest that in the remaining 37 verbs it is possible to define the meaning of PRE- in terms of the subcategories that are attested for PERE-.

My analysis brings me to the conclusion that the rich polysemies of both prefixes can be described via a single model – a radial network of submeanings shown in Figure 2. Each submeaning of the two prefixes is represented here as a rectangular box which indicates the name of the subcategory, shows how many individual verbal lexemes represent this submeaning of the prefix in the Russian lexicon (see the numbers in parentheses), and provides an illustrative example with an English gloss. The prototypical submeaning is placed in the center of the network. The lines between rectangular boxes represent semantic links between submeanings.

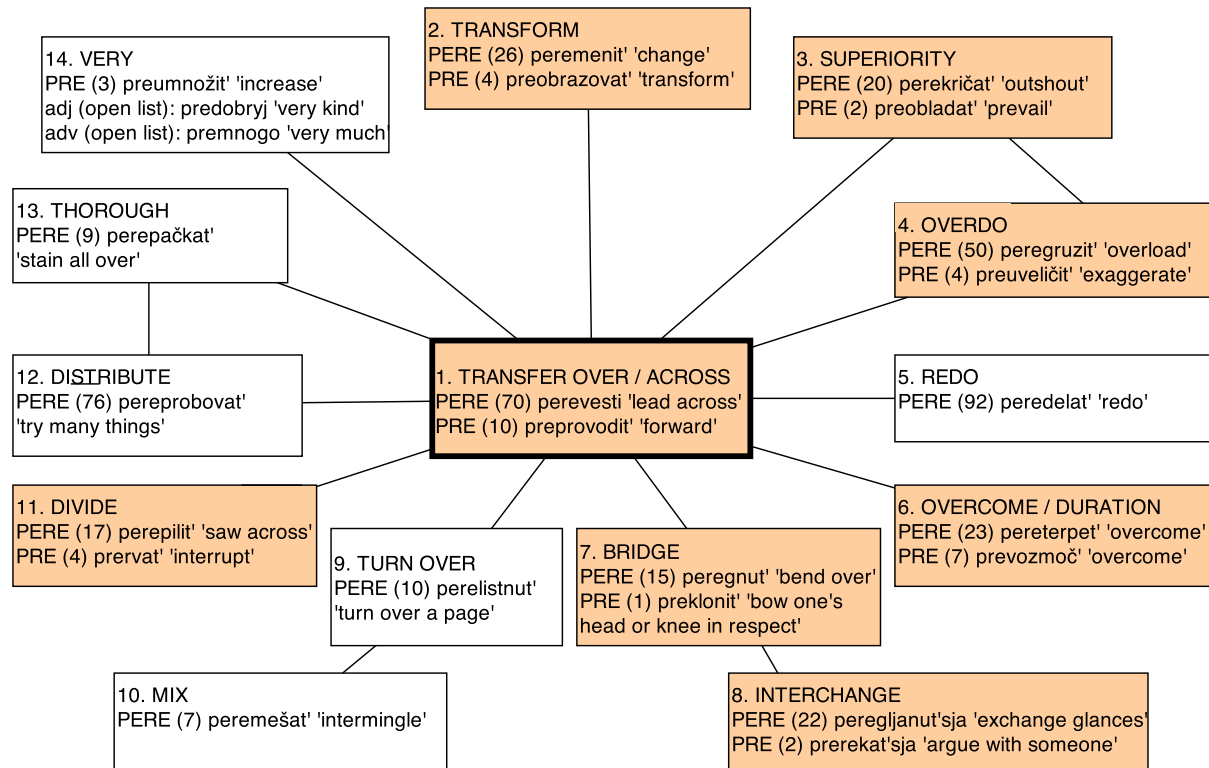


Figure 2: Radial category model for the prefixes PERE- and PRE- (numbers of standard lexemes in parentheses).

This diagram demonstrates that the prefix PRE- occupies a subset of the semantic network attested for the prefix PERE-. In particular, PERE- can refer to a total of thirteen submeanings, while PRE- is attested for eight of them. The shading in Figure 2 highlights those submeanings that are attested for both prefixes. Thus, eight submeanings (more than half) constitute a zone of semantic overlap for PERE- and PRE-. In the remaining six submeanings the two prefixes differ from each other. Five submeanings of PERE- are not attested for PRE- (4.'REDO', 9.'TURN OVER', 10.'MIX', 12.'DISTRIBUTE', and 13.'THOROUGH'), and one submeaning of PRE- (14.'VERY') is not attested for PERE-. I will now address each submeaning in turn.

Subcategory 1.TRANSFER OVER/ACROSS

The radial network of PERE-'s polysemy is organized around the central subcategory 1.TRANSFER OVER/ACROSS (Janda 1986: 138; Endresen et al. 2013: 273; comparable with 'cross' in Kagan 2013: 494). This meaning serves as a semantic prototype, which motivates all submeanings in the network, directly or indirectly. In this submeaning, the prefix PERE- refers to the physical transportation of an object (Trajector) to a new location. I suggest that this concrete spatial meaning can be visualized by means of three related image schemas shown in Figure 3:

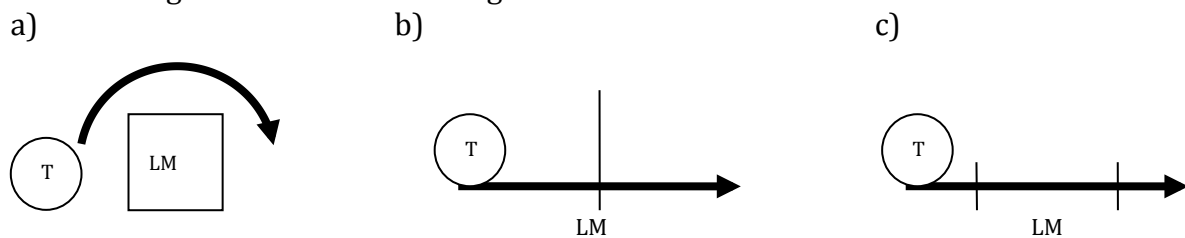


Figure 3: Variations of the image schema of the prototypical subcategory 1.TRANSFER OVER/ACROSS.

Figure 3 depicts three variations of the same image schema, where the Trajector (T) moves from one place to another place proceeding over (3a), through (3b), or across (3c) a physical Landmark (LM). The three variations can be illustrated by the verbs of motion: a) *perelezt' čerez zabor* 'climb over a fence'; b) *perenesti čerez porog* 'carry across a threshold'; c) *perejti ulicu* 'cross the street'. The image schema in (3a) implies a three-dimensional space as opposed to a one-dimensional schema in (3c). Janda (1986: 167) makes a distinction between (3a) and (3c) and describes them as different configurations. However, Janda mentions that "the only difference between <over> and <across> lies in the fact that <over>'s landmark is a barrier rather than just a space" (Janda 1986: 167). Brugman and Lakoff (1988: 113) in their analysis of the English *over*, which is semantically comparable with the Russian PERE- (cf. Janda 1986: 134; Kagan 2013: 513), combine these two types of trajectories – ABOVE and ACROSS – in the central sense of *over*. In this light, I consider all three schemas in Figure 3 as variations of a single schema. Indeed, the three schemas differ only in terms of the type of Landmark: in (3a) the Landmark is a prominent solid physical object, in (3b) the Landmark is a physical boundary, whereas in (3c) the Landmark is an intervening distance stretched out in space. Crucially, in prefixed Russian verbs the type of Landmark is encoded in the verbal stem and its syntactic arguments rather than in the prefix PERE- itself. I suggest that the prefix in its turn encodes the path, or trajectory, that transgresses, or

“overcomes”, an obstacle, but the shape and the type of the obstacle is not necessarily specified by the prefix.

The label TRANSFER OVER / ACROSS highlights the idea of transportation that a Trajector undergoes, the change of its initial location, and embraces the relevant senses ‘OVER’ and ‘ACROSS’. The same verbs of motion like *pereletet’* ‘PERE-fly’, *perepolzti* ‘PERE-crawl’, *perevesti* ‘PERE-lead’ can refer to both senses ‘OVER’ and ‘ACROSS’ depending on the lexical filler of the argument position in the verbal phrase: e.g. *pereletet’ čerez zabor* ‘fly over the fence’ (‘OVER’) vs. *pereletet’ okean* ‘cross the ocean by flying’ (‘ACROSS’), *perevesti čerez most* ‘take over the bridge’ (‘OVER’) vs. *perevesti čerez dorogu* ‘help someone to cross the street’ (‘ACROSS’).

In some prefixed verbs the focus of attention is shifted from overcoming an obstacle to a simple change of physical location. For example, in *perevesit’ kartinu* ‘hang a painting to a different place’ the idea of overcoming an obstacle is completely neutralized and overpowered by the idea of simple transportation, spatial relocation. This can be visualized as in Figure 4. This image schema should be regarded as an additional possible variation of the same schema for subcategory 1.TRANSFER OVER / ACROSS.

d)

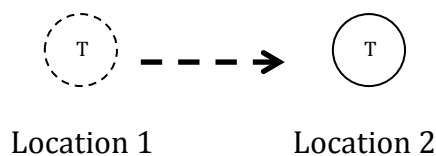


Figure 4: Image schema variation of the prototypical subcategory 1.TRANSFER OVER / ACROSS.

Such a neutralization of the obstacle can be observed in the semantics of other verbs of position too: compare *pereložit’* ‘lay to a different place’ (< *ložit’*, *klast’* ‘lay’), *peresadit’* ‘change a seat’ (< *sažat’* ‘sit’) and their intransitive counterparts *pereleč’* ‘lie down to another place’ (< *leč’* ‘lie down’) and *peresest’* ‘change a place for sitting’ (< *sest’* ‘sit down’). The same idea of a bare transportation (relocation) regardless of ‘OVER’ and ‘ACROSS’ in the semantics of the prefix is present in the umbrella terms *peremestit’* ‘transport to a different location’ (< *mesto* ‘place’) and *perebazirovat’* ‘locate to a different place’ (< *bazirovat’* ‘locate’).

Moreover, the prefix PERE- expresses the bare submeaning ‘TRANSFER’ (under the neutralization of senses ‘OVER’ and ‘ACROSS’) when it attaches to verbal bases that emphasize a change of location like *slat’* ‘send’, *dat’* ‘give’, *dvinut’* ‘move’, *sypat’* ‘strew’, *lit’* ‘pour’, *teč’* ‘flow’, *kačat’* ‘pump (oil); download’. Compare the corresponding prefixed derivatives: *pereslat’* ‘send to a different place’, *peredat’* ‘forward to someone’, *peredvinut’* ‘move to a different place’, *peresypat’* ‘strew from one place to another’, *perelit’* ‘pour to a different container’, *pereteč’* ‘flow to another place’, and *perekačat’* ‘transfer oil; download from one place to another’. The image schema behind these verbs is the one in Figure 4: a bare transfer of the Trajector from Location 1 to Location 2.

Similarly, the prefix contributes the bare submeaning ‘TRANSFER’ to those simplex predicates that denote a localization of an object by means of a certain activity like *kleit’* ‘glue’, *selit’* ‘house’, *gruzit’* ‘load’. In the prefixed derivatives of these verbs the spatial contribution of the prefix is clearly analyzable, and an obstacle which should be overcome by the Trajector is absent: *perekleit’* ‘glue from one place to another’, *pereselit’*

'house into a different place', *peregruzit'* 'trans-ship, load from one place to another (e.g. from a train to a ship)').

There is one more group of verbs that is indifferent to physical obstacles (Figure 3) and illustrative of the bare 'TRANSFER' sense (Figure 4) of the prefix PERE-. These are verbs that refer to copying an image from one place to another like *pererisovat'* 'copy by drawing' (< *risovat'* 'draw'), *perepisat'* 'copy by writing (e.g. from a textbook to a notebook)' (< *pisat'* 'write'), *peredrat'* '(colloquial) copy, imitate' (< *drat'* 'tear'), etc. Thanks to the prefix, the act of copying an image that these verbs refer to is construed as a spatial transportation of a physical object to a new location. Although the image remains in the original location after it was copied, the image nevertheless becomes transferred in space:

- (10) *Sverxu položite list prozračnoj bumagi i pererisujte na nee izobraženie s fotografii.*
[A. Kalinin. Iščite ženščinu // «Nauka i žizn'», 2009]
'Put a sheet of transparent paper on top and **copy** the image from the photograph **by drawing.**'

To sum up, there are many bases that neutralize the senses 'OVER' and 'ACROSS' that PERE- can express, and so the semantic contribution of the prefix often does not imply overcoming an obstacle but rather implies a bare change of a physical location. However, in my view, this fact does not undermine the label 1.TRANSFER OVER/ACROSS for naming the central prototypical subcategory of the prefix PERE-.

While various verbs of position and copying prefixed in PERE- lack a syntactic argument that would lexicalize an obstacle, many other verbs behave differently. Some simplex predicates encode actions that imply a vertical dimension of movement, and with the prefix, the semantics of the base arguably activates PERE-'s sense 'OVER' (Figure 3a) in addition to 'TRANSFER'. We observe the sense 'OVER' in the synonymous verbs *perebrosit'* and *perekinut'* 'throw to another place over smth' (< *brosit'*, *kinut'* 'throw'), as well as in *perelezt'* 'climb over' (< *lezt'* 'climb') and *peremaxnut'* 'jump over smth' (< *maxnut'* 'wave, skip over'). Note that these verbs normally require the prepositional phrase *čerez* + *noun* 'over something' which lexicalizes an obstacle (Landmark) that is overcome by the Trajector. An illustrative example is given in (11):

- (11) *Polkovnik ustojal. <...> Sxvatil velosiped, podnjaj i perebrosil čerez ogradu v kusty.*
[S. Jurskij. Na dačax. (1974-1983)]
'The colonel held his ground. He grabbed the bicycle, lifted it and **threw** it **over** the fence into the bushes.'

Note that other verbs, especially verbs of motion that refer to movement in a horizontal dimension, actualize the sense 'ACROSS' (in addition to 'TRANSFER') of the prefix PERE-: e.g. *pereplyt'* 'swim across' (< *plyt'* 'swim'), *perepolzti* 'crawl across' (< *polzti* 'crawl'), *perebežat'* 'run across' (< *bežat'* 'run'). The same preposition *čerez* can be used with these verbs, but it means 'across', and the Landmark it refers to is a distance (Figure 3c) or a boundary (Figure 3b). Moreover, with verbs of horizontal movement the preposition *čerez* is often optional, and the Landmark is frequently indicated by a direct object: *perejti / perebežat' (čerez) dorogu*_{ACC.SG} 'cross the street by walking / running', *perešagnut' / perestupit' (čerez) porog*_{ACC.SG} 'cross the threshold'. It is in this submeaning that the prefix PERE- can turn an intransitive verbal simplex into a transitive prefixed derivative: *plyt'*_{INTR} 'swim' > *pereplyt'*_{TR} 'cross X (direct object) by swimming', *polzti*_{INTR}

'crawl' > *perepolzti*_{TR} 'cross X (direct object) by crawling', *bežat'*_{INTR} 'run' > *perebežat'*_{TR} 'cross X (direct object) by running'. An illustrative example of PERE- with the actualized sense 'ACROSS' is given in (12), where the Landmark is explicitly named by the direct object and refers to a space – the town and the potato field:

- (12) *No slon ušel iz zooparka noč'ju, bystro **peresek** gorod, **perebežal** kartofel'noe pole i skrylsja v lesu.* [S. Kozlov. Pravda, my budem vseгда? (1969-1981)]
 'However, the elephant left the zoo at night, quickly **went across** the town, **ran across** a potato field, and disappeared in the forest.'

Summing up, the senses 'OVER' and 'ACROSS' are closely related and possible for the prefix PERE-. The actualization of either of these senses depends on the semantics of the verbal base and its argument structure. Some verbal predicates support both senses, others support only one of them. Moreover, I have shown that some verbal bases with spatial semantics support neither 'OVER' nor 'ACROSS' and instead foreground PERE-'s meaning of bare transportation. These observations lead me to conclude that the central meaning of PERE- should be labeled TRANSFER OVER/ACROSS, which potentially embraces all three senses.

The submeaning 1.TRANSFER OVER/ACROSS of the prefix PERE- is attested in 70 standard verbs and 40 marginal verbs in the database. These lexemes include motion verbs and also other verbs that refer to a concrete transfer of a physical object in its physical environment.

The spatial submeaning 1.TRANSFER can be contributed by the prefix PRE- also (a total of 10 verbs). We can find this meaning of the prefix in those predicates that refer to spatial transportation of an object like *preprovodit'* 'forward' (< *provodit'* 'accompany'), *prevoznesti(s)* 'eulogize, lift' (< *voznesti* 'raise'), *presledovat'* 'pursue' (< *sledovat'* 'follow'). Among marginal verbs with the 1.TRANSFER submeaning of the prefix PRE- we notice an obsolete verb *prepožalovat'* 'arrive' (< *požalovat'* 'come'). In addition, the 1.TRANSFER meaning of PRE- is present in verbs that denote giving or delivering the Trajector to a certain destination (Landmark) which is often conceptualized as a person: *prepodnesti* 'present with' (< *podnesti* 'bring to'), *predostavit'(sja)* 'grant' (< *dostavit'* 'deliver'), *prepodat'(sja)* 'give (a lesson)' (< *podat'* 'give'), *preporučit'(sja)* 'entrust smth to someone' (< *poručit'* 'give a task'), *predat'(sja)* (e.g. *ognju*) 'commit (to fire)' (< *dat'* 'give'). Note that in the two latter verbs the prefix PRE- can be replaced with PERE- with the same semantics: compare the examples given above in (3) and (4) and below in (13) and (14):

- (13) *Džon vyšel na ulicu, čtoby poznačit'sja s verbljudami, posle čego **preporučil** pomoščniku ustroit' nas.* [A. Šimanskij. Avstralija glazami ruskogo // «Zvezda», 2002]
 'John went outside in order to make acquaintance with camels, whereafter he **consigned** us to his assistant to give us accommodations.'
- (14) *Vospitanije syna otec tože polnost'ju **pereporučil** žene.* [I.Ė. Kio. Illjuzii bez illjuzij (1995-1999)] 'The upbringing of his son the father also **handed over** to his wife.'

The verb *prelstit'(sja)* 'lure, entice' (< *l'stit'* 'flatter') prefixed in PRE- arguably also represents submeaning 1.TRANSFER: this verb is parallel to some similar lexemes prefixed in PERE- like *peremanit'* 'attract to another side' (< *manit'* 'attract').

In the light of the discussion of the spatial senses ‘TRANSFER’, ‘OVER’, and ‘ACROSS’ of the prefix PERE-, it is reasonable to pay attention to which of these senses are shared by PRE-. Note that none of the verbs prefixed in PRE- mentioned above can be used with the preposition phrase *čerez* + *noun* ‘over/across something’. It appears that the verbs in PRE- lack an argument that conceptualizes a spatial obstacle which should be overcome. The verbs *presledovat* ‘pursue’ and *prepodnesti* ‘present with’ imply movement in space, while *predostavit’(sja)* ‘grant’ and *preporučit’(sja)* ‘entrust something to someone’ suggest a conceptual rather than spatial distance. This brings me to the conclusion that even in the most spatial uses of the prefix PRE- it can only refer to the simple transportation of an object (‘TRANSFER’), while the senses ‘OVER’ and ‘ACROSS’ are vague and uncertain.

Submeaning 2.TRANSFORM

The image schemas of the central spatial submeaning 1.TRANSFER OVER/ACROSS have rich semantic potential and serve to motivate various extensions of the prototype.

Submeaning 2.TRANSFORM is a metaphorical extension of the prototype 1.TRANSFER OVER/ACROSS⁹⁶. This submeaning results from mapping TRANSFER from the domain of spatial relations to the domain of qualitative states. TRANSFORM is based on the spatial image schema (d) in Figure 4: the two locations of an object are metaphorically reinterpreted as two qualitative states of an object – the initial and the final state. Thus, the sense of a spatial transportation TRANSFER gives rise to the sense of a qualitative transformation of an object – TRANSFORM. This meaning of the prefix PERE- can be observed in the verbs *peremenit’(sja)* ‘change’ (< *menjat* ‘change’), *pereinačit* ‘change’ (< adverb *inače* ‘differently’), and in the verbs *pererodit’(sja)* ‘transform into a different quality’ (< *rodit’(sja)* ‘give birth’), *pererasti* ‘grow into smth else’ (< *rasti* ‘grow’), *perevoplotit’(sja)* ‘reincarnate’ (< *voplotit* ‘embody, materialize’).

Likewise, in verbs of speech, the content can be partially altered through the process of translation, interpretation, or retelling, and so PERE- in this sense is used in the verbs *peretolkovat’(sja)* ‘interpret’ (< *tolkovat’(sja)* ‘interpret’), *pereskazat’(sja)* ‘retell’ (< *skazat* ‘say’), *perevesti* ‘translate; convert currency’ (< *vesti* ‘lead’), *perevrat’(sja)* ‘misrepresent, retell inaccurately’ (< *vrat* ‘tell lies’).

Another group of predicates that employ PERE- in this submeaning points to affecting physical properties of an object such as consistency, integrity, and moistness: *pererabotat’(sja)* ‘work into’ (< *rabotat* ‘work’), *perevarit’(sja)* ‘digest’ (< *varit* ‘cook’), *pereteret’(sja)* ‘rub a solid object into a powder’ (< *teret* ‘rub’), *peremolot’(sja)* ‘grind into flour’ (< *molot* ‘grind’), *peremolotit* ‘thresh into flour’ (< *molotit* ‘thresh’), *peretopit* ‘make a substance melt into a different consistency or quality’ (< *topit* ‘make melt’), *peregnat* ‘(chem.) distil; sublimate’ (< *gnat* ‘distill’), *peresoxnut* ‘dry out, parch’ (< *soxnut* ‘dry’), *perepret* ‘rot’ (< *pret* ‘rot’), *perekisnut* ‘sour’ (< *kisnut* ‘sour’), and *peregnit* ‘turn into humus’ (< *gnit* ‘rot’).

Overall, the database contains 26 standard verbs and 22 marginal verbs with this meaning of PERE-.

The prefix PRE- can also express submeaning 2.TRANSFORM. Consider the predicates that denote transformation of an object from one quality into a different new

⁹⁶ Unlike Janda 1986, I consider 2.TRANSFORM as a distinct subcategory for the purposes of comparison of PERE- and PRE-. Although some other subcategories are metaphorical extensions of the prototype 1.TRANSFER too, they are more specific and cannot include those lexemes where the prefix expresses submeaning 2.TRANSFORM.

quality: *prevratit'(sja)* 'turn, transform into' (< *vorotit'* 'turn'), *pretvorit'(sja)* 'convert into' (< *tvorit'* 'create'), *preobrazovat'(sja)* 'transform' (< *obrazovat'* 'form'). In addition, 2.TRANSFORM is expressed by PRE- in the verb *prestavit'sja* 'pass away' (< *stavit'(sja)* 'set') that implies a qualitative change of an object and conceptualizes death as a travel from one world to another, or a transformation from one state to another.

Note that predicates in PERE- and PRE- that refer to a qualitative transformation (2.TRANSFORM), often occur in the same context:

- (15) *Gosudarstvo obladalo dostatočnou moščju, čtoby nanovo pereigrat' to, čto uže bylo odnaždy i na veky vekov soveršeno, preobrazovat' i perevoplotit' granit, bronzu, otzvučavšie reči, izmenit' raspoloženie figur na dokumental'nyx fotografijax.* [V. Grossman. *Žizn' i sud'ba.* (1960)]

'The state had enough power to replay anew what had once and for ever been performed, to **transform** and **transmute** granite, bronze, given speeches, and change the location of characters in the documentary photographs.'

Submeaning 3.SUPERIORITY

Submeaning 3.SUPERIORITY of the prefix PERE- "compares the trajector's performance with that of another agent" (Janda 1986: 148) and is also called in the literature "a comparative PERE-" (Kagan 2013: 504). The list of predicates that represent this submeaning of the prefix indicates that it applies to a variety of activities and characteristics: e.g. *peregnat'* 'outdistance, leave behind' (< *gnat'* 'drive'), *perexitrit'* 'outwit' (< *xitrit'* 'play cunning'), *perekričat'* 'outshout' (< *kričat'* 'shout'), *pereščegoljat'* 'be more attractive than someone, while showing off' (< *ščegoljat'* 'show off'), *pereuprjamat'* 'outdo in being stubborn' (< *uprjamyj'* 'stubborn'), *peretjanut'* 'outweigh' (< *tjanut'* 'draw, pull'), *peretancevat'* 'win in dancing over someone' (< *tancevat'* 'dance'), *peresilit'* 'outdo in being strong' (< *sila* 'force, strength'), *pereborot'* 'master' (< *borot'(sja)* 'struggle, fight'), *perepljunut'* 'outdo in doing smth' (< *pljunut'* 'spit'), etc.

Submeaning 3.SUPERIORITY is based on the image schema (a) in Figure 3: the solid obstacle on the way of the Trajector serves as a point of reference (in other terms, a standard of comparison, cf. Kagan 2013: 505) that the winner (the Trajector) conquers and overpasses in his performance. This is another metaphorical extension of the prototype.

I suggest that we can look at submeaning 3.SUPERIORITY as an intensification of the activity named by the verbal base. Figure 5 juxtaposes two propositions – the proposition of a simplex verb (image to the left) and the proposition of the verb prefixed in PERE- in the meaning 3.SUPERIORITY (image to the right).

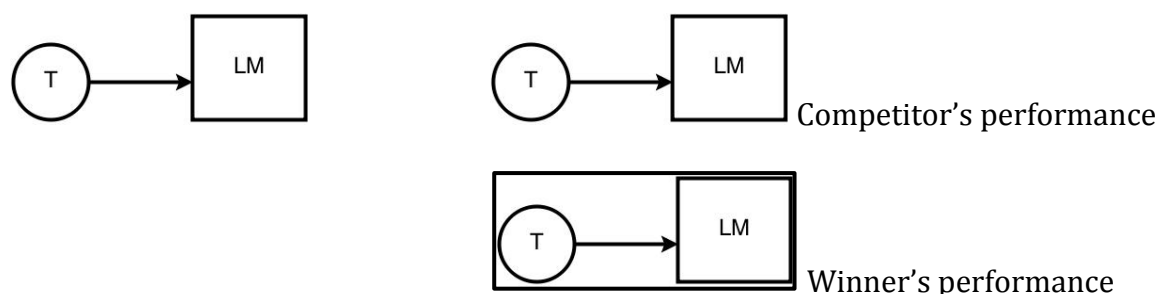


Figure 5: Submeaning 3.SUPERIORITY as intensification of activity.

Figure 5 shows that the proposition of the prefixed verb implies a comparison of the foregrounded winner's performance with the performance of his competitor taken place previously. The winner's performance quantitatively reduplicates and qualitatively outrivals the performance of the competitor. The competitor's performance stays in the background and is a presupposition of the superior performance carried out by the winner. In other words, PERE- in submeaning 3.SUPERIORITY performs the function of an intensifier.

Submeaning 3.SUPERIORITY of PERE- is highly productive and very frequent across verbs, both standard and marginal. PERE- is used in this sense in 20 standard lexemes and in 46 marginal verbs attested in the corpus.

Corpus data shows that a predicate prefixed with comparative PERE- can theoretically be formed from any verbal simplex that denotes an activity (consider examples (16)-(20)). Moreover, it is possible to create such predicates also from non-verbal bases like adjectives and nouns (examples (21) and (22)). Below I provide a few curious examples of marginal verbs from the corpus. They all show that the comparative use of PERE- alters the argument structure of the base: a derivative always implies a direct object position to lexicalize the competitor. The simplex bases of the verbs in the corpus examples of *šumet'* 'make noise', *čudit'* 'act odd', *gljadet'* 'stare', *gresti* 'paddle', *skandalit* 'make a scene', *spletničat'* 'gossip' are intransitive verbs, but the attachment of PERE- used in submeaning 3.SUPERIORITY turns these simplex verbs into transitive perfectives:

- (16) *Dnem čajnik tixij – parit sebe, ne v silax **perešumet'** mašiny za oknom* [Z. Prilepin. San'ka (2006)]
'At the daytime the tea-kettle is quiet – it steams, unable to **be louder than** the cars in the street.'
- (17) *Vydumščika kamer-junkera udalos' **perečudit'** tol'ko peterburgskomu žitelju E.F. Ganinu.* [A. Maksimova. Čudaki (2000)]
'Only E.F. Ganin, the resident of St. Petersburg, managed to **act more odd than** the tale-teller chamber junker.'
- (18) *Nekotoroe vremja èto napominalo detskuju igru v "gljadelki": kto kogo **peregljadit, ne morgaja.*** [G. Polonskij. Ne pokidaj (1998)]
'For a while it looked like a child's staring contest: who **outperforms** whom **in staring** and does not blink.'
- (19) *Teper' vyxod iz Virmy naproč' zakryt: nam javno ne **peregresti** vstrečnyj veter.* [V. Galenko. Zabytyj bereg. // «Vokrug sveta», 1990]
'Now the way out of Virma was shut entirely: we could not **win over** the wind **by paddling** against it.'
- (20) *My s vami svideteli: načalis' bega. Kto kogo **pereperčit, peresolit, perekompromatit, peregvozdit, pereskandalit, perespletničat, pereobložit, perekričit:** splošnye PERE.* [V. Agranovskij. Vtoraja drevnejšaja. Besedy o žurnalistike (1976-1999)]
'You and me are eyewitnesses: racing has begun. Who **outdoes** whom **in peppering, salting, digging out blackmail material, knocking, making a scene, gossiping, swearing, shouting** – PERE all over.'

- (21) [L]iteratorskij cinizm kogo xočeš' **pereciničit**: ty dumaješ', čto ispol'zueš' literatora, an gljad' – on uže ispol'zoval tebjja. [D. Bykov. Orfografija (2002)]
 'The cynicism of a writer can **outdo** anybody **in being cynical**: when you think you are using a writer, stop and look – he has already used you.'
- (22) V pis'me Korolenko Gor'kij pisal: «Ja ne ljublju Arcybaševa <...>. Naxvatal u Tolstogo, Dostoevskogo, delaet èto tol'ko dlja togo, čtoby **pereleonidit**' Andreeva v pessimizme». [Ju. Bezeljanskij. V sadax ljubvi (1993)]
 'Gor'kij wrote in the letter to Korolenko: "I dislike Arcybašev <...>. He adopted things from Tolstoy, Dostoevsky and he does it only in order to **outdo Leonid Andreev in being pessimistic**".'

Like PERE-, the prefix PRE- can also express submeaning 3.SUPERIORITY. This use of PRE- is attested at least in two lexemes *preobladat* 'prevail' and *prevzoi* 'surpass, excel in' formed from the verbs *obladat* 'possess' and *vzoi* 'mount, ascend' respectively. Note that only the latter verb becomes transitive through the attachment of PRE- (*vzoi*_{INTR} > *prevzoi*_{TR}), whereas the intransitivity of the former base *obladat*_{INTR} 'possess' is not altered by the prefix PRE-: *preobladat*_{INTR}.

Submeaning 4.OVERDO

Submeaning 4.OVERDO is another metaphorical extension of the prototype: it activates the image schema (b) in Figure 3, where the Trajector crosses a boundary. This boundary in 4.OVERDO is a norm, or an expected, prescribed, or ideal standard for performance. The verbs prefixed with PERE- in this submeaning refer to excessive performances, and therefore this use of PERE- is called EXCESS in (Janda 1986: 149) and "pere- of excess" in (Kagan 2013: 500). Since the norm is violated, the verbs that represent PERE- in this sense often have negative connotations: *pereperčit* 'add too much pepper' (< *perčit* 'use pepper'), *pereest* 'overeat' (< *est* 'eat'), *pereučit'sja* 'study too much' (< *učit'sja* 'study'), *pereplatit* 'overpay' (< *platit* 'pay'), *pereslastit* 'make too sweet' (< *slastit* 'sweeten'). Attachment of PERE- in this submeaning does not alter the transitivity of the verbal base. The Trajector (the normal performance) is not lexicalized in the context but is implied in the submeaning of the prefix.

I propose that submeaning 4.OVERDO can be also interpreted as an intensification of activity. We observe a contrast of two propositions (Figure 6), similar to the one discussed with regard to 3.SUPERIORITY (Figure 5). Again, the proposition of the simplex base implies a single performance (image on the left), whereas the proposition of the prefixed verb compares two performances (image on the right).

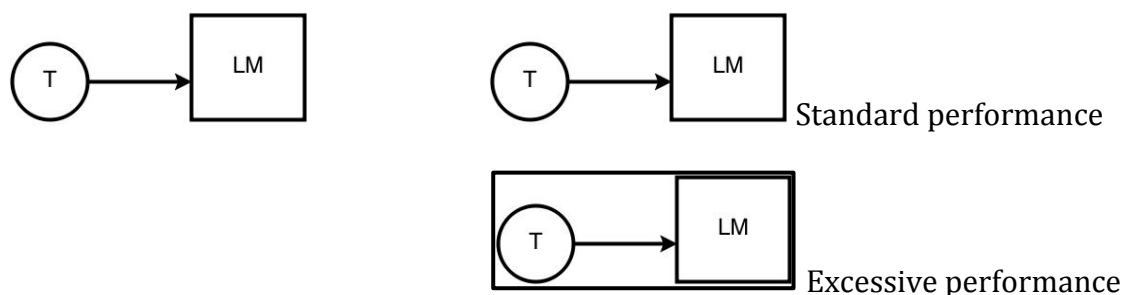


Figure 6: Submeaning 4.OVERDO as intensification of activity.

The focus is on the performance carried out by a subject (Trajector) in reality. This foregrounded performance exceeds the standard, set up by an imaginary performance. The standard is presupposed and serves as a reference point for the comparison and the evaluation of the real performance as excessive.

This use is the third most productive for PERE- and is attested in 50 standard verbs and 65 marginal verbs. Corpus data includes examples of occasionally coined predicates like *peregustit'* 'make a forest be too dense', *perekraxmalit'* 'starch too strongly', *perekupat'sja* 'overdo in bathing', *perenjančit'* 'over-nurse', *pereogromit'* 'make too gigantic', *pereostorožničat'* 'overdo in acting safe', *peretoropit'* 'make hurry too much', and others. Consider also an example of such a use given in (23) that contextualizes a marginal factitive verb *peretjaželit'* 'make too heavy' formed from the adjective *tjaželyj* 'heavy':

- (23) *Izbrana ploščad' kryla, no esli peretjaželit' samolet, ego prosto ne budet.* [F. Čuev. Il'jušin (1998)]
 'The surface space of the wing is chosen, but if one **makes** the airplane **too heavy**, it will simply not exist.'

Similarly to PERE-, the prefix PRE- expresses submeaning 4.OVERDO in the standard verbs *preuveličit'(sja)* 'exaggerate, overstate' (< *uveličit'* 'increase, enlarge'), *preumen'shit'(sja)* 'underestimate, belittle' (< *umen'shit'* 'reduce'), *presytit'* 'sate, overfeed with' (< participle *sytyj* 'satiated'), and *prevysit'* 'exceed' (< noun *vys'* 'height'). In these verbs the prefix PRE- contributes the semantics of excess and points to the unwanted outcome of overdoing an action.

Submeaning 5.REDO

Submeaning 5.REDO⁹⁷ refers to the image schema (d) in Figure 4 and partially resembles 2.TRANSFORM, which turns two locations of the Trajector into two states or qualities. The scenario of 5.REDO indicates that an activity carried out previously is performed one more time and presumably in a better way. As a result, as Janda (1986: 153) points out, "the product of an action (landmark) is either repaired or changed fundamentally".

I suggest that in submeaning 5.REDO the prefix performs the function of an intensifier of the activity named by the base. In particular, the prefix refers to repetition, or reduplication, of the proposition. Figure 7 demonstrates the contrast between the proposition denoted by the simplex base (image to the left) and the proposition denoted by the verb prefixed with PERE- in submeaning 5.REDO (image to the right).

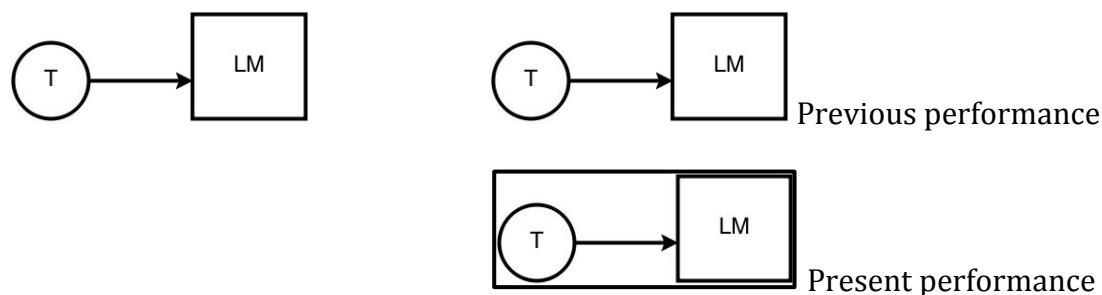


Figure 7: Submeaning 5.REDO as intensification of activity.

⁹⁷ Other terms used in the literature for this submeaning are "repetitive *pere-*" (Tatevosov 2008) and "iterative *pere-*" (Kagan 2013: 507).

The Trajector (T) affects the Landmark (LM) the second time, and achieves a better result compared to the previous performance. The previous performance is in the background as a presupposition of the new, better, and foregrounded performance.

Submeaning 5.REDO is the most frequently attested use of PERE- in terms of individual verbal lexemes. We observe this submeaning in 92 standard verbs and 111 marginal predicates in the database. Among standard verbs there are lexemes like *peredelat* 'do anew' (< *delat* 'do'), *perenumerovat* 're-number' (< *numerovat* 'number'), *pereoborudovat*(*sja*) 're-equip' (< *oborudovat* 'equip'), *pereodet* 'dress anew' (< *odet* 'dress'), *pereprjač* 're-harness (a horse)' (< (*za*)*prjač* 'harness (a horse)'), *pereprjatat* 're-hide, hide anew' (< *prjatat* 'hide'), *pereosmyslit* 're-think' (< *osmyslit* 'think'), *perepelenat* 're-swaddle' (< *pelenat* 'swaddle'), *perekrasit* 're-paint' (< *krasit* 'paint'), *perezdat* 're-publish' (< *izdat* 'publish'), *perezagruzit* 'restart, reupload computer' (< *zagruzit* 'upload'), and others. Example in (24) illustrates a standard verb *perečitat* 're-read', and the sentence in (25) provides a context for a marginal coinage *pereaxat* 'say aah again':

(24) *Emu zaxotelos' perečitat' pis'mo, i on vytaščil složennyj vdvoe konvert iz karmana.* [V. Pelevin. *Želtaja strela* (1993)]
 'He wanted to **read** the letter **again**, and he took the folded envelope out of his pocket.'

(25) — *Ne tak axaeš', pereaxaj snova, — s udovol'stvijem obronil Uxmyl.* [E. Lukin. *Katali my vaše solnce* (1997)]
 'You are saying aah in a wrong way, **say aah again**, — Uxmyl remarked with pleasure.'

Unlike PERE-, the prefix PRE- is not attested with submeaning 5.REDO.

Submeaning 6.OVERCOME/DURATION

Submeaning 6.OVERCOME/DURATION of PERE- is a metaphorical extension of the prototype 1.TRANSFER OVER/ACROSS to the domain of time (cf. on TIME IS SPACE metaphor in Lakoff & Johnson 1980; Haspelmath 1997). As Janda (1986: 143) puts it, in this submeaning "the landmark is a period of time during which the trajector pursues a given activity". The double label of this submeaning indicates two facets – OVERCOME and DURATION. Below I explain this distinction within the group of verbs that temporally reinterpret the spatial TRANSFER OVER/ACROSS (cf. also Endresen et al. 2012: 274-275 for discussion). Note that this distinction is also supported by the analysis by Kagan (2013: 495-6) who draws attention to two major subtypes of the "temporal" PERE-.

If a verb activates the image schema (c) in Figure 3, the Trajector proceeds from one boundary of a period of time to another. This scenario can be called 'DURATION', and we find it in the verbs *perezimovat* 'pass the winter' (< *zimovat* 'pass the winter'), *perenočevat* 'pass the night' (< *nočevat* 'pass the night'), *peresidet* 'sit during some period until it is safe' (< *sidet* 'sit'), *perezdat* 'wait through some time until it is safe' (< *ždat* 'wait'). Similar verbs denote 'spend a period of time doing something': *perekusit* 'have a light meal' (< *kusat* 'bite') and *perekurit* 'smoke for a while' (< *kurit* 'smoke').

Other verbs activate the image schema (a) of 1.TRANSFER OVER/ACROSS (Figure 3), with an obstacle in the way of the Trajector that it has to OVERCOME. The obstacle can be

unwanted trouble, danger, or physical pain. We can observe this use of PERE- in *pereterpet* ‘endure some time suffering’ (< *terpet* ‘endure’), *perebit’sja* ‘overcome’ (< *bit’(sja)* ‘beat’), *peremoč’(sja)* ‘overcome’ (< *moč’* ‘be able’), *perebolet* ‘recover from illness, become well again’ (< *bolet* ‘be ill’), *perenesti(s’)* ‘overcome (difficulties)’ (< *nesti* ‘carry’), *perežit* ‘live through, outlive, survive’ (< *žit* ‘live’), and *perekantovat’sja* ‘survive’ (< *kantovat’sja* ‘wait until better time’). These verbs are normally transitive unless they contain the postfix *-sja*.

Additional group of verbs representative of PERE- in the temporal sense imply both OVERCOME and DURATION and denote an intensive activity or a state that lasted during a period of time but came to an end: *perexotet’(sja)* ‘stop wanting’ (< *xotet* ‘want’), *perebesit’sja* ‘stop behaving crazy, overcome being crazy’ (< *besit’sja* ‘be furious’), *peregoret* ‘burn out, stop burning or wanting something’ (< *goret* ‘burn’), *perekipet* ‘boil through some time and stop’ (< *kipet* ‘boil’), *perestat* ‘stop’ (< *stat* ‘become’). These verbs are intransitive. Examples in (26) and (27) demonstrate that in this group of verbs the Landmark is a state or an activity which is extended in time and which is simultaneously an “obstacle” that the agent has to overcome:

(26) *Pokupka delaetsja ne ottogo, čto xočetsja, a dlja togo, čtoby bol’she ne xotet’. Čtoby «perexotet’».* [E. Piščikova. Pjatiètažnaja Rossija (2007)]
 ‘People buy stuff not because they want to have it but rather in order not to want it anymore. In order **to stop wanting**.’

(27) *I voobščè ja mnogo čego xoču. I ne perexočivaetsja, i ne pereterplivaetsja.* [Ju. Daniël’. Pis’ma iz zaključenija (1966-1970)]
 ‘And in general I want many things. And it does not feel like I **stop wanting and overcome it**.’

Overall, submeaning 6.OVERCOME/DURATION is attested in 23 standard verbs prefixed with PERE- and 19 marginal verbs which indicate productivity of this use of the prefix. Marginal verbs with PERE- in this meaning include coinages that realize ‘OVERCOME’ (*perebedovat* ‘outlive, survive’ < *bedovat* ‘suffer, live in misery’) or ‘DURATION’ (*pereletovat* ‘pass the summer’ < *letovat* ‘spend summer’). A number of marginal verbs mean ‘spend some time doing something’: *peredremat* ‘spend some time taking a nap’ (< *dremat* ‘take a nap’), *peregodit* ‘wait through’ (< *godit* ‘wait’). Some coinages imply both OVERCOME and DURATION and refer to an interval of time “filled” with an intense activity or a state that is over: *perešumet* ‘stop being noisy after a while’ (< *šumet* ‘make noise’), *perecvesti* ‘stop blossoming’ (< *cvesti* ‘blossom’), *pereserdit’sja* ‘stop being angry after a while’ (< *serdit’sja* ‘be angry’), *pereljubit* ‘love someone during some time and stop loving’ (< *ljubit* ‘love’), *perekolbasit’sja* ‘be done with some activity’ (< *kolbasit’sja* ‘do something intensively’).

The prefix PRE- expresses submeaning 6.OVERCOME/DURATION in several verbs that feature ‘OVERCOME’. These predicates include *preterpet* ‘undergo’ (< *terpet* ‘endure’), *preodolet’(sja)* ‘overcome’ (< *odolet* ‘conquer, master’), *prevozmoč’* ‘overcome’ (< *vozmoč’* ‘be able’), *prenebreč’* ‘neglect’ (< *ne bereč’* ‘not take care’). In addition, there is one verb with PRE- ‘DURATION’ – *prebyt’_{PF}* ‘last, remain constant’ (< *byt* ‘be’), which is mostly used in its imperfective form *prebyvat’_{IPF}* ‘be in the state of’.

Submeaning 7.BRIDGE

Submeaning 7.BRIDGE implies that in the image schema (a) (Figure 3) the Trajector coincides with its trajectory: “The Trajector progresses, extending itself across a gap, and leaving a solid path” (Janda 1986: 157). In this scenario, the Trajector is usually a concrete physical object like a bandage, belt, or a piece of cloth stretched over an object. Standard verbs that represent PERE- in this sense include *perebintovat’* ‘put a bandage over’ (< *bintovat’* ‘bandage’), *peremotat’* ‘put a bandage over’ (< *motat’* ‘wind’), *perekryt’(sja)* ‘cover over’ (< *kryt’* ‘cover’), *perevjazat’(sja)* ‘bandage, tie up’ (< *vjazat’* ‘tie’), *perevit’(sja)* ‘weave across’ (< *vit’* ‘weave’), *perepojasat’(sja)* ‘tie a belt around’ (< *pojas’* ‘belt’), *peretjanut’* ‘bridge’ (< *tjanut’* ‘pull’)⁹⁸, *perexlestnut’(sja)* ‘whip over; float over’ (< *xlestnut’* ‘whip; float’), *perekinut’* (e.g. *most*) ‘span over (a bridge)’ (< *kinut’* ‘throw’), and *perekrestit’(sja)* ‘make the sign of the cross over’ (< *krestit’* ‘baptize’). As pointed out by Janda (1986: 157), this submeaning is very dependent on the context, and can be best perceived through the semantic support of the words that fill in argument positions of the subject and direct object:

- (28) *Ruka ego byla akkuratno peretjanuta bintami, pod kotorye japonec položil mnogo kakix-to trav.* [A. Gelasimov. *Stepnye bogi* (2008)]
‘His arm was carefully **tied** with bandages under which the Japanese put lots of herbs.’

Among marginal verbs that feature PERE- in this submeaning we find *peregorbit’sja* ‘crook over’ (< *gorbit’sja* ‘crook’), *perepolosovat’* ‘put a bandage’ (< *polosovat’* ‘put stripes’), and *pereplesnut’(sja)* ‘splash over (about water)’ (< *plesnut’* ‘splash’). Example in (29) illustrates the use of the latter verb. Here, the Trajector that is expanded along its trajectory is a wave thrown over a boat:

- (29) *Štorm načalsja vnezapno; my ne ponjali, čto proisxodit, kogda vysokaja volna pereplesnula palubu.* [I. G. Ėrenburg. *Ljudi, gody, žizn’* (1960-1965)]
‘The storm began suddenly; we did not understand what was going on, when a high wave **splashed over** the deck.’

I suggest that the same alignment of the Trajector and the trajectory accounts for the semantics of verbs that refer to solid objects that are bent rather than tied across: *peregnut’(sja)* ‘bend over’ (< *gnut’* ‘bend’). Therefore, in my analysis in the submeaning 7.BRIDGE I merge together two of Janda’s subcategories: <Bend> (ibid: 168) and <Bridge> (ibid: 157). Consider the example in (30) which differs from (28) and (29) only in terms of the flexibility and texture of the object:

- (30) *Ustroivšis’ poudobnee, on peregnul gazetv dvovoe i pogruzilsja v čtenie.* [V. Pelevin. *Želtaja strela* (1993)]
‘After he had set comfortable, he **folded** the newspaper in half and immersed himself in reading.’

⁹⁸ I distinguish this sense of *peretjanut’* with PERE- ‘7.BRIDGE’ from four other uses of this verb – *peretjanut’(sja)* ‘pull to another place’ with PERE- ‘1.TRANSFER’, *peretjanut’* ‘outweigh’ with PERE- ‘3.SUPERIORITY’, *peretjanut’* ‘pull in too tight, wait for too long’ with PERE- ‘4.OVERDO’, and *peretjanut’* ‘stretch anew’ with PERE- ‘5.REDO’.

The prefix PRE- expresses the same submeaning 7.BRIDGE in the verb *preklonit'(sja)* 'bow (one's head or knee) in respect or worship' formed from the simplex *klonit'* 'bend, incline'. This verb is parallel to *peregnut'(sja)* 'bend over' (< *gnut'* 'bend'), and it is reasonable to say that PRE- and PERE- in these predicates refer to the same situation.

Submeaning 8.INTERCHANGE

Submeaning 8.INTERCHANGE is a metaphorical extension of 7.BRIDGE – it applies the same idea from the spatial domain to the domain of communication. The trajectory of BRIDGE is duplicated and reversed in direction. Most verbs that represent PERE- in this sense refer to exchange of sounds or other signals for the purposes of communication across two or more living beings (Trajectors). Examples are numerous: *perešeptjvat'sja* 'exchange wispers' (< *šeptat'* 'wisper'), *pereščelkivat'sja* 'exchange clacking sounds' (< *ščelkat'* 'clack'), *perestukivat'sja* 'exchange signals by knocking on something' (< *stučat'* 'knock'), *peresvistjvat'sja* 'exchange whistling sounds' (< *svistet'* 'whistle'), *perekrikivat'sja* 'exchange loud sounds' (< *kričat'* 'shout'), *perezvanivat'sja* 'exchange phone calls' (< *zvonit'* 'call'), *perešučivat'sja* 'exchange jokes' (< *šutit'* 'make jokes'), *peregljadyvat'sja* 'exchange glances' (< *gljadet'* 'look at'), *peresmeivat'sja* 'exchange laughing sounds' (< *smejat'sja* 'laugh'), etc.

As pointed out by Janda (1986: 158), the space between the Trajectors constitutes a gap (the Landmark). The signals travel across this gap (both spatial or imaginary) in the manner of exchange, sent from one participant to the other and back again. This reduplicated trajectory of 7.BRIDGE and the exchange of signals can be visualized as shown in Figure 8:

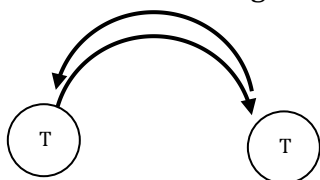


Figure 8: Image schema of submeaning 8.INTERCHANGE (from Janda 1986: 158).

By means of the cognitive notions Trajector and Landmark and their configuration, Janda captures precisely the schema of subcategory 8.INTERCHANGE. This cognitive model is entirely consistent with subtle nuances of native speaker's intuition, as we can see from the description provided by Kornej Čukovskij. In the fragment of his book cited in (31), Čukovskij explains the semantics of the noun *perelaj* created in child language (lit. 'PERE-bark' – 'the process of exchanging barking signals by dogs') which is a nominalization of the verb *perelaivat'sja* 'exchange barking sounds' (< *lajat'* 'bark'):

- (31) *Čtoby ob'jasnit' «perelaj» inostrancu, prišlos' by pribegnut' k takoj mnogoslovnoj opisatel'noj reči: lajut dve sobaki (ili bol'se) s dvux protivopoložnyx storon, pričem ne srazu, a poperemmeno – edva umolkaet odna, totčas že prinimaetsja lajat' drugaja.* [K. I. Čukovskij. *Ot dvux do pjati.* (1933)]

'In order to explain a foreigner what «**perelaj**» (lit. **PERE-bark**) is, one would have to employ such a wordy periphrastic description: there are two (or more) dogs that are barking from two opposite sides. Moreover, they are barking not simultaneously but by turns – as soon as one dog becomes silent, the other dog begins.'

According to Čukovskij (1933), the novel coinage *perelaj* is produced by a child by analogy with other nominalizations that illustrate the same productive pattern of derivation – *perezvon* ‘jingle-jangle’ (< *zvonit* ‘ring the bell’), *perebranka* ‘cross-talk’ (< *branit* ‘curse’), *peregovory* ‘negotiations’ (< *govorit* ‘talk’), *perepiska* ‘exchange of letters’ (< *pisat* ‘write’). In all these words the prefix PERE- denotes both reciprocity and discontinuity.

Note that all verbs cited above for PERE- of 8.INTERCHANGE are morphologically imperfective and reflexive. Indeed, PERE- in this submeaning is normally accompanied with the reciprocal intransitivizing postfix *-sja* and a marker of the imperfective aspect *-iva-* which points to iterativity of the exchange. As an example of this combination of morphemes, consider the verb *pere-pis-yva-t'-sja* ‘exchange letters’ (< *pisat* ‘write’). However, the perfective and non-reflexive counterparts of this verb are not representative of the meaning 8.INTERCHANGE of the prefix: the verbs *perepisat'sja* and *perepisat* exist, but with other senses of PERE- – ‘TRANSFER’ and ‘REDO’: ‘copy from one place to another by writing’ and ‘re-write’ accordingly. Regarding predicates like *pere-pis-yva-t'-sja* ‘exchange letters’, Janda (1986: 160) fairly pointed out that “these verbs are not derived imperfectives, but rather iteratives formed directly from base verbs by the addition of both prefix and suffix (and particle) at once”. In other words, the combination of the three morphemes – the prefix PERE-, the imperfective suffix, and the reciprocal postfix – form a morphological construction that functions as a single morphological structure.

This construction with PERE- of 8.INTERCHANGE is very productive and can be frequently observed among both standard verbal predicates (22 lexemes) and marginal coinages attested in the corpus (27 lexemes). Marginal verbs with PERE- of 8.INTERCHANGE expand this use to various kinds of animal communication and denote exchange of purring sounds (*peremjaukivat'sja* and *peremurlykivat'sja* < *mjaukat* ‘murr’, *murlykat* ‘purr’), cackling signals (*perekudaxtyvat'sja* < *kudaxtat* ‘cackle’), quacking (*perekvakivat'sja* < *kvakat* ‘quack’), bleating (*perebleivat'sja* < *blejat* ‘bleat’), and roaring (*pererykivat'sja* < *rykat* / *ryčat* ‘roar’). Moreover, marginal words collected from the corpus make it possible to revise earlier accounts based primarily on dictionary data. For example, it was established that the use of PERE- in submeaning 8.INTERCHANGE is only compatible with iterative actions and therefore is restricted to imperfective verbs (Janda 1986: 159). However, corpus data indicates that marginally there can be formed perfective alternatives of standard imperfective verbs and if they occur they always refer to a single act of the exchange. For example, consider semelfactive occasional verbs attested in the corpus, such as: *perešepnut'sja* ‘exchange wispers once’ (< *šepnut* ‘wisper once’) in (32), *perexrjuknut'sja* ‘grunt once at each other’ (< *xrjuknut* ‘grunt once’) in (33), and *perekivnut'sja* ‘nod to each other once’ (< *kivnut* ‘nod once’) in (34):

(32) *My s Anatolijem perešepnulis'PF – vodki vzjato s pereborom.* [S. Jurskij. Počem v Pariže kartoška? (1991)]

‘Anatolij and I **exchanged wispers** – they went over the top with the supply of vodka.’

(33) *Svin'ja i dva porosenka <...> ne udostoili Bima vnimaniem, a prosto perexrjuknulis'PF meždu soboj ironičeski.* [G. Troepol'skij. Belyj Bim černoje uxo (1971)]

‘The pig and two piglets did not honor Bim with their attention, instead they just ironically **exchanged grunting sounds** with each other.’

- (34) [K]ak tol'ko oni razgljadeli Mašu, to *peregljanulis'*_{PF} i *perekivnulis'*_{PF}.
[E. Evtušenko. «Volčij pasport» (1999)]
'As soon as they saw Maša, they **exchanged glances** and **nods**.'

Moreover, there are a few perfective verbs with PERE- of 8.INTERCHANGE among standard lexemes as well: *peremolvit'sja* 'exchange words' (< *molvit* 'utter'), *peregljanut'sja* 'exchange glances' (< *gljanut* 'take a look at'), *peremignut'sja* 'exchange blinking signals' (< *mignut* 'blink'), *perekinut'sja* 'exchange signals' (< *kinut* 'throw'), and *perebrosit'sja* 'exchange words' (< *brosit* 'throw')⁹⁹. Among other perfective verbs that also represent the PERE- of 8.INTERCHANGE are *peremnožit'(sja)* 'multiply two or more numbers with one another' (< *množit* 'multiply'), *peregovorit* 'have a word with someone' (< *govorit* 'talk'), and *peredraznit* 'mimic' (< *draznit* 'tease'). These lexemes as well as a marginal rare verb *pereopylit* 'pollinate with each other' (< *opylit* 'pollinate') suggest that the presence of the reflexive postfix *-sja* is also optional, and that the reciprocity of the action can be expressed by the prefix PERE- alone.

The prefix PRE- forms two verbs that resemble predicates prefixed in PERE- of 8.INTERCHANGE. Both verbs are reciprocal and refer to verbal communication: *prerekat'sja* 'argue with someone' (< *reč* 'speak') and *prepirat'sja* 'squabble with someone' (< *peret* 'push')¹⁰⁰. Note that there are semantically similar verbs in PERE- – a quite frequent verb *pererugivat'sja* 'exchange arguings' (< *rugat'(sja)* 'argue') (155 corpus examples) and a marginal verb *perekorjat'sja* 'exchange curses' (< *korit* 'reprove'). In addition, there exists a rarely used verb *perepirat'sja* 'argue with each other' (< *peret* 'push, make one's way forcefully') which is claimed to be semantically identical to *prepirat'sja* (Efremova 2000):

- (35) *Poka oni perepirajutsja s Sergeem Klimenko o pravax na prioritet poseščenija Pantjuxinskoj, snimaju vsju kompaniju na video* [K. Serafimov. Èkspedicija vo mrak (1978-1996)]
'While they are **arguing** with Sergej Klimenko about the priority of visiting Pantjuxinskaja, I am filming the whole group.'

Parallel examples of PERE- and PRE- with semantics of verbal quarrelling suggest that submeaning 8.INTERCHANGE can be expressed by both prefixes in question.

Submeaning 9.TURN OVER

Submeaning 9.TURN OVER is spatial and accounts for those verbs prefixed in PERE- that refer to an object which flips over and turns around its own axis (Janda 1986: 170).

⁹⁹ In her thorough account, Janda (1986: 160) does not entirely exclude the possibility of semelfactive verbs with PERE- of 8.INTERCHANGE. Janda mentions that "[t]here are few actions of which a single exchange constitutes a realistic and potentially useful unit" (ibid: 160). The perfective verbs that I found constitute a minority in the group of predicates with PERE- in this sense: they definitely indicate the possibility but to an even larger extent they conform the opposite tendency manifested by imperfective reciprocal verbs described by Janda.

¹⁰⁰ Isačenko (2003/1965: 285) supports the idea of semantic similarity between *prerekat'sja* / *prepirat'sja* and the verbs prefixed with PERE- of 8.INTERCHANGE. At the same time, I suggest that the prefix PRE- in these verbs might simultaneously refer to breaking or blocking the signal sent by the conversation partner.

Schematically it is visualized in Figure 9. This submeaning makes another modification of the original image schema of 1.TRANSFER OVER/ACROSS: the Trajector coincides not with its trajectory (as in 7.BRIDGE) but with the Landmark, and as a result of the motion it turns over.

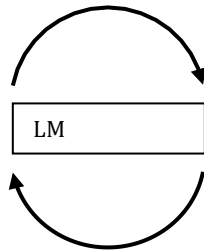


Figure 9: Image schema of submeaning 9.TURN OVER (from Janda 1986: 170).

Apart from the verb *perevernut'(sja)* 'turn over' discussed by Janda, this submeaning of PERE- is present in other similar predicates like *perelistnut'(sja)* 'turn over a page' (< *listnut'* 'turn a page'), *perevorotit'(sja)* 'turn over' (< *vorotit'* 'turn'), *perekuvyr(k)nut'sja* 'turn a somersault over oneself' (< *kuvyr(k)nut'sja* 'turn a somersault'), and *perelicevat'* 'turn (an article of clothing)' (< *licevat'* 'turn'). Additionally, verbs that refer to twisting and distorting an object also might feature this use of PERE-: *perekrutit'(sja)* 'twist over' (< *krutit'* 'twist'), *perekosobočit'(sja)* 'warp, distort' (< *kosobočit'(sja)* 'squint'), and *perekosit'* 'warp' (< *kosit'* 'twist').

Overall, PERE- in submeaning 9.TURN OVER is attested in 10 standard lexemes and 11 marginal verbs. The latter include *pereprokinut'* 'turn over' (< (*o*)*prokinut'* 'throw'), *perevertet'(sja)* 'turn over' (< *vertet'* 'turn'), *perekrjučit'* 'crook over' (< *krjučit'* 'twist'), *perekružít'sja* 'spin over' (< *kružít'sja* 'spin'), and other verbs of similar semantics.

The prefix PRE- is not attested with submeaning 9.TURN OVER.

Submeaning 10.Mix

Whereas submeaning 9.TURN OVER is "reflexive" and limited to verbs of turning and twisting, submeaning 10.MIX employs the same spatial image schema but repeats it multiply, shuffling and mixing multiple trajectors with each other, as shown in Figure 10:

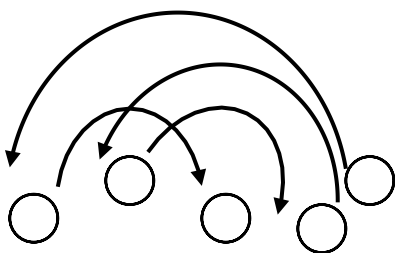


Figure 10: Image schema of submeaning 10.MIX (from Janda 1986: 172).

PERE- refers to 10.MIX, when it forms verbs *peretasovat'(sja)* 'change the order of cards' (< *tasovat'* 'shuffle'), *peremešat'(sja)* 'intermix' (< *mešat'* 'mix'), *pereputat'(sja)* 'mix up' (< *putat'* 'mix'), *pereplesti(s')* 'interlace, interweave' (< *plesti* 'plait'), and also those verbs that imply mixing particles or layers: *perepaxat'* 'furrow' (< *paxat'* 'plow'), *perekopat'* 'dig and mix soil, making it good for planting something' (< *kopat'* 'dig'), *perevorošit'*

'intermingle, mix over' (< *vorošit* 'turn'). In addition to 7 standard verbs, the same use of PERE- can be found in 9 marginal verbs including *peresloit*'(*sja*) 'interlace' (*sloit* 'make a layer'), *perevejat* 'mix by blowing' (< *vejat* 'blow'), *peregresti* 'mix by raking' (< *gresti* 'rake'), and *peremutit*'(*sja*) 'mix' (< *mutit* 'trouble'). 10.MIX is the most restricted submeaning of PERE- in terms of the number of lexemes that represent it. Yet, corpus examples of marginal verbs like (36) indicate that this use is not unproductive. Quite the opposite, it is applicable to various bases that support this use:

- (36) [O]dnako tjaželoje èto delo – **pereboltat**' veslami celoe more. [Č. Ajtmatov. Pegij pes, beguščij kraem morja (1977)]
 'However, this is a difficult task to **stir up** the entire sea with oars.'

The prefix PRE- is not attested with submeaning 10.MIX in my data.

Submeaning 11.DIVIDE

In submeaning 11.DIVIDE the prefix PERE- refers to situations when the integrity of the Landmark is violated: the Trajector cuts it across and divides it in two parts, as shown in Figure 11:

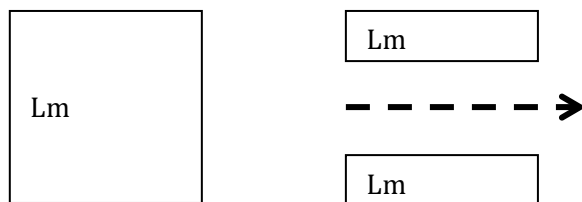


Figure 11: Image schema of submeaning 11.DIVIDE (from Janda 1986: 162).

The prefix PERE- acquires this sense in verbs that denote various ways of forced physical impact on a solid object – cutting (*pererezat*'(*sja*) 'cut off' < *rezat* 'cut'), chopping (*pererubit*'(*sja*) 'chop off' < *rubit* 'chop'), hitting (*perebit* 'break' < *bit* 'hit'; *perešibit*'(*sja*) 'hit across'), tearing apart (*perervat*'(*sja*) 'tear apart' < *rvat* 'tear'), sawing (*perepilit* 'saw apart' < *pilit* 'saw'), breaking (*perelomit*'(*sja*) 'break in two parts' < *lomit* 'break'), pressing (*peredavit* 'press across' < *davit* 'press'; *perežat*'(*sja*) 'divide by squeezing' < *žat* 'squeeze'), rubbing (*pereteret*'(*sja*) 'divide by rubbing across' < *teret* 'rub'), biting (*perekusit*'(*sja*) 'bite in two parts' < *kusat* 'bite'), and gnawing (*peregryzt* 'gnaw apart' < *gryzt* 'gnaw') as well as other manners of partitioning like creating a barrier: *peregorodit* 'partition off, barricade' < *gorodit* 'fence', *peremknut* 'block by a barrier', *pereryt* 'dig across' < *ryt* 'dig'. In all these verbs PERE- contributes to simplex bases the submeaning 11.DIVIDE in a very concrete spatial physical sense.

Interestingly, the prefix PRE- attaches to the same bases with concrete spatial semantics of loss of integrity *seč*' 'cut', *rvat*' 'tear', *gorodit*' 'barrier', and *lomit*' 'break', but it turns these verbs into abstract lexemes that denote putting an end to an activity or a process. Compare derivatives in PRE-: *preseč*'(*sja*) 'cut short, stop an activity', *prervat*'(*sja*) 'break off, interrupt':

- (37) *Tak polučilos', čto otdyx na kurorte prišlos' prervat*'. [N. Sverčok. My vinovaty už tem, čto xočetsja nam exat' (2002) // «Soči», 2002.08.22]
 'It happened that we had to **interrupt** our vacations at the seaside resort.'

So, in these verbs PRE- expresses submeaning 11.DIVIDE, but instead of a spatial facet of this meaning observed for PERE-, PRE- transforms it into an abstract figurative sense applicable to activities rather than concrete objects. This observation supports the widely spread view that the Slavonic prefix PRE- has a bookish flavor and is often used metaphorically (Vinogradova 1972: 177). Yet, a few questions remain.

Can PRE- express 11.DIVIDE in the concrete spatial sense similar to PERE-? The verb *pregradit'* 'obstruct, block the way' suggests a positive response. Indeed, *pregradit'* is semantically very close to *peregorodit'* 'barricade' and even can be used in the same contexts as we saw in examples (5) and (6). Moreover, the verb *prelomit'(sja)*, parallel to *perelomit'(sja)*, has not only an abstract sense 'interpret into a new quality' (2.TRANSFORM) but also a concrete spatial meaning 'refract (a ray)', featuring PRE-11.DIVIDE.

On the other hand, the prefix PERE- can acquire a more abstract reading of 11.DIVIDE too: e.g. the verb *perebit'* (< *bit'* 'hit') can refer not only to spatial partitioning ('breaking a body part') but also to blocking a process or an activity (e.g. 'interrupt someone talking').

Summing up, it would be a simplification to claim that in submeaning 11.DIVIDE PERE- exclusively refers to spatial Landmarks like physical objects, whereas PRE- implies only abstract Landmarks like activities and processes. This generalization accounts for the tendency, but as we see, both PRE- and PERE- can contribute concrete and abstract senses of 11.DIVIDE.

Submeaning 12.DISTRIBUTE

Submeaning 12.DISTRIBUTE¹⁰¹ is the second most frequently attested use of PERE- and is not available for the loan prefix PRE-. We find this meaning of PERE- in 76 standard Russian verbs and 79 marginal occasional verbs in the corpus.

Distributive use of PERE- is very liberal and can apply to a variety of simplex bases. Predicates involved often denote spatial movement (*peretaskat'* 'drag many objects one by one' < *taskat'* 'drag, carry') or physical impact like repairing, washing, or damaging: *pereštopat'* 'darn many pieces of clothes' < *štopat'* 'darn', *perečinit'* 'repair all or many things' < *činit'* 'repair', *peremyt'* 'wash many objects' < *myt'* 'wash', *pereportit'* 'spoil many things' < *portit'* 'spoil'. Other verbs express different ways of physical contact: *perekusat'* 'bite many objects' < *kusat'* 'bite', *perecelovat'* 'kiss many objects' < *celovat'* 'kiss'. Moreover, 12.DISTRIBUTE can be expressed in verbs that denote social activities like reading, dancing and playing: *perečitat'* 'read all books or a large number of books' < *čitat'* 'read', *peretancevat'* 'dance many dances or dance with many partners' < *tancevat'* 'dance', *pereigrat'* 'perform all or many roles; play in all games' < *igrat'* 'perform a role; play'. More abstract applications of distributive PERE- involve physical, emotional, or mental experiences like illnesses, feelings, or thoughts: *perebolet'* 'be sick with many diseases' < *bolet'* 'be sick', *perečuvstvovat'* 'feel many feelings' < *čuvstvovat'* 'feel', *peredumat'* 'think of all or many things' < *dumat'* 'think'. An umbrella term for all possible predicates is the verb *peredelat'* 'do many or all things' < *delat'* 'do'.

Distributive use of PERE- is elaborately explained in cognitive terms in Janda (1986: 161): "the landmark is multiplied a finite number of times. The set (LM₁, LM₂, ... LM_n) represents all of a series of objects, each of which is subjected to the action of the

¹⁰¹ This submeaning is called SERIATIM in Janda 1986: 161 and Flier 1984: 143 and distributive *pere-* in Isačenko 2003/1965: 287-289; Kagan 2013: 513; Tatevosov 2013.

verb. The landmarks are dealt with one after the other and these separate units are summed up as one large landmark, all of which has been affected.” Therefore, the set of affected objects can be marked with quantifiers either in plural (*vse* ‘all’), counting the individual objects, or in singular (*vsě* ‘everything’), taking all affected objects as a single exhaustive set. Both ways of marking the Landmark are present in the example (38):

- (38) *Naverno, vsě na svete perečital, vsech pereslušal, vsě peresmotrel i perežil.*
 [L. Gurčenko. *Aplodismenty* (1994-2003)]
 ‘He **has** probably **read** everything in the world, **listened** to everybody, **seen** and **experienced** everything.’

Multiplication of the Landmark is a phenomenon in prefixal semantics not exclusive to PERE-. A similar use is possible for the prefix O(B)- (cf. Chapter 5) in submeaning 9.AFFECT A “CIRCLE” OF OBJECTS realized in verbs like *obzvonit* ‘call a “circle” of people’ (< *zvonit* ‘call’) and *oprostit* ‘interview a “circle” of people’ (< *prosit* ‘ask’). The difference between O(B)- and PERE- regarding a multiplied Landmark can be explained in the following way: O(B)- indicates that all objects that belong to a certain group (“circle”) are affected, while PERE- suggests a broader scope of impact and spreads the influence to all objects that exist (39) or to a large number of objects that provides an agent with a rich experience (40):

- (39) *Vy uže vse pomidory pereščupali!* – *A vam žalko?* [E. Piščikova. *Pjatiëtažnaja Rossija* (2007)] ‘You **have** already **pinched** all tomatoes! – Have you no mercy?’
- (40) *Ja ix stol’ko perevidela, ètix myšej i krys, – nikakogo straxa u menja pered nimi ne bylo, nu nikakogo.* [L. Gurčenko. *Aplodismenty* (1994-2003)]
 ‘I **had seen** so **many** of them – mice and rats – that I was no longer afraid of them, not at all.’

Distributive use of PERE- is rather grammaticalized and is only remotely related to the spatial prototype 1.TRANSFER OVER/ACROSS. I suggest that in submeaning 12.DISTRIBUTE PERE- performs the function of an intensifier – it amplifies the meaning of the base by means of pluralization of the Landmark. In Figure 12, I visualize the contrast between two propositions.

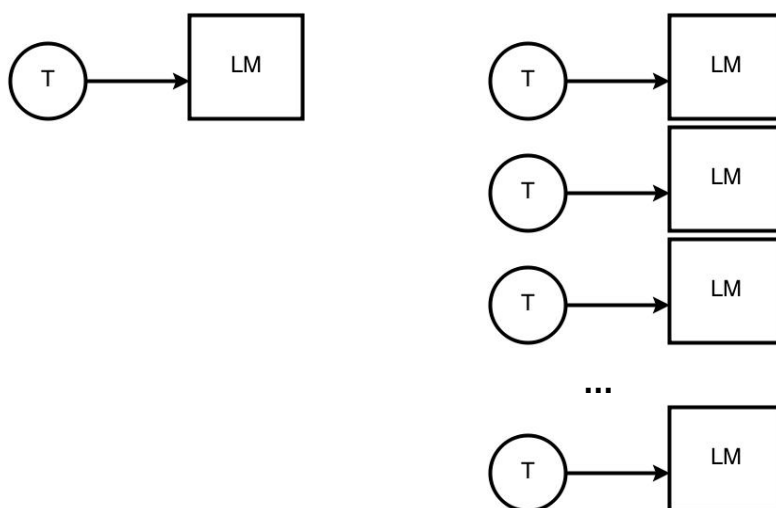


Figure 12: Submeaning 12.DISTRIBUTE as intensifier of activity.

The image to the left depicts the proposition of a verbal simplex that denotes an impact produced by the Trajector (T) on the Landmark (LM). The impact is represented by an arrow. The image to the right represents the proposition of the verbal derivative prefixed with PERE- in the sense 12.DISTRIBUTE. The prefix multiplies the entire proposition of the verbal base and in this sense intensifies its semantics quantitatively.

Submeaning 13.THOROUGH

Submeaning 13.THOROUGH of the prefix PERE- suggests a scenario similar to that of 12.DISTRIBUTE with the only difference that in 13.THOROUGH the Landmark is not subdivided into series of objects but is conceptualized as a mass. Therefore, instead of multiplication of the Landmark, the prefix increases the intensity of the impact on the Landmark. Often the same predicate can feature either submeaning 12.DISTRIBUTE or submeaning 13.THOROUGH of the prefix depending on the context. For example, in (41) the verb *perepačkat'* denotes 'stain many things' (< *pačkat'* 'stain') with PERE- 'DISTRIBUTE', whereas in (42) the same verb means 'stain one and the same object all over' and thus contains PERE- 'THOROUGH':

(41) – *Sejčas vsě perepačkaet!* – *Nalivajte vannu, kupat' ego budem.* [D. Doncova. Dollary carja Goroxa (2004)]
 '– Now he **will stain everything!** – Fill the bathtub, we will bathe him.'

(42) *S golovy do nog kapitan byl perepačkan grjaz'ju.* [N. Leontjev, A. Makeev. Grossmejster syska (2003)]
 'The captain **was stained** with dirt **all over** from top to toe.'

Similarly, 12.DISTRIBUTE and 13.THOROUGH are expressed by PERE- in verbs that refer not to physical impact as in (41) and (42) but emotional influence. For instance, depending on the context, the verb *perevolnovat'* can bear the meaning 'make many people worried' ('DISTRIBUTE') or 'make someone worried very much' ('THOROUGH'). the same holds for the predicate *perepugat'* 'frighten many people' ('DISTRIBUTE') and 'frighten someone a lot' ('THOROUGH') formed by PERE- from the base *pugat'* 'frighten'. A similar alternation of 12.DISTRIBUTE and 13.THOROUGH readings of the same verb we can find in novel coinages attested in the corpus: compare the predicate *pereissledovat'* formed from the simplex *issledovat'* 'investigate' with PERE- of 12.DISTRIBUTE in (43) and PERE- of 13.THOROUGH in (44):

(43) *Ja stal rabotat' dal'se, pereissledoval mnogo alkaloidov i drugogo roda veščestv.* [A. Ja. Danilevskij. Issledovanija nad spinnym i golovnym mozgom ljaguški (1866)¹⁰²]
 'I began to work further and **examined** a lot of alkaloids and substances of other kinds.'

¹⁰² Although this particular example comes from the 19th century and is therefore not included in the database, the phrasing is very up-to-date and, to my intuition, could be produced by a modern speaker today as well.

- (44) *Literaturnyj èpizod v «Bednyx ljudjax» uže tak začitán nami i pereissledovan; no, kak byvaet <...>, obširnyj smysl ego i kolossal'nost' roli ostajutsja ne vskrytymi.* [S.G. Bočarov. Vstupitel'nye slova. (2000)]
 'The episode of "Poor Folk" has been read to tatters and so **thoroughly investigated** by us, but its meaning and role remain unrevealed.'

Similarly to 12.DISTRIBUTE, in submeaning 13.THOROUGH PERE- functions as an intensifier of activity denoted by the verbal base. Figure 13 shows the contrast between the proposition of the simplex base (image to the left) and the proposition of the prefixed derivative with PERE- 'THOROUGH' (image to the right). An arrow represents the impact of the Trajector (T) on the Landmark (LM).

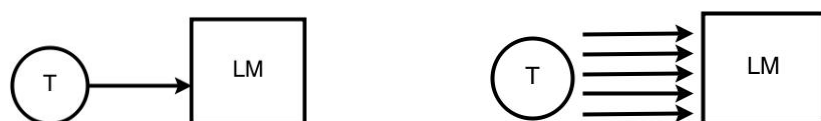


Figure 13: Submeaning 13.THOROUGH as intensifier of activity.

In the proposition of the prefixed verb multiple arrows visualize the increased intensity of the impact which is the semantic contribution of the prefix.

Submeaning 13.THOROUGH of PERE- is typically manifested in verbs that denote physical impact on the look of an object as in *peremjat'* 'rumple something a lot' (< *mjat'* 'rumple'), *peremazat'* 'smear a lot' (< *mazat'* 'smear'), *peremarat'* 'stain all over' (< *marat'* 'stain'). Another type of verbs that represent PERE- 'THOROUGH' are lexemes that refer to negative emotional experiences of fear, exhaustion, and uneasiness: *peretrusit'* 'be afraid very much' (< *trusit'* 'be afraid'), *peretrevožit'* 'make worried a lot' (< *trevožit'* 'make worried'), *perekonfuzit'* 'confuse to a large degree' (< *konfuzit'* 'confuse'), *peremajat'* 'exhaust a lot' (< *majat'* 'exhaust'), *perebalamutit'* 'trouble a lot' (< *balamutit'* 'trouble'), and *perepoložit'* 'trouble someone a lot' (< (*vs*)*položit'* 'trouble').

The prefix PRE- does not express submeaning 13.THOROUGH.

Submeaning 14.VERY

By contrast, submeaning 14.VERY is attested only for the prefix PRE-. We can observe this use of PRE- in three standard predicates *preuspet'* 'succeed in' (< *uspet'* 'manage to do in time'), *preumnožit'(sja)* 'increase a lot' (< *umnožit'* 'increase, multiply'), and *preispolnit'(sja)* 'fill up to the maximal degree' (< *ispolnit'(sja)* 'fill, fulfill') and also in two marginal verbs *preizobilovat'* 'abound in a large degree' (< *izobilovat'* 'abound') and *preukrasit'* 'decorate a lot' (< *ukrasit'* 'decorate'). These verbs do not bear a negative connotation and do not refer to overdoing an activity.

While the previously discussed submeanings 3.SUPERIORITY, 4.OVERDO, 5.REDO, 12.DISTRIBUTE, and 13.THOROUGH realize intensification of activity, submeaning 14.VERY brings about intensification of a property. That is why this meaning is very rarely manifested in verbs and is so productive beyond the verbal domain – in those grammatical classes that lexicalize properties: adjectives and adverbs (e.g. *premilij* 'very nice' < *milij* 'nice'; *premnogo* 'very many' < *mnogo* 'many').

A closer look at verbs with PRE- meaning 'VERY' reveals their peculiar semantics. The verb *preumnožit'* is a factitive and is associated not only with the verb *umnožit'* 'multiply' but also with the adverb *mnogo* 'a lot of'. *Preumnožit'* bears the meaning 'make

something increase in quantity' rather than 'multiply several times', so I conclude that the prefix in this verb denotes an increase in property (quantity) rather than intensification of the activity named by the verbal base. Similarly, *preispolnit'* implies that a container becomes more filled, and *preuspet'* suggests that a person becomes more successful. These verbs also encode intensified resulting properties rather than intensified activities.

Recall that for other meanings of intensification 3.SUPERIORITY, 4.OVERDO, and 5.REDO – there is a point of reference, or a standard of comparison, in terms of Kagan 2013. 3.SUPERIORITY compares the performance of an activity carried out by the subject with its performance by someone else (a competitor). 4.OVERDO compares the performance carried out by the subject with what a norm suggests. 5.REDO puts the result of repeated action in correspondence with a previous performance, because the action is repeated in order to achieve a better result. 12.DISTRIBUTE and 13.THOROUGH do not imply a comparison of this kind but it is possible to contrast the proposition of the verbal base with the proposition of the derivative verb intensified by the prefix. I suggest that 14.VERY is different in that it constitutes an end-point metonymy of the prototype 1.TRANSFER OVER/ACROSS: the Trajector does not proceed over the Landmark, and the focus is restricted to the resulting state of being “over” some norm, i.e. into the state of exceeding a neutral degree of a property. Thus, 14.VERY is a metaphoric extension from the domain of spatial relations to the domain of properties and states. Also, 14.VERY corresponds to the end point of the trajectory that transgresses over the Landmark.¹⁰³

A different morphological context modifies the meaning of the prefix. In adjectives and adverbs – the word classes that denote gradable qualities – the meaning 14.VERY of the prefix PRE- is more regular and grammaticalized than it is in the domain of verbal derivation. This is not surprising, because adjectives and adverbs are those grammatical classes that are specialized for the expression of quality and where the function of intensifier (marker of a high degree) is most appropriate and welcome.

Summing up, the analysis I have pursued in this section suggests that the two prefixes in question are semantically comparable. Both PERE- and PRE- are highly polysemous and share a network of interrelated submeanings, where eight of fourteen submeanings can be expressed by both prefixes, and the remaining six submeanings are specific for either PERE- or PRE-.

In other words, the two prefixes exhibit a partial semantic overlap and a partial divergence. This makes it difficult to evaluate their mutual semantic identity according to the semantic criterion of Standard Allomorphy. If we take into account the entire semantic network that these prefixes represent, it becomes clear that PERE- and PRE- are neither completely identical in meaning nor absolutely different. One has to consider

¹⁰³ An additional argument in favor of the proposed semantic connection between the excessive 14.VERY and the spatial sense 1.‘TRANSFER OVER/ACROSS’ comes from linguistic typology. The same combination of meanings expressed by a single cognate linguistic formant is attested in other Slavic languages as well as in Latin and Ancient Greek. The Slavic prefixes related to the Church Slavonic PRE- – *prze-* in Polish, *pře-* in Czech, and *pre-* / *prije-* in Serbo-Croatian, according to Herman (1975), can express both ‘transition, transfer, transformation’ and ‘extreme degree, intensity, superiority, and act of exceeding’. In Latin, the formant *per-*, historically related to PRE-, served as a preposition meaning ‘cross over’ and as a prefix of intensification, with the meaning ‘very’ (e.g. *permagnus* ‘very big’), while in Ancient Greek the same combination of meanings was attested for *peri* (Černyx 1993: v.2, 21).

the degree of their commonality. Since those semantic subcategories that are shared by both prefixes include the prototype and yield a total of eight submeanings, they account for the largest part of the semantic network. We can conclude that PERE- and PRE- are arguably more similar to each other than different. In terms of individual lexemes that represent these eight shared submeanings in the standard Russian lexicon, there are 243 verbs in PERE- (55% of all verbs in PERE-) and 34 verbs in PRE- (87% of all verbs in PRE-).

The overall general picture can be further sharpened by means of a quantitative analysis of this distribution. In the next section I focus on the Radial Category Profiles of the two prefixes and test their semantic overlap and mismatch for statistical significance.

6.4 Radial Category Profiling: PERE- vs. PRE- in the domain of verbs

In this subsection I focus on the following question: what submeanings are the most productive and prominent for PERE- and PRE-? I will show that the submeanings form semantic blocks that reflect the distinct semantic profiles of the two prefixes.

The distribution of all verbs in the database across the two prefixes in question is summarized in Table 2 and visualized in Figure 14.

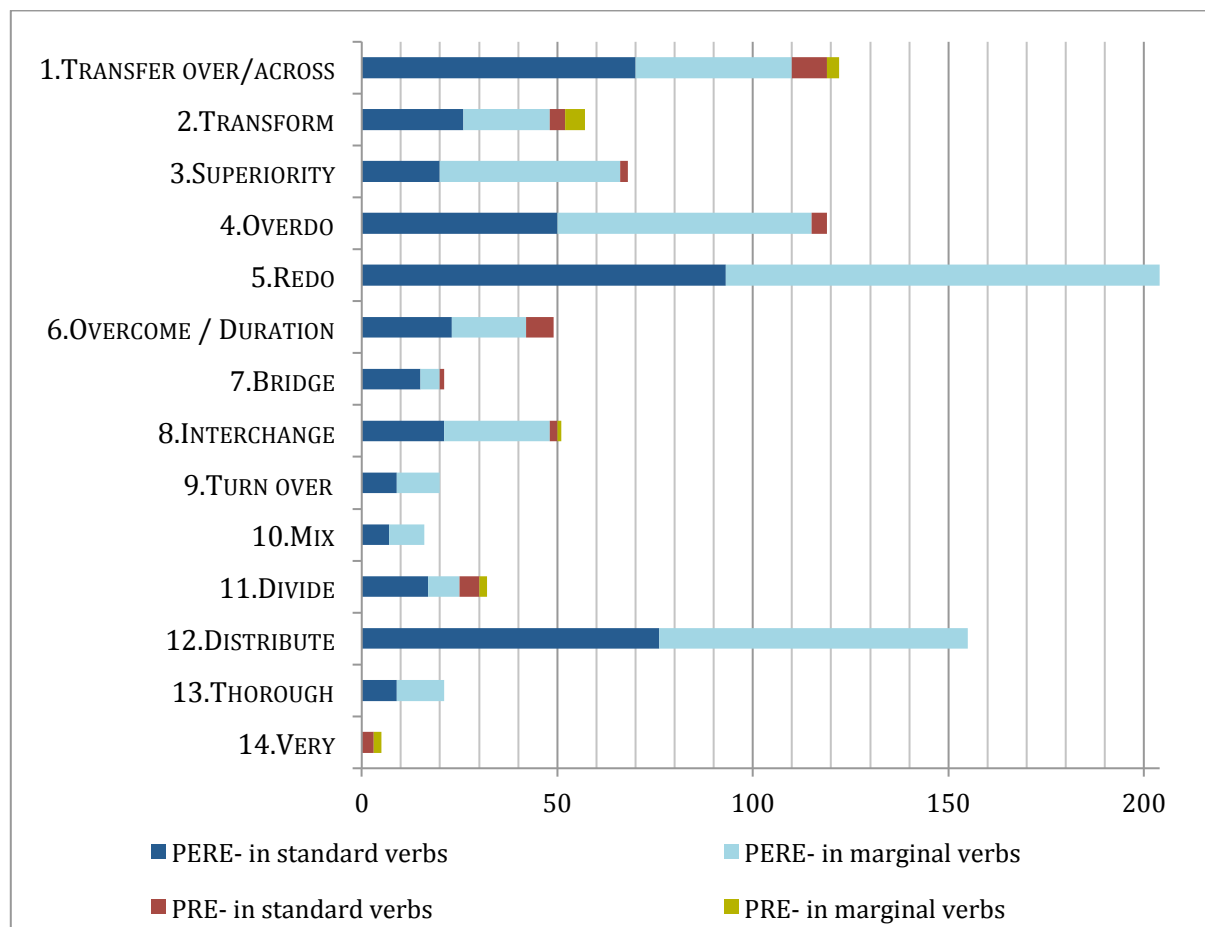


Figure 14: Distribution of verbs with PERE- and PRE- in the database (raw numbers).

	PERE- standard	PERE- marginal	PRE- standard	PRE- marginal
1.TRANSFER OVER/ACROSS	70	40	10	3
2.TRANSFORM	26	22	4	5
3.SUPERIORITY	20	46	2	0
4.OVERDO	50	65	4	0
5.REDO	92	111	0	0
6.OVERCOME / DURATION	23	19	7	0
7.BRIDGE	15	5	1	0
8.INTERCHANGE	22	27	2	1
9.TURN OVER	10	11	0	0
10.MIX	7	9	0	0
11.DIVIDE	17	8	4	2
12.DISTRIBUTE	76	79	0	0
13.THOROUGH	9	12	0	0
14.VERY	0	0	3	2
Deetymologized	0	0	2	2
Total:	437	454	39	15

Table 2: Distribution of verbs with PERE- and PRE- in the database (raw numbers).

I suggest that the fourteen submeanings form three groups, as shown in Table 3. The first group combines spatial submeanings 1.TRANSFER OVER/ACROSS, 7.BRIDGE, 9.TURN OVER, 10.MIX, and 11.DIVIDE. The second group includes those submeanings that intensify semantics of the base: 5.REDO, 12.DISTRIBUTE, 4.OVERDO, 3.SUPERIORITY, 13.THOROUGH, and 14.VERY. The third group is more heterogeneous and consists of the remaining submeanings that are neither spatial nor intensifying but all metaphorical: 2.TRANSFORM, 6.OVERCOME/DURATION, and 8.INTERCHANGE.

Semantic Group	Submeanings	PERE- standard	PERE- marginal	PRE- standard
Spatial domain	1.TRANSFER OVER/ACROSS 7.BRIDGE 9.TURN OVER 10.MIX 11.DIVIDE	119 (27%)	73 (16%)	15 (41%)
Intensification	3.SUPERIORITY 4.OVERDO 5.REDO 12.DISTRIBUTE 13.THOROUGH 14.VERY	247 (57%)	313 (69%)	9 (24%)
Other (metaphorical)	2.TRANSFORM 6.OVERCOME / DURATION 8.INTERCHANGE	71 (16%)	68 (15%)	13 (35%)
Total		437 (100%)	454 (100%)	37 (100%)

Table 3: Distribution of verbs in PERE- and PRE- across semantic groups of submeanings.

In Table 3, I present quantities of verbs for each of three semantic groups with regard to standard verbs prefixed with PERE- and PRE- indicative of the lexicon and marginal verbs prefixed with PERE-, characteristic of productivity. I omit marginal verbs in PRE- as they are very sparse and not informative for the general picture that I am interested in here.

I suggest that the distribution of verbs across these groups of submeanings highlight the centers of gravity that PERE- and PRE- have within a network of polysemy (Radial Category). I follow Nessel et al. 2011 and apply the methodology of Radial Category Profiling, where a Radial Category Profile of a prefix is its relative frequency distribution across the subcategories of a radial network.

Shading in Table 3 reflects the profiles of the two prefixes: in verbs, PERE-'s center of gravity is located in the group of intensification submeanings, whereas PRE-'s center of gravity is in the submeanings of the spatial domain and in the metaphorical submeanings other than intensification. Different centers of gravity that PERE- and PRE- have in terms of semantics frame a pattern of their distribution across verbs.

In Figure 15, I visualize the numbers presented in Table 3 and show the profiles of PERE- and PRE- according to their use in standard Russian verbs.

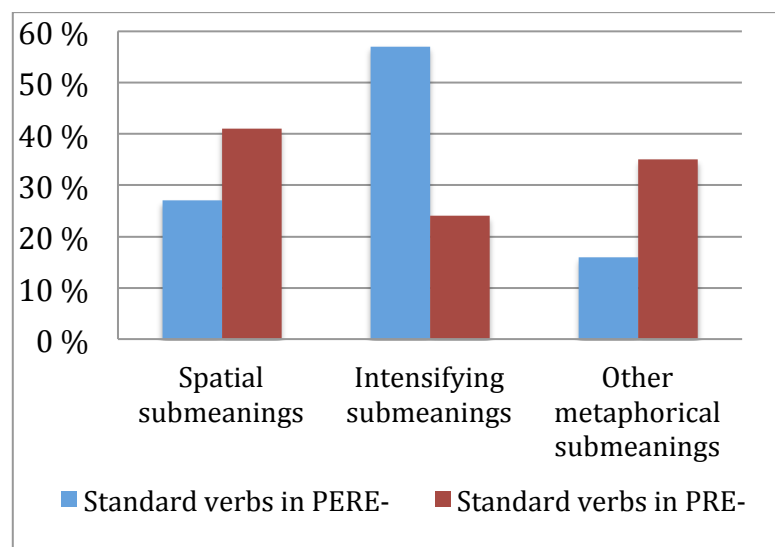


Figure 15: Radial Category Profiles of the prefixes PERE- and PRE- (%).

A chi-square test shows that the difference between the distributions of PERE- and PRE- across three semantic groups (Spatial, Intensifying, Other) is statistically significant ($X^2 = 15.6$, $df = 2$, $p\text{-value} < 0.001$) and has a small-to-medium effect size (Cramer's $V = 0.2$). This suggests that in standard lexemes PERE- and PRE- have distinct Radial Category Profiles.

The difference between the distributions of PERE- in standard vs. marginal verbs is also found statistically significant ($X^2 = 18.5$, $df = 2$, $p\text{-value} = 9.39e-05$), with a somewhat smaller effect size (Cramer's $V = 0.14$). This indicates that standard and marginal verbs in PERE- differ significantly in terms of what submeanings of this prefix they tend to employ. In particular, in marginal verbs the intensifying submeanings of PERE- are even more frequently attested than in standard verbs with this prefix, whereas spatial submeanings of PERE- are less frequently employed in marginal new coinages as opposed to standard verbs. Thus, the outcome of the statistical analysis justifies our distinction of standard vs. marginal lexemes, as they behave differently in terms of realization of PERE-'s submeanings.

An interesting outcome of Radial Category Profiling is the generalization about the use of PERE-. Semantic grouping of submeanings indicates that in standard verbs PERE- is more frequently attested in intensifying submeanings than in spatial submeanings in terms of individual lexemes (compare 57% vs. 27% respectively in Table 3), and this tendency is even more robust in marginal verbs: 69% vs. 16% respectively.

I suggest that there are five submeanings in which the prefix PERE- functions as an intensifier of the activity denoted by the verbal base: 5.REDO, 12.DISTRIBUTE, 4.OVERDO, 3.SUPERIORITY, and 13.THOROUGH, ordered according to their productivity. In Figure 16, these submeanings are highlighted by rectangles with thick frames.

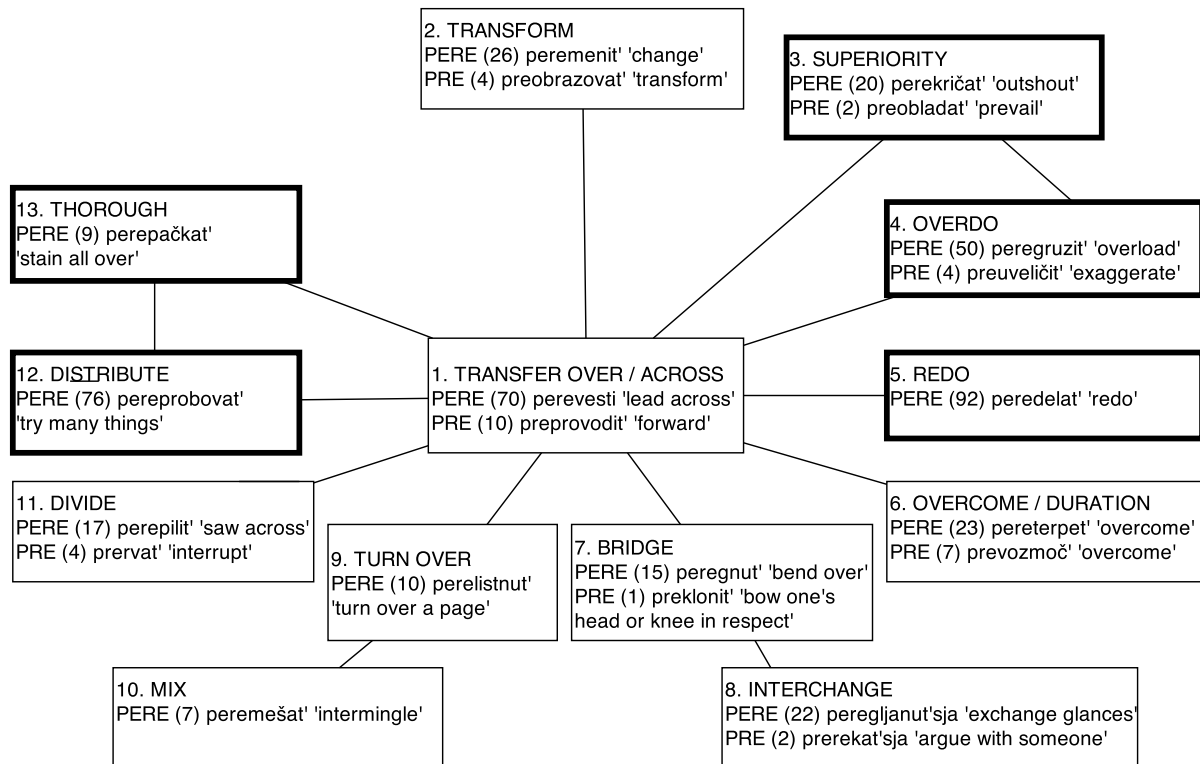


Figure 16: Prefix PERE- as a highly productive verbal morphological intensifier of activity.

In Table 4, I list the intensifying submeanings of PERE- and provide their type frequencies – the numbers of standard (first number) and marginal (second number) lexemes that represent each of these submeanings in parentheses.

This data shows that PERE- functions as a productive intensifier of activity in the verbal domain. This function accounts for more than half of PERE-'s uses in the verbs attested in the corpus. Overall, the prefix PERE- performs the function of intensifier in 247 standard lexemes (57% of standard verbs in PERE-) and 313 marginal verbs (69% of marginal verbs in PERE-).

Submeaning	Description	Examples
REDO (92; 111)	'repeat an activity to achieve a better result'	<i>peredelat</i> 'redo' < <i>delat</i> 'do'
DISTRIBUTE (76; 79)	'apply an activity to a number of objects'	<i>pereprobovat</i> 'try many things' < <i>probovat</i> 'try'
OVERDO (50; 65)	'perform an activity more than a norm suggests'	<i>peregruzit</i> 'overload' < <i>gruzit</i> 'load'
SUPERIORITY (20; 46)	'perform an activity better than someone else'	<i>perekričat</i> 'outshout' < <i>kričat</i> 'shout'
THOROUGH (9; 12)	'thoroughly affect with an activity the whole object'	<i>perepačkat</i> 'stain all over' < <i>pačkat</i> 'stain'

Table 4: Submeanings of the prefix PERE- where it intensifies an activity denoted by the base.

Interestingly, some verbal bases can employ the intensifying PERE- in different submeanings. Consider the prefixed verb *pereigrat* 'PERE-play' which has four different readings depending on the sense of the prefix: 'win over someone' with PERE- 3.'SUPERIORITY'; 'overact' with PERE- 4.'OVERDO'; 're-play, play anew' with PERE- 5.'REDO'; and 'perform all or many roles' with PERE- 12.'DISTRIBUTE'. The ambiguity is always resolved in the context.

Note that intensification of an activity is not exclusive for PERE-. Two of its submeanings of intensification can be expressed by the prefix PRE- too: 3.SUPERIORITY and 4.OVERDO. However, if we sum up verbal predicates in which PRE- intensifies the meaning of the base (submeanings 3.SUPERIORITY, 4.OVERDO, 14.VERY) this yields 9 lexemes, or only 24% of those few verbs where this prefix remains in Modern Russian. The function of intensification is performed by this prefix less in the verbal domain but primarily in the classes of adjectives and adverbs (see Section 6.6). In verbs the prefix PRE-, by contrast, is more frequent in spatial meanings, first of all 1.TRANSFER (9 standard lexemes), 11.DIVIDE (5 standard lexemes), and metaphorical extensions of the prototype – submeanings 6.OVERCOME / DURATION (7 standard lexemes) and 2.TRANSFORM (4 standard lexemes).

Note that the two most productive intensification submeanings of PERE- (see Figure 16) – 5.REDO and 12.DISTRIBUTE – are not available for the loan prefix PRE-. I suggest that both of these submeanings feature quantitative intensification of an activity: in these submeanings the prefix PERE- reduplicates and multiplies the proposition of the base. On the contrary, the other two intensification meanings of PERE- – 3.SUPERIORITY and 4.OVERDO – which are attested for the loan prefix PRE- arguably suggest rather qualitative intensification of an activity. These submeanings focus on the qualitatively different outcome of the subject's performance. This might be the reason why 3.SUPERIORITY and 4.OVERDO can be expressed by the prefix PRE- too: they might be conceptually closer to intensification of properties, typical for PRE- in adjectives and adverbs.

6.5 Grammatical properties of PERE- vs. PRE-: Aspect and Transitivity

In this section I will compare the two prefixes in terms of their grammatical properties. In particular, I am concerned with the following question: how often is the attachment of

PERE- or PRE- accompanied with a change of aspectual characteristics and transitivity of the verbal base? In addition, I will also examine the phenomenon of prefix stacking and how it correlates with the shift of aspect and transitivity.

In order to address these issues, I will consider only standard verbs that have a clear verbal base – that is 418 lexemes prefixed in PERE- and 36 lexemes in PRE-. I set aside factitive verbs formed by these prefixes from non-verbal bases like *pereuprjamit'* 'outdo someone in being stubborn' (< adjective *uprjamyj* 'stubborn'), *pereinačit'* 'change' (< adverb *inače* 'differently'), and *prevysit'* 'exceed' (< noun *vys* 'height'). I also exclude from this analysis two deetymologized verbs in PRE- *presmykat'sja* 'creep' and *preminut'* 'fail to'. In total, the data that was set aside yields 5% of standard verbs prefixed in PERE- (21 lexemes) and 7% of standard verbs in PRE (3 lexemes).

6.5.1 Shift in aspect

Do PERE- and PRE- differ in terms of how they can affect the aspect of the base verb? I address this question quantitatively and assume that the more verbs a prefix can turn from an imperfective (IPF) to perfective (PF), the stronger perfectivizer it is.

The distribution of data is shown in Table 5 and is visualized in Figure 17. They aggregate raw values and also present a percentage ratio. The numbers of verbs indicate that PERE- is over two times more "aspectual" as a prefix, or a better perfectivizer, than its loan Slavonic counterpart PRE-. PERE- shifts aspect of the base verb in 76% cases, whereas PRE- does so only in 36% cases: e.g. *plyt'*_{IPF} 'swim' > *pereplyt'*_{PF} 'swim across', *seč'*_{IPF} 'cut' > *preseč'*_{PF} 'stop'.

Shift in aspect	Subtype	Verbs in PERE-		Verbs in PRE-	
Shift	IPF -> PF	316	76%	13	36%
No shift	IPF -> IPF	12	24%	5	64%
	PF -> PF	90		18	
Total		418	100%	36	100%

Table 5: Prefixes PERE- and PRE- with regard to aspectual shift in the derivatives.

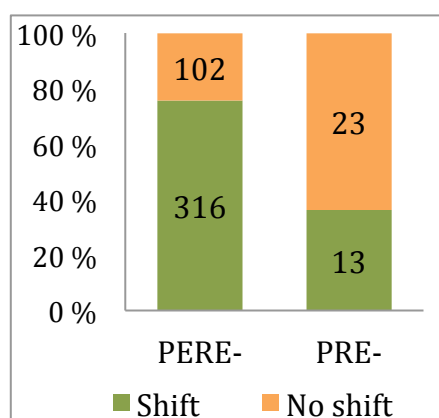


Figure 17: Prefixes PERE- and PRE- with regard to aspectual shift in the derivatives.

PERE- does not shift the aspect of the base in 24% cases. Mostly, this concerns perfective bases that maintain the perfective aspect under attachment of the prefix (PF -> PF), as in *brosit'*_{PF} 'throw' > *perebrosit'*_{PF} 'throw over' and *dat'*_{PF} 'give' > *peredat'*_{PF} 'pass over'. In these verbs the prefix PERE- appears in different submeanings. The most frequent of them is 1.TRANSFER OVER/ACROSS illustrated above and 5.REDO: *izdat'*_{PF} 'publish' > *pereizdat'*_{PF} 're-publish'. The few verbs in PERE- that remain imperfective after the prefix is attached are formed via the morphological construction *pere-...-yva-t'-sja* with submeaning 8.INTERCHANGE: *šeptat'*_{IPF} 'wisper' > *perešeptyvat'sja*_{IPF} 'exchange wispers'. Strictly speaking, because there is an imperfective suffix *-yva-* present in these verbs, PERE- is not responsible for the imperfective aspect of the derivatives.

By contrast, the loan prefix PRE- does not shift the aspect of the base almost three times more frequently than PERE- – in 64% cases. Among them, there are 18 verbs that remain perfective (e.g. *vzojti*_{PF} ‘mount’ > *prevzojti*_{PF} ‘excel in’) and 5 verbs that remain imperfective (e.g. *obladat*_{IPF} ‘possess’ > *preobladat*_{IPF} ‘prevail’). The latter type is unique to PRE-, as there are no such examples for PERE-.

The difference between the distributions of PERE- and PRE- with regard to aspectual shift is statistically significant¹⁰⁴. This suggests that our generalization is robust and PERE- functions as a much stronger perfectivizer of verbs than the loan prefix PRE-.

6.5.2 Shift in transitivity

Do PERE- and PRE- affect transitivity of the base verb? Are they different in this regard? Answering these questions, I exclude from my calculations all reflexive verbs that lack a non-reflexive counterpart, because they contain a morphological intransitivizer *-sja* which overrules any possible effect of the prefix.

Table 6 aggregates the numbers of verbs that characterize the distribution of PERE- and PRE-. Figure 18 visualizes the same data in two bars.

Shift in transitivity	Subtype	Verbs in PERE-		Verbs in PRE-	
Shift	INTR -> TR	33	9%	4	15%
	TR -> INTR	4		1	
No shift	INTR -> INTR	40	91%	4	85%
	TR -> TR	313		24	
Total		390	100%	33	100%

Table 6: Prefixes PERE- and PRE- with regard to transitivity shift in the derivatives.

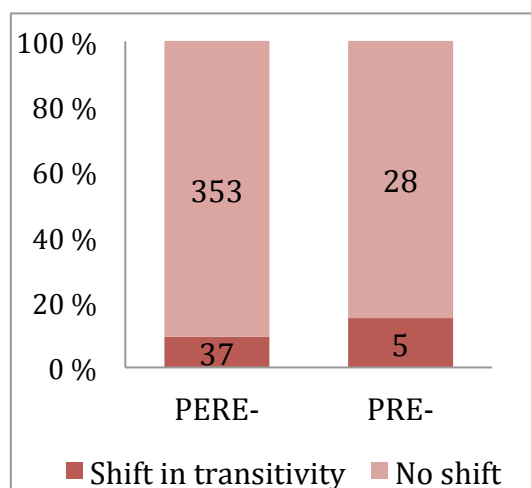


Figure 18: Prefixes PERE- and PRE- with regard to shift in transitivity.

The difference between the two prefixes appears to be very small with regard to how often they shift or do not shift transitivity of the verbal base. Both prefixes mostly fail to affect transitivity: in 91% verbs with PERE- and 85% verbs with PRE- the attachment of the prefix is not accompanied with any changes of transitivity. This means that intransitive verbs mostly remain intransitive, as in *goret*_{INTR} ‘burn’ > *peregoret*_{INTR} ‘burn out’ and *obladat*_{INTR} ‘possess’ > *preobladat*_{INTR} ‘prevail’. Likewise, transitive verbs mostly remain transitive, as in *ubedit*_{TR} ‘persuade’ > *pereubedit*_{TR} ‘change someone’s opinion’ and *uveličit*_{TR} ‘increase’ > *preuveličit*_{TR} ‘exaggerate’.

However, in a few cases the attachment of PERE- and PRE- is accompanied with a shift in transitivity. In the majority of such shifts, an intransitive base verb turns into a transitive derivative. This happens when the prefix

¹⁰⁴ I applied the Fisher’s Exact Test to the raw values shown on the bars in Figure 17. According to this analysis, p-value = 2.093e-06, confidence interval is 2.6 to 12.2, odds ratio 5.4.

introduces an argument which lexicalizes a spatial or metaphorical Landmark, i.e. an obstacle that is situated on the way of the Trajector. Therefore, an intransitive base verb often turns into a transitive, when the prefix denotes 1.TRANSFER OVER/ACROSS: *bežat'*_{INTR} 'run' > *perebežat'*_{TR} 'run across', *stupit'*_{INTR} 'step' > *prestupit'*_{TR} 'transgress, violate', *l'stit'*_{INTR} 'flatter' > *prel'stit'*_{TR} 'lure someone'. Similarly, the submeaning 3.SUPERIORITY also requires a direct object argument that lexicalizes a competitor in an activity: *xitrit'*_{INTR} 'use cunning' > *perexitrit'*_{TR} 'outsmart someone', *vzajt'*_{INTR} 'mount' > *prevzajt'*_{TR} 'outdo someone in smth'. In addition, many verbs with PERE- of 6.OVERCOME/DURATION imply a direct object that denotes a period of time that one has to overcome or wait through: *zimovat'*_{INTR} 'pass the winter' > *perezimovat'*_{TR} *zimu* 'pass the winter', *sidet'*_{INTR} 'sit' > *peresidet'*_{TR} (*obstrel*) 'sit safe during (e.g. shooting attack)'.

Very rarely, a transitive base verb becomes an intransitive derivative: *uspet'*_{TR} 'manage to do in time' > *preuspet'*_{INTR} 'succeed in', *govorit'*_{TR} 'talk, say smth' > *peregovorit'*_{INTR} 'talk over about smth'. These cases are rather exceptional and constitute very few examples.

Overall, the difference between the distributions of PERE- and PRE- with regard to changes in transitivity of the base verb is small and not statistically significant¹⁰⁵. Crucially, both prefixes can turn an intransitive verb into a transitive one in particular submeanings that imply a lexicalized obstacle. Yet, the relevant examples in PERE- are more numerous than the ones in PRE-.

6.5.3 Prefix stacking

Prefix stacking takes place when a verb attaches a prefix over another one. Among verbs with stacked prefixes, PRE- is generally more common than PERE-: *pre-voz-nesti* 'extol' and *pre-pod-nesti* 'present with'. However, examples with PERE- exist as well: *pere-iz-brat'* 're-elect' and *pere-na-selit'* 'overpopulate'. The stacked PERE- and PRE- are distributed across verbs as shown in Figure 19.

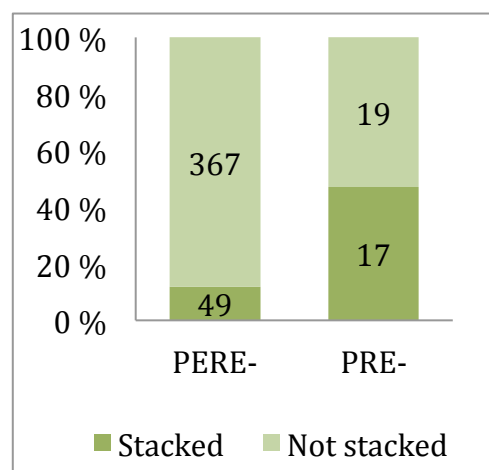


Figure 19: Prefixes PERE- and PRE- with regard to prefix stacking.

PERE- is stacked only in 12% of standard verbs, whereas PRE- stacks in 47% of verbs. The Fisher's Test suggests that the difference between PERE- and PRE- in terms of stacking is highly significant¹⁰⁶.

The next question is: does prefix stacking correlate with aspectual shift and transitivity shift in prefixed verbs? I summarize the distributions of data according to these parameters in Table 7 and Table 8. Table 7 shows that in most verbs where PERE- and PRE- stack over another prefix they fail to alter the aspect of the base. However, when PERE- and PRE- are not stacked over another prefix, they usually shift the aspect to perfective: 313 out of 369 lexemes with PRE- and 12 out of 19 verbs with PERE-.

¹⁰⁵ According to the Fisher's Exact Test applied to raw numbers shown in Figure 18, p-value = 0.4, confidence interval is 0.2 to 2, odds ratio 0.6.

¹⁰⁶ P-value = 8.65e-07, confidence interval 0.07 to 0.3, odds ratio 0.15.

Prefix stacking	Subtype	Verbs in PERE-		Verbs in PRE-	
Stacked	Aspect shift	3	49 (12%)	0	17 (47%)
	No aspect shift	46		17	
Not stacked	Aspect shift	313	369 (88%)	12	19 (53%)
	No aspect shift	56		7	
Total		418	418 (100%)	36	36 (100%)

Table 7: Stacking of PERE- and PRE- and shift in aspect of the base.

I subjected these distributions of verbs¹⁰⁷ to the Fisher’s Exact Test and found that the difference between them is statistically highly significant¹⁰⁸. This suggests that stacking of the prefixes PERE- and PRE- correlates strongly with their ability to shift the aspect of the base verb. In particular, these prefixes are more likely to shift aspect if they are not stacked.

Table 8 addresses the other possible correlation – prefix stacking and transitivity shift. It aggregates the numbers of verbs where the stacked or non-stacked prefix is or is not accompanied with the shift in transitivity.

Prefix stacking	Subtype	Verbs in PERE-		Verbs in PRE-	
Stacked	Transitivity shift	0	49 (12%)	2	17 (47%)
	No transitivity shift	49		15	
Not stacked	Transitivity shift	37	369 (88%)	3	19 (53%)
	No transitivity shift	304		13	
	Not applicable	28		3	
Total		418	418 (100%)	36	36 (100%)

Table 8: Stacking of PERE- and PRE- and transitivity shift.

Table 8 demonstrates that the shift in transitivity is attested more frequently in verbs where PERE- and PRE- are not stacked over another prefix: compare 37 and 3 verbs in PERE- and PRE- respectively in case they are not stacked with 0 and 2 verbs in PERE- and PRE- in case they are stacked. I tested the difference in distributions of verbs depending on prefix stacking and transitivity shift for statistical significance¹⁰⁹, and the Fisher’s Exact Test¹¹⁰ pointed to a positive response. We can conclude that stacking of PERE- and PRE- over another prefix holds a robust correlation with their ability to affect transitivity of a base verb.

It is noteworthy that the p-value in this analysis is much larger than it is in the case of correlation between stacking and aspect. This means that the correlation of aspect shift and prefix stacking is stronger and more robust than the correlation of prefix stacking with shift in transitivity.

¹⁰⁷ The numbers that I tested are based on Table 7 and involve the following opposition: prefix stacking with aspect shift (3+0=3) and with no aspect shift (46+17=53) vs. no prefix stacking with aspect shift (313+12=325) and with no aspect shift (56+7=63).

¹⁰⁸ P-value < 2.2e-16, confidence interval 0.002 to 0.04, odds ratio 0.01.

¹⁰⁹ Similarly to my calculations based on Table 7, I addressed the following distribution: prefix stacking with transitivity shift (0+2=2) and with no transitivity shift (49+15=64) vs. no prefix stacking with transitivity shift (37+3=40) and with no transitivity shift (304+13=317). Note that I leave aside the verbs where the analysis of transitivity shift is not applicable (due to the presense of the reflexive marker *-sja* in the base and the derivative) – a total of 31 lexemes.

¹¹⁰ P-value = 0.04, confidence interval 0.03 to 1, odds ratio 0.2.

When PERE- and PRE- are stacked over another prefix, they appear in different submeanings. Table 9 provides an overview of their different semantic uses in the stacked position. Eight submeanings are represented. Most verbs with stacked PERE- and PRE- employ these prefixes in the intensifying function in submeanings 5.REDO and 4.OVERDO, which account for a total of 43 verbs. In addition, we observe other intensifying submeanings of stacked PERE- and PRE-: 3.SUPERIORITY, 12.DISTRIBUTE, and 13.VERY.

Submeaning	# verbs in PERE-: Example	# verbs in PRE-: Example
1.TRANSFER OVER/ACROSS	4: <i>peredoverit'</i> 'delegate'	6: <i>preprovodit'</i> 'forward'
2.TRANSFORM	3: <i>perevoplotit'</i> 'reincarnate'	1: <i>preobrazovat'</i> 'transform'
3.SUPERIORITY	0	2: <i>prevzoiiti</i> 'surpass'
4.OVERDO	10: <i>pereutomit'</i> 'overdrive'	2: <i>preuveličit'</i> 'exaggerate'
5.REDO	31: <i>pereosmyslit'</i> 'rethink'	0
6.OVERCOME/DURATION	0	3: <i>preodolet'</i> 'overcome'
12.DISTRIBUTE	1 <i>perezabyt'</i> 'forget lots of'	0
13.VERY	0	3: <i>preumnožit'</i> 'increase'
Total	49	17

Table 9: Submeanings of stacked prefixes PERE- and PRE-.

Interestingly, the stacked prefixes in question can also express a spatial submeaning 1.TRANSFER OVER/ACROSS as well as metaphorical submeanings 2.TRANSFORM and 6.OVERCOME/DURATION. This contradicts the traditional assumption that only superlexical prefixes can stack over other prefixes (Recall the discussion of this issue in Chapter 4, Section 4.6.3; cf. Svenonius 2004: 206, 229; Romanova 2004: 257, 265).

Summing up the discussion of grammatical properties, it is reasonable to conclude that PERE- is a more "verbal" prefix than PRE-, as it is a stronger perfectivizer. Regarding transitivity, the difference between PERE- and PRE- is found insignificant. A statistical analysis suggests that stacking of PERE- and PRE- over another prefix decreases their ability to affect the verbal morphological categories of aspect and transitivity. PRE- is stacked significantly more often than PERE-, and its strength as a perfectivizer and transitivizer of verbal bases is significantly smaller.

6.6 Beyond verbs: The role of Slavonic prefix PRE-

So far we have looked in detail at the domain of verbs which features the productive use of the East Slavic prefix PERE-, native for Russian, and the unproductive use of its Slavonic equivalent PRE-, the loan prefix. In this section I address a different question: is the prefix PRE- productive in any other morphological domains of Modern Russian? What kind of role does this prefix have: is it just an obsolete morphological anachronism or does it form any new words?

As opposed to PERE-, the prefix PRE- has an additional domain of use beyond verbs – in adjectives and adverbs. This application of PRE- is much more prominent than its role in prefixed verbs. In this domain "beyond verbs", PRE- functions as a highly productive intensifier of quality which amplifies the meaning of the base. PRE- is often employed to form novel adjectives and adverbs. This use of PRE- has not been thoroughly investigated. The goal of this section is to look at it in greater detail.

6.6.1 Two domains of PRE-

All uses of the prefix PRE- are distributed across two morphological domains – the verbal and the adjectival. In verbs, the prefix PRE- is infrequent and relatively unproductive. By contrast, in the adjectival domain, this prefix is very productive and can produce a potentially open list of words. In particular, PRE- is attested in a wide range of adjectives and adverbs, and also in some nouns and even individual words of some other classes. Interestingly, these two uses of PRE- are quite different in terms of register: the verbal PRE- has a bookish flavor and is mostly preserved in Slavonic loan words which sound very elevated, literary and formal due to their origin. However, beyond verbs, the prefix PRE- is very colloquial. Finally, the two uses of PRE- differ in terms of semantics as well: in verbs this prefix demonstrates rather diverse semantics which is to some degree comparable with the polysemy of PERE-. In adjectives and adverbs, however, PRE- has very stable and consistent semantic content: it expresses a high degree of a quality and intensifies the meaning of the simplex base. These differences between the two applications of the prefix PRE- are summarized in Table 10:

Criteria	Verbal PRE-	Non-verbal PRE-
Productivity	Rare Unproductive Closed list of words	Frequently used Very productive Open list of words
Parts of speech	Verbs	Adjectives, adverbs, some nouns
Register	Bookish	Colloquial
Semantics	Polysemous 'TRANSFER' & extensions	Meaning 'VERY' Intensifier of quality

Table 10: Two facets of the prefix PRE- in Modern Russian.

6.6.2 PRE- as a productive intensifier of quality

Most frequently the prefix PRE- attaches to qualitative scalar adjectives (45) and adverbs (46). These words refer to a gradable property and imply a scale.

- (45) *predobryj* 'very kind' < *dobryj* 'kind'
prebol'soj 'very big' < *bol'soj* 'big'
- (46) *prespokojno* 'very calmly' < *spokojno* 'calmly'
predostatočno 'more than enough' < *dostatočno* 'enough'
prebol'no 'very painfully' < *bol'no* 'painfully'
premilo 'very nicely' < *milo* 'nicely'

In addition, PRE- intensifies those adverbs that refer to a degree themselves, as I illustrate in (47):

- (47) *premnogo* 'very much/many' < *mnogo* 'much'
očen'-preočen' lit. 'very much' < *očen'* 'very'
sovsem-presovsem 'to the very extreme' < *sovsem* 'to an extreme'

čutočku-prečutočku ‘very little’ < *čutočku* ‘a little bit’

The intensifying PRE- is also attested in some nouns that refer to a quality or quantity:

- (48) *preneprijatnost* ‘very unpleasant event’ < *neprijatnost* ‘unpleasant event’
prelentjaj ‘very lazy person’ < *lentjaj* ‘lazy person’
preizbytok ‘large abundance’ < *izbytok* ‘surplus’

However, in general, nouns are rather marginal context for this use of PRE-.

Often the use of the prefix PRE-, which expresses a higher degree of a quality, combines with another intensification device – reduplication (e.g. *dobryj* ‘kind’ > *dobryj-dobryj* ‘very kind’).¹¹¹ In combination they compose the pattern X-PRE-X, where the position of X is most frequently filled with an adjective or an adverb: e.g. *dobryj-predobryj* ‘very kind’, *mnogo-premnogo* ‘very much’.

- (49) *Idu obratno. Vižu, devuška – xorošen’kaja-prexorošen’kaja...* [A. Bitov. *Bezdel’nik* (1961-1962)] ‘I am going back. I see a girl – lit. **cute-INTENS-cute.**’

The examination of corpus examples of this reduplicative morphological pattern with PRE- shows that it applies to a variety of semantic types of adjectives. Most frequently, we observe characteristics of age (e.g. *staryj-prestaryj* ‘very old’), personality (e.g. *dobryj-predobryj* ‘very kind’), and color (e.g. *sinij-presinij* ‘very blue’). Other types include characteristics of spatial and temporal length (e.g. *dlinnyj-predlinnyj* ‘very long’, *dolgyj-predolgyj* ‘very long’), size (e.g. *ogromnyj-preogromnyj* ‘very huge’, *malen’kij-premalen’kij* ‘very small’), shape (e.g. *tolstyj-pretolstyj* ‘very fat’), taste (e.g. *vkusnyj-prevkusnyj* ‘very delicious’, *gor’kij-pregor’kij* ‘very bitter’), smell (e.g. *dušistyj-predušistyj* ‘very fragrant’), temperature (e.g. *teplyj-preteplyj* ‘very warm’), as well as physical properties (e.g. *tverdyj-pretverdyj* ‘very solid’) and intellectual properties (e.g. *glupyj-preglupyj* ‘very stupid’). This is not, however, an exhaustive list, more semantic types of adjectives are involved and even more can be employed, because this pattern can potentially be used for any qualitative adjective.

The use of PRE- in such a reduplicative pattern is very productive and extends to words of other classes, for example, *požalujsta-prepožalujsta* ‘INTENS-please’ (50, 51) and *spasibo-prespasibo* ‘INTENS-thank-you’ (52):

- (50) *Timka maxal rukami i kričal: – Požalujsta-prepožalujsta, spasite menja!* [V. Postnikov. *Šapka-nevidimka* (1997)]
‘Timka was waving his arms and shouting: “**Please! Please!**, rescue me!”’

- (51) – *Uvažajmyj kollega, ne rasskažete li vy rebjatkam interesnuju skazku?*
– *Požalujsta-prepožalujsta – kivnul borodač. – Odin dobryj <...> volk s’el odnu devočku.* [Ju. Družkov. *Volšebnaja škola* (1984)]
‘– Dear colleague, could you tell the children an interesting fairy-tale?’

¹¹¹ Reduplication is often iconic: repetition of the formal shape of a word manifests intensification of its semantics and can mark plurality, intensity, or repetition (Bauer 2001: 26). Bauer provides some examples of reduplication in Afrikaans that are typologically similar to Russian, although Afrikaans does not use a prefix: *amper* ‘nearly’ > *amper amper* ‘very nearly’; *dik* ‘thick’ > *dik dik* ‘very thick’ (Bauer 2001: 25).

– “Lit. **here-you-go-INTENS-here-you-go**”, the bearded man nodded. “A kind wolf ate a girl.”

- (52) *Gena! Gena! Vot spasibo! Vot **spasibo-prespasibo!*** [V. Rozov. V poiskax radosti // «Teatr», 1957] ‘Gena! Gena! Thank you! **Thank you very much!**’

Interestingly, the prefix PERE- is also often used in a reduplicative pattern of a similar type X-PERE-X, where the position of X is always filled with a participle. In this use, the prefix PERE- can carry submeanings 4.OVERDO (53), 5.REDO (54), or 13.THOROUGH (55) that intensify semantics of the verbal base:

- (53) *Kakim-to čudom serebrjanoe bljudo [katka «Dinamo»] umestilos’ v gustotiščje **zastroennogo-perezastroennogo** centra Moskvy.* [Ju. Nagibin. T’ma v konce tunnelja (1994)]
‘Somehow, they managed to place the silver plate [the skating-rink “Dinamo”] in the density of **the overbuilt** Moscow downtown.’
- (54) *Kniga byla staraja, **čitanaja-perečitanaja**, s massoj podčerkivanij.* [Ju.O. Dombrovskij. Xranitel’ drevnostej (1964)]
‘The book was old, **read and re-read many times**, with lots of underlinings.’
- (55) ***Stiranoje-perestiranoje** obmundirovanie <...> davno poterjalo blagorodnyj zelenyj cvet.* [A. Azol’skij. Diversant // «Novyj Mir», 2002]
‘The uniform, **washed many times**, lost its noble green color long ago.’

The intensifying PERE- normally never applies to adjectives, although marginally it is attested in the reduplicative pattern typical for PRE-. Consider a corpus example in (56) of the word *perezanjat(yj)* ‘very busy, too busy’ formed from *zanjat(yj)* ‘busy’, an adjectivized participle *zanjat(yj)* ‘occupied’ of the verb *zanjat’* ‘occupy’:

- (56) *Detjam udeljaetsja men’še i men’še vnimanija. Ljudi **zanjaty** i **perezanjaty**.* [P.A. Sorokin. Zametki sociologa // «Volja naroda», 1917]
‘Children get less and less attention. People are **busy** and **over/very-busy**.’ (lit. busy and INTENS-busy)

Finally, PERE- and PRE- can be used in the same reduplicative pattern in the same context, intensifying the meaning of the base each in its own way. Example in (57) illustrates, that PRE- applies to a qualitative scalar adjective and intensifies the property (14.VERY), whereas PERE- applies to a participle and intensifies the activity denoted by its verbal base (13.THOROUGH):

- (57) *On prygal v **staroj-prestaroj, latannoj-perelatannoj** šipovke, v kotoroj pobil četyre mirovyx rekorda.* [A. Srebnickij. Tri goda pjat’ mesjacev vosem’ dnejj... // «Junost’», 1969]
‘He jumped in **very old, patched all over** hobnailed boots, which he was wearing when he beat four world records.’

6.7 Wrapping up the proposal

6.7.1 Non-standard grammatically conditioned allomorphy

The status of PERE- and PRE- in Modern Russian is not trivial. I propose that it can be best captured by extending the traditional notion of allomorphy. I suggest that PERE- and PRE- should be viewed as non-standard allomorphs of a single prefix rather than as different morphemes.

On the one hand, PERE- and PRE- satisfy neither of the two criteria of standard allomorphy: their polysemies are comparable yet not identical. Their distribution across verbs approaches complementarity, but there is evidence of overlap – minimal pairs of verbs (see Appendix 5).

Yet, on the other hand, the two prefixes are historically related and still preserve the formal similarity of their phonological shape. Moreover, the correlation of PERE- and PRE- is supported by a synchronically relevant morphophonological rule and a spelling rule that explicitly associate the two prefixes with each other. In addition, the two prefixes are related in terms of semantics: their semantic contribution in verbs is comparable and can be modelled by means of a single network of polysemy, as I demonstrated in Section 6.3. In the Modern Russian lexicon, PERE- and PRE- share the central spatial prototypical meaning 1.TRANSFER OVER/ACROSS as well as seven other submeanings which are metaphorical extensions of the prototype.

However, my major argument concerning the allomorphic status of PERE- and PRE- is that they share a function of intensification of the base. Both prefixes function as prominent and highly productive morphological intensifiers, each in its own domain – PERE- in verbs and participles, PRE- in adjectives and adverbs. Therefore, I propose that PERE- and PRE- are grammatically conditioned allomorphs of a single morpheme. They have different centers of gravity within shared semantics and different domains of productivity in word classes.

How exactly are PERE- and PRE- conditioned grammatically? Their use is distributed across morphological classes of words, and in terms of productivity this distribution is near-complementary. PERE- is productively used exclusively in the verbal domain, whereas PRE- is productively used beyond it – in adjectives and adverbs. I observe that PRE- is a productive marker of excessiveness in the non-verbal domain, while PERE-, by contrast, is productively employed in verbal morphology.

As a consequence, the two prefixes in question are significantly different in terms of their grammatical properties – aspectual shift, transitivity, and prefix stacking. Overall, PERE- is a more “verbal” prefix than PRE-, it is able to affect such verbal categories as aspect and transitivity. In terms of aspect, PERE- is a productive perfectivizer and usually shifts the aspect of a simplex imperfective to perfective. In terms of transitivity, in certain submeanings (i.e. 1.TRANSFER OVER/ACROSS and 3.SUPERIORITY) PERE- shifts an intransitive base into a transitive derivative. PRE-, on the contrary, is unproductive in verbs and when used, often stacks over another prefix. PRE- is a weak perfectivizer and often does not affect the aspect of the verbal base, so that imperfective verbs often remain imperfective, and perfective verbs remains perfective. Whereas PERE- has a stronger impact on verbal grammatical properties, PRE- is a productive marker of excess in quality, as it is frequently manifested in newly coined adjectives and adverbs.

The Cognitive Linguistic framework and quantitative analysis of corpus data make it possible to account for both the subtle differences and the important similarities

of the two prefixes. This study has an important theoretical implication. It suggests that the traditional definition of allomorphy cannot account for gradient and controversial empirical data and should be reconsidered.

6.7.2 Allomorphy via borrowing

My argument for the allomorphic relationship of the prefixes PERE- and PRE- in Modern Russian leads to a somewhat controversial yet unusual conclusion: I suggest that a non-native affix can be an allomorph of a native cognate affix.

Note that the two prefixes that I examined are etymologically related and phonologically similar. Both factors play a role in the synchronic status of these prefixes. A common etymological source to a large degree motivates the semantic comparability of the native and the loan prefix. Moreover, the phonological similarity of PERE- and PRE- supports their association with each other. The morphophonological alternation *pleophony ~ metathesis*, which relates dozens of roots in the native (Russian) and non-native (Slavonic) lexical strata, also brings the two prefixes into a prominent and perceptible correlation. Yet, the very idea of allomorphy for a pair consisting of a native and a non-native formant requires additional attention.

The unified allomorphic account of such “unequal” affixes as PERE- and PRE- does not come for free. The formal similarity and historical kinship alone do not justify postulating allomorphy, as I illustrated by comparing PERE-/PRE- with PRI- and PRED-. I suggest that it is the synchronic distribution based on significant semantic profiles that reveals the synergy of the two prefixes. My analysis of their most abstract uses in terms of intensification indicates a near-complementary distribution of the two prefixes across the two morphological domains of verbs and adjectives / adverbs. It is crucial that this prominent pattern of distribution captures the most productive uses of PERE- and PRE- in Modern Russian.

In this light, I propose that it is possible for a native affix and a loan affix to become allomorphs in the host language if 1) these affixes are historically related and 2) the host language recycles their former relationship and employs both affixes in a synchronically productive way.

Is this account possible elsewhere than in Russian? Is allomorphy of a native and a loan variant of a single morpheme attested across other languages? Similar cases are sporadically discussed in the literature. I can offer two examples of similar empirical phenomena in English and Spanish. These examples suggest that this type of allomorphy is rather common across languages.

The example from English is described in Spencer 1991. Spencer discusses partially suppletive allomorphs, that is allomorphs which bear phonological similarity but are not related by any phonological rule in Modern English. In particular, Spencer attributes this type of allomorphy to borrowing: “we find that a morpheme (or strictly a word containing the morpheme) has been borrowed from another language at two different times and assimilated in two different forms, or in more complex cases, we find the same morpheme has been borrowed from two different but related languages” (ibid: 8). Spencer illustrates this phenomenon with stem allomorphy of the words *France*, *French*, *Franco*, and *Frank* in Modern English: “the word *France* was borrowed from Norman French (cf. Modern French *France*). But English has also borrowed a morpheme *franco-* (as in *francophile*, *Franco-Prussian*, *francophone*) from Latin (though the Romans themselves got the word from Germanic). Historians use somewhat closer to the original Germanic word when they use the term *Frank* and *Frankish* to refer to early periods of

French history. Finally, the adjective from *France*, namely *French*, is another example of allomorphy, which cannot sensibly be explained by phonological rules of current (synchronic) English. The *France* ~ *French* alternation is the case of a morpheme changing shape through the ravages of historical sound changes” (Spencer 1991: 8).

Fábregas & Scalise (2012: 17) provide a similar example from Spanish. They propose an arguably allomorphic relationship “between a word borrowed from Latin and the result of the historical evolution of that same Latin word”. For example, the words *leche* ‘milk’ and *lácteo* ‘dairy’ used in Modern Spanish are forms of the same root that are historically related to each other, in particular “the first is a historically regular phonological evolution of the second” (Fábregas & Scalise 2012: 17). However, the authors point out that the association of such words with each other depends on speakers’ historical knowledge of this relationship.

Both Spencer 1991 and Fábregas & Scalise 2012 discuss their examples in terms of stem allomorphy. In this regard, my data is different because it presents allomorphy of affixes, but the principle is the same.

6.8 Morphological intensifiers among other types of intensification

A key concept in my analysis of PERE- and PRE- presented in this chapter is intensification. In this section I discuss this phenomenon in detail.

It is common knowledge that linguistic intensifiers refer to increase – an increase of quantity or quality (Dressler & Barbaresi 1994: 416). The effect of intensification can be compared to what a loudspeaker does: it amplifies the meaning of whatever it applies to – be it an entire proposition of a sentence or a meaning of a single word. One can paraphrase an intensifier as ‘very-X’, ‘more-X’, ‘extremely-X’.

Labov (1984: 43) pointed out that intensity lies “at the heart of social and emotional expression”, and “its very nature is not precise”. According to Labov, intensity is gradient and dependent on other linguistic structures. In particular, there are different facets of intensification depending on how it is manifested. Indeed, the very term intensifier often refers to different although related phenomena.

Most studies of intensification focus on lexical intensifiers of propositional semantics, i.e. adverbs like *very*, *absolutely*, *really* (Lorenz 2002: 148) as well the intensive use of quantifiers *all*, *never*, *ever*, *always* (Labov 1984: 48). Others explore the emphatic use of reflexive pronouns like German *selbst*, Russian *sam*, English *myself* (cf. the role of such markers of discourse prominence in contexts like *The President himself will give the opening speech*, cf. König 2001: 749).

Apart from that, intensification can also be marked by expressive phonology, in particular by pitch or segmental length. For instance, in Russian, the prolongation of an initial consonant usually increases a negative evaluation of the adjective, as in *mmmerzkiĭ* ‘filthy’ as opposed to positively-evaluated prolongation of the stressed vowel in *ma-a-alen’kiĭ* ‘small’ (Berkov 1996: 116).

Another widely-studied device of intensification is reduplication of a base, which can be illustrated by Afrikaans: *amper* ‘nearly’ > *amper amper* ‘very nearly’ (Bauer 2001: 25).

Finally, many languages have intensifiers that operate on a more granular level of derivational morphology, or morphological intensifiers. They include affixes that amplify the meaning of a stem and can be paraphrased as ‘very-X’ or ‘more-X’, for example the Italian augmentative suffix *-one* and elative suffix *-issimo* (Dressler & Barbaresi 1994).

In this light, my interpretation of the prefixes PERE- and PRE- as morphological intensifiers is motivated by the analysis of their rich semantics. The contribution of my proposal about their allomorphic status lies in the perspective on these prefixes from the point of view of morphological intensification. I suggest that their distribution and properties are grounded in the key difference: PRE- intensifies a property, whereas PERE- intensifies an activity, or the proposition, denoted by the base predicate. At the same time, both functions (both types of intensification) are semantically developed from a single spatial prototype TRANSFER OVER/ACROSS shared and expressed by both prefixes. In terms of intensification, it becomes possible to provide a unified account of PERE- and PRE- and explain why it is more fruitful to consider them as allomorphs of a single morpheme.

Since intensification is defined as an increase of a quantity or quality, it necessarily refers to a scale and is most naturally applied to adjectives that are typically gradable as a class (Croft 2001). Yet, the use of Slavonic prefix PRE- in Modern Russian demonstrates that morphological intensifiers can be found beyond the adjectival domain. Adverbs are generally less prototypical bases for morphological intensification than adjectives (Dressler & Barbaresi 1994), while nouns can be intensified as long as they have some “adjectival dimension” (van Oss 1989: 77). PRE- applies to both classes and even words of other types. Intensification of verbs generally seems very problematic and hardly possible since they profile temporal relations – processes and events (Langacker 2008: 102)¹¹². Yet Modern Russian provides plenty of examples where prefixation of a verb with PERE- signals intensity of an activity in a number of different ways (3.SUPERIORITY, 4.OVERDO, 5.REDO, 12.DISTRIBUTE, and 13.THOROUGH).

The loan Slavonic prefix PRE- is a more typical intensifier in terms of the morphological domain (adjectives and adverbs) that it applies to. Why is it so that PRE- is specialized for intensification of words that denote atemporal properties, and PERE- is used for intensification of activities and events? Why is this distribution not the other way around? This might be explained by the fact that PRE- is a prefix borrowed from Church Slavonic. Similar to PRE-, other Slavonic loan prefixes like SO- and VOZ- demonstrate weak ability to affect verbal morphological categories, mainly aspect, in Modern Russian. Instead, they are “recycled” and re-employed for different linguistic purposes.

On the other hand, generally, loan formants can easily acquire the intensifying function because they have a privilege over native linguistic affixes, being more “exotic” and therefore more suitable for an expressive use. Compare the prefix *über* ‘over, very’ borrowed from German to Modern English and frequently used as an intensifier. Consider the examples of *über* in (58)-(61) that are culled from the Corpus of Contemporary American English (COCA, <http://corpus.byu.edu/coca/>):

¹¹² Langacker (2008) persuasively argues that the traditional parts of speech, or grammatical classes, such as noun, verb, adjective, etc., are not just empty linguistic labels bound to certain morphological and syntactic behavior but rather refer to different cognitive ways of conceptualizing human experience. He shows that the traditional notions of grammatical word classes are grounded in conceptual archetypes such as ‘thing, or physical object’ for a noun, ‘action, or event’ for a verb, and ‘property’ for an adjective (Langacker 2008: 93-94). The variety of parts of speech employed in human languages witnesses our cognitive ability to construe a situation in alternative ways. Grammatical classes thus differ in terms of their construals. According to Langacker (2008: 102), adjectives profile atemporal relations as opposed to verbs, which profile temporal relations – processes and events.

- (58) *Spider walked in. And I was trying to be like **uber cool**.* [2010, Reality Show, transcript]
- (59) *Two **uber liberal** Congress people, Elijah Cummings and Nancy Pelosi, allow Ms. Fluke access to Capitol Hill.* [2012, Santorum: FOX Favors Romney]
- (60) *Consider the **uber model** of the moment, Gisele Bundchen. Her hair is full of healthy waves...* [2000, Vegetarian Times]
- (61) *Our family was in a bad way. We yelled, threatened... A friend said, "You should call **Ubernanny**"* [2012, The Southern Review]

These examples illustrate the expressive use of *über* employed instead of *super* and *hyper* – semantically comparable intensifiers that arguably have become less expressive in modern English slang.¹¹³

6.9. Conclusions

My contribution in this chapter is twofold – first, to the study of the prefixes PERE- and PRE-, and second, to the study of the phenomenon of allomorphy in general.

First, I presented a first corpus-based contrastive study of the prefixes PERE- and PRE- in Modern Russian. I produced a comprehensive database of Russian verbs formed by these prefixes that can be used in future studies as well. I analyzed this comprehensive dataset from two perspectives – semantic and morphological characteristics. I adopted a set of submeanings most of which were established in earlier studies of PERE- and compared the two prefixes meaning by meaning and word by word. Thus, my findings are based on both the analysis of individual lexemes and a general picture of data distribution. A quantitative account of the list of verbs made it possible to characterize the overall pattern of distribution of PERE- and PRE-. For each submeaning expressed by these prefixes in verbs, I inferred a quantitative measure (index) of their prominency and productivity – a type frequency, that is a number of standard and marginal novel coinages that employ a particular prefix in this submeaning. As a result, I identified submeanings that stand out as the most prominent and best represented across standard prefixed verbs and also the submeanings that are most productive among marginal coinages in the corpus. I proposed a new unified account of productive uses of PERE- and PRE- in terms of the function of intensification. I argued that the two most productive intensifying submeanings of PERE- that are also unattested for PRE- (5.REDO and 14.DISTRIBUTE) indicate quantitative intensification of an activity (proposition) denoted by a base verb. Meanwhile, the two intensifying submeanings that are shared by the prefixes in question (3.SUPERIORITY and 4.OVERDO) represent a qualitative intensification of an activity. In addition, I investigated grammatical properties of the two prefixes and concluded that PERE- is a more “verbal” prefix as it usually alters the aspect of the simplex base, while the loan prefix PRE- is a weak perfectivizer. Instead, PRE- is productive beyond verbs, where it pursues its semantic potential as a marker of a high degree of a property. A statistical analysis revealed a robust correlation of prefix stacking and the ability of a prefix to affect aspect and transitivity of the base.

¹¹³ There are comparable examples of such a use of the German *über* in Norwegian as well: 25-åringen <...> opplevde stor suksess med de **überfengende** singlene «I Spy» og «Jive Babe» – “The 25-year-old had experienced great success with the **übercatchy** songs “I Spy” and “Jive Babe””.

Second, and more generally, in this chapter I revisited an old issue concerning the relationship between native Russian and loan Slavonic morphological elements and defined their modern status in terms of a Non-Standard Allomorphic relation. In this study I demonstrated that it is possible to go beyond a rigid narrow definition of this phenomenon and benefit from examination of non-standard data that does not fit the traditional criteria. As a result, I arrived at the non-trivial conclusion that allomorphic relations can be found between a native local variant and a loan variant of a single morpheme. Although they acquire different roles and become redistributed in the system of the host language, they are compatible alternatives that are associated with each other.

There is a number of arguments in favor of allomorphic status for PERE- and PRE-, yet these arguments have different weight and definitional power. Certainly, it is an important fact that the two prefixes are etymologically related and preserve a formal similarity. However, these two characteristics alone are not sufficient for proposing allomorphy. Moreover, these two factors can be even misleading, as I argued regarding the non-allomorphic relations of PRE- and PRED-.

Apart from a shared diachronic source and synchronically perceptible formal resemblance, PERE- and PRE- are also related by a historically-grounded morphophonological alternation which aligns Russian and Slavonic shapes of a number of other morphemes.

In addition, a crucial role in postulating allomorphic status for the two prefixes belongs to their semantic compatibility in verbs. In this regard, I proposed a new interpretation of several metaphorical submeanings in terms of different facets of semantic intensification. This account revealed that PERE- and PRE- productively perform an identical function of intensification of the morphological base which adapts to the context of a particular morphological domain. As intensifying prefixes, PERE- and PRE- are specialized for two opposite domains – broadly speaking, verbs and adjectives. Apart from this productive pattern of distribution that relates PERE- and PRE- as two alternatives of a single entity, they also share a spatial prototype and the majority of submeanings within the network of polysemy. Thus, in this chapter I showed how to approach the criterion of semantic compatibility in the case when the elements being compared are highly polysemous¹¹⁴. I have shown that it is fruitful to analyze the distribution of prefixes not only in terms of individual verbs but in the light of more general patterns – i.e. morphological types (verbs and adjectives/adverbs). This reveals that the distribution of PERE- and PRE- is in fact grammatically conditioned. An important conclusion that I make in this study is that this distribution is not complementary but contains an overlap in the use of the two prefixes at the peripheries of their productive semantic areas: e.g. PRE- is most productive in adjectives and adverbs but is preserved in the domain of verbs which is peripheral for this prefix. PERE-, by contrast, demonstrates the highest productivity in verbs but can sporadically be used in adjectives. Thus, the distribution is near-complementary when it concerns the domains of their productive application, and the overlap (co-existence) is tolerated in the periphery.

¹¹⁴ PERE- is regarded as one of most polysemous Russian prefixes (Kagan 2013: 490).

Chapter 7

The prefixes VZ- and VOZ- ‘up’

7.1 Introduction

Regarding the prefixes VZ- and VOZ-, Russian presents us with a unique situation in the Slavic domain: the native prefix VZ- and the loan prefix VOZ- have been coexisting since their formal differentiation emerged in the 14th c. In other words, the two prefixes have coexisted for over six centuries in Russian. What are they today? What differences have they gained over time and what properties do they still share? What predominates in their behavior: similarity or divergence? Are they allomorphs of a single morpheme or two distinct morphemes in Modern Russian?

The analysis of these prefixes in terms of allomorphy may appear problematic at least because VOZ- is a loan prefix. However, as I showed in the case study of PERE- and PRE- in Chapter 6, a borrowed prefix can be an allomorph of a native one.

The analysis of VZ- and VOZ- in terms of distinct morphemes would have to prove that the two phonological variants underwent a morphological split and turned into distinct units dissociated from one another.

In scholarly works on Modern Russian the prefixes VZ- and VOZ- are usually described as distinct morphemes, different in their origin (Russian vs. Church Slavic), phonological shape, register, and set of meanings (Baranovskaja 1974: 122; Pilipenko 2001: 68; Tabačenko 2011: 98; Efremova 2010: 533-534; Offord 1996: 201; Matveeva 1983; Golovin 1965). In this chapter,¹¹⁵ I go beyond this traditional view and show that the modern status of these two prefixes is not straightforward.

The main objective is to provide a thorough detailed account for differences and similarities between VZ- and VOZ- and turn the controversy about their morphemic relationship into an empirical question that can be addressed quantitatively. Therefore, I focus on finding out how similar and how distant the two prefixes are and what predominates in their distribution: similarity or divergence.

I discuss the historical relations of the two prefixes in 7.2 and outline the challenging properties of their synchronic behavior in 7.3. In 7.4, I present corpus data collected for this study and explore the distribution of VZ- and VOZ- in standard and marginal verbs. Semantic comparison of the two prefixes is provided in 7.5 and is followed by the quantitative analysis in terms of radial category profiling in 7.6. In 7.7, I examine minimal pairs of verbs prefixed in VZ- and VOZ- and analyze the factors that determine the choice of the prefix. In 7.8, I summarize my findings and relate them to Non-Standard Allomorphy.

¹¹⁵ In this chapter I elaborate the analysis of VZ- and VOZ- carried out by Svetlana Sokolova and myself based on 97 perfective verbs. We analyzed all Natural Perfectives and those Specialized Perfectives that have more than 100 attestations in the corpus. The results of that joint project are described in Endresen et al. 2012: 264-266 and Janda et. al 2013: 56-58. The earlier versions of this analysis were presented at the November seminar at the University of Tromsø (Baydimirova & Sokolova 2010) and at the 44th Annual Meeting of the Societas Linguistica Europaea at the University of La Rioja, Logroño in Spain.

7.2 Etymological relationship and formal similarity

In Modern Russian the prefixes VZ- and VOZ- coexist but belong to different lexical strata: the prefix VZ- is a native Russian prefix, while the prefix VOZ- is a borrowing from Church Slavic (Townsend 1968: 59; Gallant 1979: 64; Vinogradova 1984: 24; Tabačenko 2011: 97).

Etymologically both prefixes VZ- and VOZ- come from a single Proto-Slavic adverbial *vъz*. The latter can be traced back to the Indo-European root **ups*: **up*, which gave rise to similar adverbial particles, prefixes and prepositions in other languages, for example Lithuanian *už* and English *up* (Vasmer 1971: v.1, 333). The initial consonant in *vъz*- is a prothetic *v* and is considered a Slavic innovation (Tabačenko 2011: 97).

Both Slavic prefixes VZ- and VOZ- result from two reflexes of a positional phonological alternation in the prefix VЪZ-. The prefix VZ- is explained by the weak position of the yer vowel *ъ*, where the vowel finally disappeared, whereas VOZ- is the reflex of the strong position of the same vowel *ъ*, where it developed into a full vowel *o* (recall the discussion of yer vowels in Chapter 3). Those cases where the yer vowel should have disappeared due to its weak phonological position but was preserved are qualified as borrowings from Church Slavonic (Tabačenko 2011: 97). However, we know that in the 11-14th centuries in Old Russian there was still a single prefix VЪZ-, and there was no semantic split between its phonological allomorphs (Baranovskaja 1974: 123; Tabačenko 2011: 97).

Filin attributes the distinction between VZ- and VOZ- and the vocalization of the vowel *o* to the role of artificial church pronunciation that emerged after the fall of yers¹¹⁶ (Filin 1981: 51). Thomas (1969: xx¹¹⁷) dates this process between the Kievan and Muscovite periods of Russian history. He suggests that this was the time when new Church Slavonic doublets were imported into Russian, and in particular the artificial church pronunciation of the vowel in the prefix VЪZ- “in places where spoken Russian now had no vowel” led to creation of new slavonicisms. This process brought about doublets like the slavonicism *vozrast* ‘age’ as opposed to Russian *vzroslyj* ‘grown-up, adult’. According to Uspenskij (2002: 289), the prefix VOZ- becomes productive at the time of the Second South Slavic Influence. In other words, the case of VZ- and VOZ- cannot be considered traditional borrowing, where an element of one language gets “transferred” to another language (Bussmann 1996: 287).

Partly because the two prefixes share their historical origin, they also maintain formal phonological similarity in Modern Russian. This pair of prefixes is not related by the same historical alternation as PERE- and PRE- discussed in Chapter 6, but the opposition of the shapes VZ- and VOZ- also involves the absence vs. presence of a vowel that was originally a yer.

Summing up, VZ- and VOZ- stem from a single etymological source and are former phonological variants, which once were related via a vowel/zero alternation. The process that brought the two forms into existence is no longer active in Modern Russian, although vowel/zero alternation is attested as a morphophonological phenomenon that applies to other morphemes in Modern Russian (for example, in RAZ- ~ RAZO-). The two prefixes are phonologically similar and can be easily associated with each other, because they often attach to the same roots, as in *vozrast* ‘age’ and *vzroslyj* ‘adult’ (< the root *rast* ~ *ros* ‘grow’).

¹¹⁶ The original text is “Ego vokalizacija odnositsja k iskusstvennomu cerkovnomu proiznošeniju, kotoroe vzniklo posle padenija reducirovannyx.” (Filin 1981: 51).

¹¹⁷ This is a page number.

Both characteristics, common etymological source and formal similarity, facilitate a close relationship between the two prefixes in Modern Russian. I address their synchronic behavior in the next section.

7.3 Distributional overlap and subtle semantic difference

The prefix VZ- has allomorphs VZ- (*vzletet'* 'fly up'), VS- (*vskipet'* 'boil'), and VZO- (*vzobrat'sja* 'climb'). As shown in Table 1, VZ- is a default major allomorph, VS- is the result of regressive devoicing in the context of the adjacent voiceless consonant, while VZO- is a vocalized allomorph, which occurs in front of consonant clusters. In other words, these variants are phonologically conditioned and represent Standard Allomorphy similar to that of RAZ-, RAS-, and RAZO-, discussed in Chapter 3.

Allomorph type:	VZ-			VOZ-		
	default allomorph	VZ-	<i>vzletet'</i> 'fly up'	<i>letet'</i> 'fly'	VOZ-	<i>voznesti</i> 'raise'
devoiced allomorph	VS-	<i>vskipet'</i> 'boil'	<i>kipet'</i> 'boil'	VOS-	<i>vospet'</i> 'praise'	<i>pet'</i> 'sing'
vocalized allomorph	VZO-	<i>vzobrat'sja</i> 'climb'	<i>brat'(sja)</i> 'take'	VOZO-	<i>vozmnit'</i> 'get a false idea of one's importance'	<i>mnit'</i> 'think much of oneself'

Table 1: Phonologically conditioned allomorphs of the prefixes VZ- and VOZ-.

In parallel, the prefix VOZ- also has three phonologically conditioned allomorphs: the major allomorph VOZ- (*voznesti* 'raise'), the devoiced allomorph VOS- in front of voiceless consonants (*vospet'* 'praise, eulogize'), and the vocalized allomorph VOZO- in front of certain stem-initial consonant clusters (*vozmnit'* 'get a false idea of one's importance').

This comparison shows that VZ- and VOZ- have similar sets of allomorphs, and the relationship between their allomorphs is rather straightforward. However, the relationship between VZ- and VOZ- is more complex.

On the one hand, these two prefixes can occur within a single paradigm of the same verb, for example, the prefix VOZ- appears in the perfective verb *vozzvat'* 'appeal to', which corresponds to the imperfective verb *vzyvat'* 'appeal to' prefixed in VZ- (cf. Ožegov & Švedova 2001, Efremova 2000). The alternation of VZ- and VOZ- within a single aspectual pair suggests that they must be variants of a single morpheme and thus exhibit an allomorphic relationship.

Yet a short glance at another aspectual pair suggests that this observation may be an illusion, and indicates that the picture is not that simple. The perfective verb *vzoiťi* 'rise' is associated with two imperfectives: *vosxodit'* 'climb, rise, date back' prefixed in VOZ- and *vsxodit'* 'climb, rise, spring up' prefixed in VZ-. These imperfective verbs have similar meaning, but are not always interchangeable. In particular, the verb *vosxodit'* can denote tracing something back to earlier distant events in the past, but such use is not possible for the verb *vsxodit'*. On the other hand, the verb *vsxodit'* can describe the growth of young sprouts above ground, whereas *vosxodit'* does not allow such use. Thus, the meaning of these verbs prefixed in VZ- and VOZ- is not identical, which violates the semantic criterion of Standard Allomorphy.

Moreover, the fact that both prefixes can be attached to the same simplex *xodit'* 'walk' suggests that they are not distributed complementarily across verbs, thus violating the second, distributional, criterion of Standard Allomorphy. There are more examples of this distributional overlap, where both prefixes can be attached to the same base verb. For example, *stat'* 'become, stand' forms derivatives *vstat'* 'stand up' and *vosstat'* 'rise, rebel' (cf. Endresen & Plungian 2011), the simplex *vesti* 'lead' is a base for the prefixed verbs *vzvesti* 'take up, cock a gun' and *vozvesti* 'elevate (e.g. to the throne); erect a building'. Similarly, the verb *ljubit'* 'love' can combine with both VZ- and VOZ-, as shown in derivatives (*ne*)*vzljubit'* 'start disliking' and *vozljubit'* 'come to love'.

Another question is: are there contexts where the two prefixes show free variation and are completely interchangeable? Returning to our example with two imperfectives *vsxodit'* and *vosxodit'*, we should notice that they overlap in the spatial meaning 'climb, rise'. Both verbs occur in contexts that describe concrete spatial unidirectional upward movement designated via the prepositional phrases *na goru* 'up onto the mountain', *na veršinu* 'uphill to the mountain top', *na tribunu* 'up to the platform', *na kryl'co* 'onto the porch'. Examples in (1) and (2) provide a context where the two alternatives that feature prefixes VZ- and VOZ- are interchangeable:

- (1) *Oni po očeredi vosxodili na tribunu.* [Je. Proškin. *Mexanika večnosti* (2001)]
'One after another they **were walking up** the tribune.'
- (2) *Vsxodja na tribunu, on vdruk načinal govorit' ne vpolne svojim golosom.*
[K. Vanšenkin. *Pisatel'skij klub* (1998)]
'When **walking up** the tribune he would suddenly start talking with not quite his normal voice'.

However, even though (1) and (2) on the face describe the same situation of people climbing the platform, one could argue that in (1) the upward motion probably went slower and took longer time than the motion described in (2). Another possibility is that the verb in (1) depicts a longer trajectory of the upward movement as opposed to the trajectory in (2), which zooms in on the final part of the trajectory.

This slight difference of the same verb prefixed in VZ- and VOZ- might seem a pure speculation. Obviously, different speakers might have different intuitions about it. Yet it is intuitively true that the verb of motion *xodit'* 'walk' prefixed in VOZ- is more likely to describe upward movement which aims to reach a high and remote point like *gora* 'mountain' or *veršina* 'mountain top'. By contrast, the counterpart derivative *vsxodit'* formed by the prefix VZ- is much more natural in situations with shorter trajectories and closer destinations like *kryl'co* 'porch', *kapitanskij mostik* 'captain's bridge', *xolm* 'hill', and *veranda* 'terrace'.

I suggest that this collocational tendency is a symptom of the fact that the two prefixes are different not only in terms of register as usually stated in the literature (Vinogradova 1972: 178; 1984: 24-26). In this chapter I show that the difference in register is only one effect of the core difference that VZ- and VOZ- have in terms of their spatial image schemas. In particular, while these prefixes share the image schema of upward movement, they differ in terms of the length of the vertical trajectory: the prefix VOZ- encodes a high vertical path, whereas VZ- categorizes a short path that does not extend far from the surface. I further suggest that this difference in spatial image schemas motivates metaphorical uses of VZ- and VOZ- in verbs of other semantic classes.

We can speculate that the difference in image schema between VZ- and VOZ- is the result of the process described by Pilipenko (2001: 68) under the term *zaraženie kontekstom*, literally ‘infecting by the context’. The variant VOZ- acquired this connotation from its frequent use in bookish Church Slavonic texts of high register. As argued by Pilipenko (2001: 68), the contexts that were highly frequent for verbs in VOZ- gradually became the only possible contexts for these verbs. As a result, the semantics of such idiomatized contexts “infected” the verbs themselves and became part of the semantics of verbs used in these contexts.

Whereas the difference between VZ- and VOZ- in terms of spatial image schemas points towards their distinct morphological status, there is another property of their behavior that suggests the opposite analysis in terms of allomorphy. In poetry, the shape of the prefix can be manipulated for metrical purposes, and VOZ- can be shortened into VZ-, as in (3), whereas VZ- can be replaced with VOZ-, as in (4).

- (3) *Idut vojska, razduvaja stjag, Idut vojska, što brali reixstg, Vojska, što v mukax bezmernyx **Vzrosli** do legendy zemnogo šara, I vižu ja sredi nix bessmertnyx Zoju, Matrosova i Laara.* [I.L. Sel’vinskij. “Dal’še Otečestvennaja pošla...” (1967)]
 ‘The troops walk and unfurl the banner, The troops that conquered Reichstag, The troops who **grew** and became the legend of the globe, And I see among them the immortal Zoja, Matrosov, and Laar.’
- (4) *Odin iz Majakovskix napisal bukval’no sledujuščee: “Prekrasnyj kon’ zdes’ ne valjalsja, On zdes’ **vosпрыgnul** na bugor, I navsegda na nem ostalsja, Pobednyj vitjaz’ Svjatogor.”* [V. Slipenčuk. Zinziver (2001)]
 ‘One of young poets wrote exactly the following: “The beautiful horse did not lie here, It **jumped up** onto the hillock And stayed there for ever, The victorious warrior Svjatogor”.’

In standard Russian, the verbs *vzrasti* ‘grow’ and *vosпрыgnut’* ‘jump up’ normally contain the opposite prefixes: *vozrasti* ‘grow’ and *vsпрыgnut’* ‘jump up’. Moreover, in the context of the global scale of the event described in (3), we expect the use of VOZ-, whereas the context of a short scale of a size of a hillock in (4) triggers the prefix VZ-. Thus, the actual use of these prefixes in (3) and (4) to some extent contradicts hearers’ expectations about the use of the prefix and lends an additional effect of creative speech. Note that in these examples the manipulation of the prefix shape for the sake of rhythm does not cause misunderstanding or confusion about the meaning of the derivative verbs. Yet, in (4), the use of the prefix VOZ- causes a comical effect, which is arguably part of the poet’s strategy in creating his flamboyant writing style.

This manipulation of the prefix shape for metrical purposes suggests a close relationship between VZ- and VOZ- and association links that exist between these prefixes in Modern Russian.

7.4 Data

I collected and analyzed a data set of 384 verbal lexemes, including 241 verbs prefixed in VZ-, 141 verbs in VOZ-, and 2 verbs with alternating VZ- ~ VOZ- in the paradigm. The database is formatted as a Microsoft Excel spreadsheet and is available at <http://hdl.handle.net/10037.1/10078>.

7.4.1 Methodology

In order to take into account as much available data as possible, the first step was to collect all verbs that start with *voz*, *vos*, *vz*, and *vs*, where these sequences are not necessarily prefixes. The verbs were extracted via the software management program MySQL¹¹⁸ from the frequency dictionary by Lyashevskaya & Šaroff (2009), based on the Modern Subcorpus of the RNC (texts created in 1950-2007).

The next step involved manually checking each verb in the corpus and dictionaries. As a result, a number of verbs were excluded from the list: 1) misspelled verbs (e.g. *vspilit'* a typo for *vspylit'* 'suddenly become angry'); 2) unprefixated verbs (e.g. *vozt'* 'transport'), 3) verbs prefixed in V- 'into' (e.g. *vstavit'* 'insert' < *stavit'* 'set').

The resulting set of datapoints was organized in the following way. Each verb included in the database had to be formed via the prefixation of a simplex base with VZ- or VOZ-. Most verbs in the database are perfectives formed by the attachment of the prefix to the unprefixated simplex imperfective like *vospet'*_{PF} 'praise' (< *pet'*_{IPF} 'sing'). The database includes also perfective verbs formed from perfective verbal bases (*vskriknut'*_{PF} 'shout loud once' < *kriknut'*_{PF} 'shout once') and from non-verbal bases (*vozglavit'*_{PF} 'take over the leadership' < noun *glava* 'head'; *vsxolmit'* 'make hilly' < noun *xolm* 'hill'). In addition, the database contains perfective verbs with multiple motivations like the verb *vzdorožat'* 'rise in price' associated with both the imperfective *dorožat'* 'become more expensive' and the adjective *dorogoj* 'expensive'.

Perfective verbs are formed via prefixation and allow for semantic comparison of the prefixed derivative with a simplex base. By contrast, most secondary imperfectives like *vospevat'* 'praise' are formed via suffixation of a prefixed perfective (e.g. *vospevat'* 'praise' < *vospet'* 'praise' + *-a-*) and therefore are not relevant for examination of prefix semantics. For this reason, all secondary imperfectives attested in the corpus that have a corresponding perfective verb are listed in the database within the entries of their corresponding perfectives. For instance, the perfective verb *vospet'* 'praise' and its secondary imperfective *vospevat'* 'praise' are coupled together into one entry. If more than one secondary imperfective is attested for the same perfective verb, all imperfectives are listed in one entry with their perfective: e.g. the perfective *vzryxlit'* 'make friable (of soil)', and its secondary imperfectives *vzryxljat'* and *vzryxlivat'*. Thus, imperfective verbs are preserved in the database for informative purposes but are not counted as separate entries if they have a corresponding perfective verb.

In case a corresponding perfective is not attested and the imperfective is formed by a morphological construction combining the prefix VZ-/VOZ- and the imperfective suffix, it is listed as a separate entry. There are 20 such imperfectives. They include four verbs that belong to the standard Russian lexicon and have more than ten attestations in the corpus (*vossylat'* 'send', *vozdyxat'* 'slobber over', *vozdymat'* 'raise high up', and *vzdymat'* 'raise') and 16 marginal verbs with less than 10 attestations (e.g. *vskripyvat'* 'creak repeatedly', *vzvjakivat'* 'produce a short clinking sound repeatedly').

Note that in the latter verbs we are dealing with fusion on the morpheme boundary: the prefix VZ- becomes partially fused with the initial consonant of the bases *skripet'* 'creak' and *zvjakat'* 'clink'. As a result, the consonant clusters *vzskr* and *vzzv*

¹¹⁸ The command used in MySQL software was `SELECT * FROM verbforms.freqlemma f1 WHERE lemma LIKE "B3%"`, where `verbforms.freqlemma` is the electronic database of verbs available in Lyashevskaya & Šaroff (2009). Because Russian orthography reflects the voicing assimilation on VZ- and VOZ-, the search included all possible graphic variants: *vz-*, *vs-*, *voz-*, *vos-* (they replaced the letters B3 in the above formula).

become simplified into *vskr* and *vzv*. The same type of fusion can be observed in the verb *vstat* 'stand up' formed from the simplex *stat* 'stand, become', as discussed in detail in Endresen & Plungian 2011.

In order to avoid duplication of data, I merged reflexive verbs with the postfix *-sja* with their non-reflexive counterparts, if the postfix plays only an intransitivizing role and the semantic contribution of the prefix in a reflexive and non-reflexive is the same. For example, the reflexive verb *vstoporščit'sja* 'bristle itself' is merged together with its non-reflexive counterpart *vstoporščit* 'bristle' into a single verbal lexeme *vstoporščit'(sja)* 'bristle', where the parentheses show that the reflexive verb in *-sja* is attested. If a non-reflexive perfective form is not attested in the corpus, only the reflexive perfective is listed in a separate entry and the postfix *-sja* is not put in parentheses: e.g. *vzbuševat'sja* 'storm, rage' (no **vzbuševat*), *vzgorjačit'sja* 'get oneself enflamed rapidly and unexpectedly' (no **vzgorjačit*).

Because reflexives are merged with their non-reflexives and secondary imperfectives are merged with their perfectives, a single entry in the database often represents not one verb like *vospitat* 'bring up' but four verbs, namely the perfective *vospitat* itself, its secondary imperfective *vospityvat* 'bring up' and corresponding reflexive verbs *vospitat'sja* and *vospityvat'sja*, or even more verbs, depending on how many secondary imperfectives and corresponding reflexive forms are attested. For the purposes of quantitative analysis, I assume that all four (or more) such verbs represent a single verbal lexeme that occupies one entry in the database.

Four verbal lexemes were added to the data collected from the corpus: *vzborozdit* 'furrow', *vz'eršit'(sja)* 'ruffle', *vskosmatit'(sja)* 'make shaggy', and *vstopyrit'(sja)* 'bristle'. Even though these verbs are not attested in the corpus, they are well-recognized Natural Perfectives present in major dictionaries and in the Exploring Emptiness database of Russian prefixes created at the University of Tromsø¹¹⁹:

7.4.2 Overview: VZ- and VOZ- in standard and marginal verbs

The database shows that the collected verbs are very diverse in terms of their token frequencies.

On the one hand, there are highly frequent verbs in VZ- and VOZ- that have thousands of corpus attestations: for example, *vzjat'(sja)* 'take', *vspomnit'(sja)* 'recall', *vstat* 'get up', *vozniknut* 'emerge', *voskliknut* 'exclaim', and *vosstanovit'(sja)* 'restore'.

On the other hand, there are numerous rare verbs like *vzryčat* 'start roaring' (4 atts.), *vozrevnovat* 'start being jealous' (9 atts.), *vospreobladat* 'win, dominate' (3 atts.) and even single occurrences like *vzdeševet* 'suddenly become cheaper' (1 att.) and *vozžalet* 'start feeling pity about' (1 att.).

Among the verbs rarely attested in the corpus there are newly created occasionalisms like *vzbyčit'sja* 'suddenly become stubborn and unfriendly' that are unlikely to be known to most speakers and obsolete verbs like *vozalkat* 'start wanting' that do not belong to actively used Modern Russian vocabulary.

The Russian National Corpus is a well-balanced collection of original Russian texts that is representative for Russian linguistic production. Therefore, I assume that the token frequencies of the prefixed verbs correspond to the degree of entrenchment of the words in language use. For this reason, rare verbs that lie outside the standard

¹¹⁹ <http://emptyprefixes.uit.no/>

Russian lexicon are symptomatic in terms of the productivity of certain prefixes but are not informative from the perspective of linguistic input shared by all speakers.

In the database, I distinguish between standard and marginal verbs. In order to distinguish between these two classes of verbs, I make use of their token frequencies. It appears that verbs that have ten or more attestations are well-known and standard, while verbs that have nine or fewer attestations in the Modern Subcorpus of the RNC are marginal. This matters especially because the database includes many verbs with few attestations. Such verbs comprise 173 lexemes and thus account for 45% of the entire database. If we fail to distinguish between standard and marginal verbs, one group can skew the distribution and the overall conclusions could be distorted.

The distribution of the two prefixes is different across standard and marginal verbs. Table 2 provides numbers on these two groups subdivided into smaller sub-types according to different token frequencies of verbs: frequent (>9), rare (9-2), single (1) and unattested (0).

Type of verbs	Token frequency type	VZ-	VOZ-	VOZ- /VZ-	Total:
Standard	Frequent (10 and more atts.)	118	91	2	211
Marginal	Rare (9-2 attestations)	61	28	0	173
	Single attestation	58	22	0	
	Unattested in the RNC	4	0	0	
Total:		241	141	2	384

Table 2: Distribution of verbs in VZ- and VOZ- across frequency types in the RNC.

Figures 1 and 2 visualize the distributions of prefixes among standard and marginal verbs respectively.

As shown in Figure 1, in standard verbs (with 10 and more attestations in the corpus), VZ- appears in 56% of derivatives, and VOZ- is used in 43% of verbs. There are two verbs (1% of standard verbs) that exhibit an alternation of the two prefixes within an aspectual paradigm: *vzojti*_{IPF} – *vsxodit'*_{IPF}/*vosxodit'*_{IPF} ‘go up, rise’ and *vozzvat'*_{PF} – *vzyvat'*_{IPF}/*vozyvat'*_{IPF} ‘appeal to’.

Figure 2 shows that the proportion of VZ- and VOZ- is significantly different¹²⁰ in marginal verbs, where VOZ- is attested only in 29% of cases (50 lexemes), while a much greater majority of lexemes is prefixed in VZ-, namely 71% (123 lexemes).

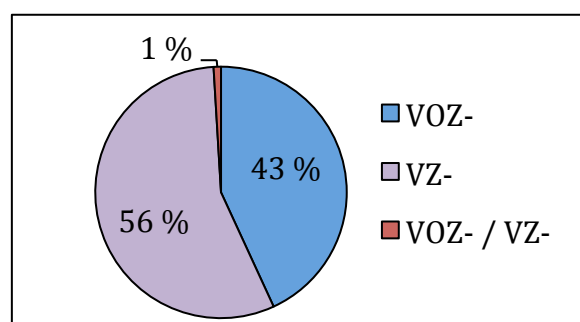


Figure 1: Distribution of prefixes VZ- and VOZ- in standard verbs.

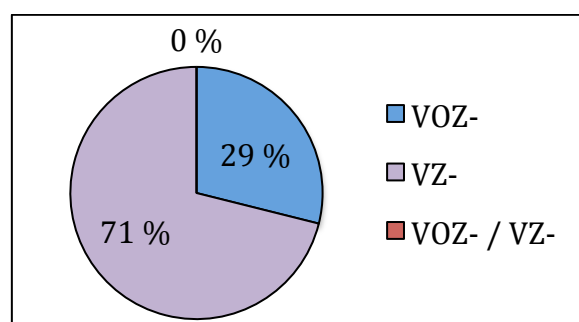


Figure 2: Distribution of prefixes VZ- and VOZ- in marginal verbs.

¹²⁰ The difference between the distributions of the two prefixes was subjected to a chi-squared test and found statistically significant (X-squared = 8, df = 1, p-value = 0.004), with a small, but reportable effect size (Cramer's V = 0.15).

What does this tell us? First, the prefix VOZ- is almost equally well represented in the standard Russian lexicon as the prefix VZ- (compare 91 vs. 118 lexemes respectively). Second, both prefixes are used to coin novel verbs in Modern Russian: consider new coinages like *vsxrjuknut'* 'give a grunt', *vznervit'* 'make someone nervous', *vzdeševet'* 'suddenly become cheaper' in VZ- and the verbs *vozbuxnut'* 'oppose, rebel' and *vospljasat'* 'dance' in VOZ-. This fact disproves the traditional assumption that the prefix VOZ- is completely unproductive (Pilipenko 2001: 71). However, in production of novel words VZ- overrules VOZ-. Moreover, most low-frequency marginal verbs in VOZ- are obsolete like *voskrylit'* 'fly up' (< *krylo* 'wing'), *vosklubit'sja* 'swirl up' (< *klubit'sja* 'swirl'), and *vozglagolit'* 'start speaking' (< *glagolit'* 'speak').

It is even more informative to look at the productivity of the two prefixes with regard to their diverse semantic uses that I identify in the next section.

7.5 Semantics of the prefixes VZ- and VOZ-

I propose that the rich polysemies of both prefixes can be modelled as a single radial network of meanings¹²¹ presented in Figure 3.

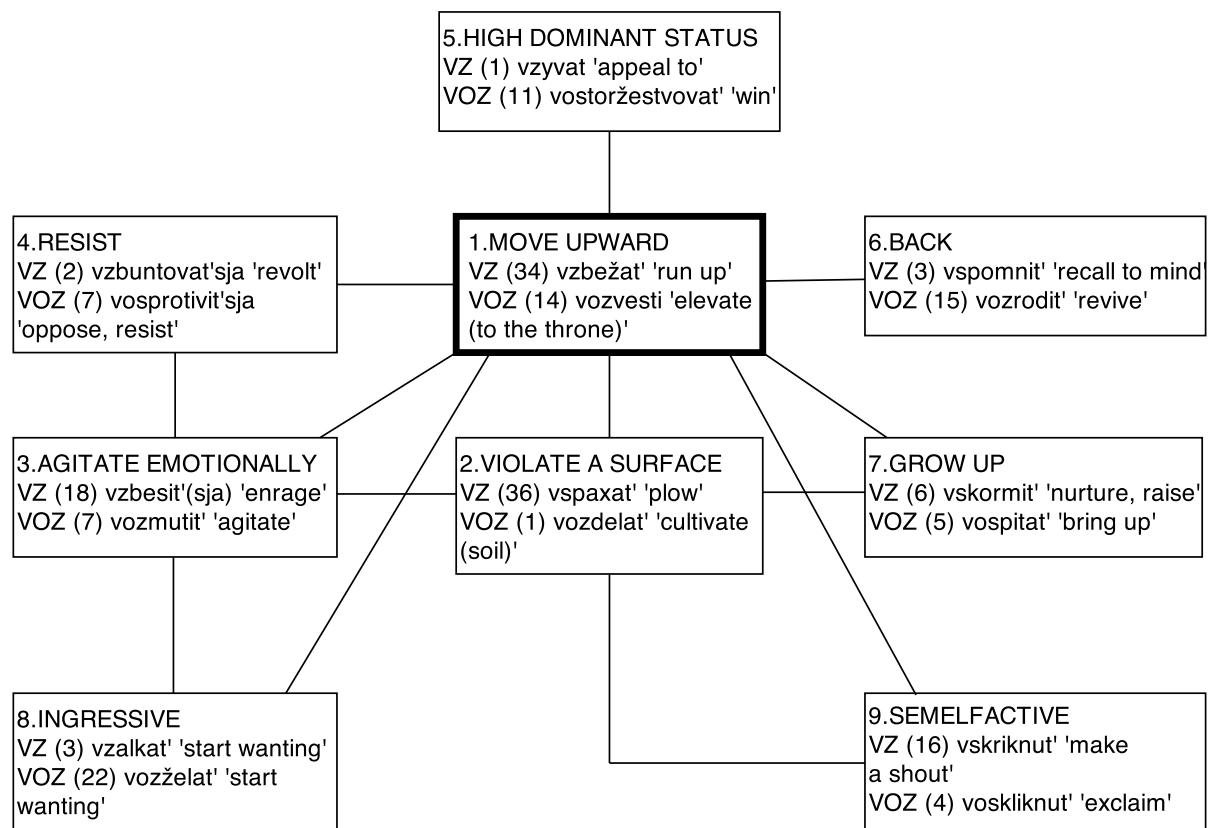


Figure 3: Radial category model for the prefixes VZ- and VOZ-.

¹²¹ The inventory of submeanings proposed in the present analysis is compatible with previous scholarship. The Old Russian prefix VЪZ- is described in dissertations by Pilipenko (2001) and Tabačenko (2011). A contrastive analysis of the Modern Russian VZ- and VOZ- and the Old Russian VЪZ- is offered in Baranovskaja 1974. Będkowska-Kopczyk (2012) examines the related prefix VZ- in Slovene verbs denoting emotions. Šarić (2012) offers a cognitive account of the cognate Croatian prefix UZ-.

In Figure 3, rectangles represent submeanings of prefixes, and the lines between rectangles show semantic links that connect submeanings via cognitive mechanisms. All nine submeanings are attested for both prefixes, although some submeanings are more typical for VZ-, whereas others are more closely associated with VOZ-. For example, 2.VIOLATE A SURFACE is almost exclusively expressed by VZ-, and 5.HIGH DOMINANT STATUS is mostly manifested by VOZ-. These and other crucial differences are highlighted and explained in the present analysis. Numbers in parentheses correspond to standard verbs that represent each submeaning of each prefix. Overall, this model comprises 211 standard lexemes in VZ- and VOZ-. Marginal verbs are also taken into account and cited in the text.

Before I describe each submeaning in detail, I outline the overall structure of this semantic network. The nine submeanings presented in Figure 3 differ in their nature and status in the overall network of polysemy.

The center of the network is the spatial prototype 1.MOVE UPWARD which is also historically primary for both prefixes (Baranovskaja 1974: 124). This concrete spatial meaning motivates one spatial meaning 2.VIOLATE A SURFACE, five metaphorical submeanings 3.AGITATE EMOTIONALLY, 4.RESIST, 5.HIGH DOMINANT STATUS, 6.BACK, 7.GROW UP, and two grammatical aspectual meanings 8.INGRESSIVE and 9.SEMELFACTIVE.

The spatial semantic extension 2.VIOLATE A SURFACE is related to the prototype 1.MOVE UPWARD in terms of metonymy: instead of encoding movement along a vertical trajectory, it focuses on the uppermost parts of objects, the surfaces (e.g. skin in *vspuxnut* 'swell up' < *puxnut* 'swell').

Submeanings 3-7 are metaphorical extensions of the prototype. They result from the cognitive mechanism of mapping: the vertical scale is mapped from the spatial domain (1.MOVE UPWARD) to the domain of human emotional and mental states (3.AGITATE EMOTIONALLY), the domain of human behavior (4.RESIST), social hierarchy (5.HIGH DOMINANT STATUS), temporal precedence (6.BACK), and growth with age (7.GROW UP). Note that these submeanings are not specific for the Russian marker of upward motion but are based on common orientational metaphors CONSCIOUS IS UP (e.g. *wake up*), HAVING CONTROL IS UP (e.g. *control over someone, being on top of the situation*), MORE IS UP (*the income rose*), and HIGH STATUS IS UP (*at the peak of career*) identified by Lakoff & Johnson (1980: 15).

The two submeanings located at the bottom of Figure 3 are most abstract and distant from the spatial prototype: the aspectual meanings 8.INGRESSIVE and 9.SEMELFACTIVE. Their conceptual relationship to the spatial prototype can be explained in terms of semantic bleaching of the initial spatial meaning which typically takes place in the process of grammaticalization. In what follows, I provide a detailed account of which aspectual submeanings are more closely associated with each of the two prefixes in question.

7.5.1 Shared prototype

Submeaning 1.MOVE UPWARD is central for both prefixes. It refers to a basic and concrete spatial experience – a physical upward motion of an object (Trajector) along a vertical path (Trajectory), as shown in Figure 4. This meaning motivates all other

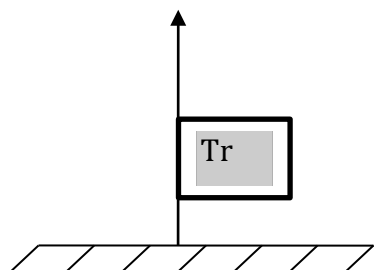


Figure 4: Image schema of the spatial prototype 1.MOVE UPWARD.

submeanings of VZ- and VOZ-, as recognized in Pilipenko 2001, Tabačenko 2011, Będkowska-Kopczyk 2012, Baranovskaja 1974, and Gallant 1979.¹²²

MOVE UPWARD is well represented in verbs of motion and verbs of position. The prefix VZ- more frequently forms verbs of motion like *vzletet'* 'fly up' (< *letet'* 'fly') and *vzoiiti* 'rise, go up' (< *iditi* 'walk'), whereas VOZ- attaches almost exclusively to verbs of position: *vozložit'(sja)* 'place on the top of' (< *klast'/ložit'sja* 'lay') and *vossedat'* (more common *vossedat'*_{IPF}) 'sit high on something (usually on a throne)' (< *sest'* 'sit').

However, VZ- is also attested in verbs of position (*vstat'* 'get up, stand up' < *stat'* 'stand, become'), and VOZ- can also be found in verbs of motion: *vozvesti* 'elevate (to the throne); raise, erect a building' from *vesti* 'lead' and *voznesti(s')* 'bring up, raise' from *nesti(s')* 'carry'.

An interesting case is the verb *vzvalit'* 'throw up, lift upon', where the semantics of both the prefix and the base are associated with the vertical scale but imply opposite directions: the simplex base *valit'* denotes 'throwing down', while the prefix suggests upward movement. In the outcome derivative the meaning of the prefix prevails over the meaning of the base: *vzvalit'* 'lift up'. Likewise, the verb *vozniknut'* 'emerge, uprising' is built out of the prefix VOZ- which encodes the vertical path and the base *niknut'* 'wilt, lend down' which denotes motion in the opposite direction. Similarly to *vzvalit'*, in *vozniknut'* the schema of the prefix outcompetes the schema of the base.

7.5.2 Different altitude

Both VZ- and VOZ- encode a path along a vertical trajectory where the Trajector MOVES UPWARD. Following Baydimirova & Sokolova 2010, I propose that it is fruitful to differentiate between VZ- and VOZ- in terms of the length of the path encoded in their image schemas.

As shown in Figure 5, the prefix VZ-, as opposed to VOZ-, implies a shorter upward trajectory, and therefore most often focuses on the top part of the Landmark, namely the surface. By contrast, the image schema of the prefix VOZ- is characterized by a longer trajectory, which takes more time to complete. Compare, for example, the contrastive pair of the verbs *voz-vesti* 'bring to an elevated position, erect a building' vs. *vz-vesti* 'raise a hammer on a gun', or the two imperfective verbs *vsxodit'* 'rise' vs. *vosxodit'* 'go up' discussed above.

¹²² Diachronic studies show that the meaning MOVE UPWARD is central not only for Modern Russian but also for earlier historical stages. Baranovskaja (1974: 124) and Pilipenko (2001) suggest that this meaning is etymologically primary and motivates other submeanings to a more or less perceptible degree. In the early Slavic translations from Greek, the Old Church Slavonic prefix V'bz- corresponds to the Greek *áva-* (*ana*) 'upwards, increase, repetition, backwards' (Mahota 1993: 292). The meaning 'motion up' of V'bz- was of Common Slavic origin and today is attested for descendants of this prefix in most of modern Slavic languages: Bulgarian V'bz-, VZ-, Czech VZ-, VZE-, Slovak VZO-, Polish WZ- (Vasmer 1971, v.1: 310), Slovene VZ- (Będkowska-Kopczyk 2012), Croatian UZ- (Šarić 2012), and Ukrainian Z-/ZI- (Basova et al 2003: 133).

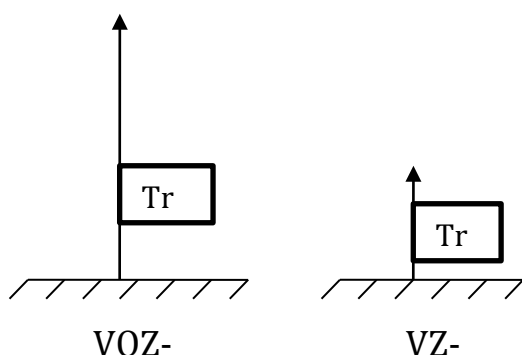


Figure 5: Spatial image schemas of the prefixes VOZ- and VZ-: different altitude.

The difference in the altitude of the trajectories is also evident from simplex bases that VZ- and VOZ- combine with. For example, the prefix VOZ- forms a verb *vozvysit'* 'raise, elevate, go up' from a base that refers to height. This derivative is multiply motivated by several possible bases: the verb *vysit'sja* 'tower, be high', the adjective *vysokij* 'high', and the noun *vys'* 'height'. Either way, it is the prefix VOZ- that is compatible with the semantics of high altitude. Example in (5) emphasizes the semantics of height by lexical means:

- (5) Nu a v maksimal'no vysokom, vtorom položenii kuzov **vozvysilsja** nad zemlej na 275 mm. [A. Konov. Rezerv olimpijca. (2004)]
 'And in the highest possible (lit. maximally high) second position of the cart its body **raised** above ground 275 millimeters.'

The low altitude of the vertical trajectory is denoted by the prefix VZ- in the verbs *vsporxnut'* 'fly up', *vskočit'* 'get up from sitting position', and *vsprygnut'* 'jump on'. The latter verb is illustrated in (6):

- (6) *Ded pytal'sja vsprygnut'* na podnožku tramvaja... [D. Rubina. Okna (2011)]
 'Grandfather tried to **jump onto** the footstep of the tram...'

The verbs *vzlezt'* 'climb up' and *vz'exat'* 'drive upward' typically also refer to destinations located at low altitude:

- (7) *Plavnym dviženiem ja svernul s šosse i <...> vz'exal vverx po travjanistomu sklonu...* [V.V. Nabokov. Lolita (1967)]
 'With a gentle movement I turned off the road and **drove up** the grassy hill.'

Not all motion verbs prefixed in VZ- can be unequivocally assigned a short vertical trajectory. However, verbs that denote flying like *vzletet'* 'fly up' or derivatives like *vzvit'(sja)* 'soar (of a bird), go up (of flags)' and *vzmyt'* 'rapidly move upwards (of a bird)' arguably have ingressive semantics and focus on the initial stage of rising into the air, which is a limited part of the path close to the surface. The verb *vsplyt'* 'surface, rise from bottom' highlights the surface area as well, but it puts in focus the final part of the trajectory and has resultative semantics.

Thus, the verb *vsplyt'* 'surface' refers to rising up to the surface, while *vzletet'* 'fly up' denotes rising from the surface: the former arrives to the surface, the latter departs from it. The difference between the two image schemas can be visualized as shown in

Figure 6, where the short black arrows mark the initial and the final parts of the trajectory, and the grey dashed background arrows show the potential complete paths.

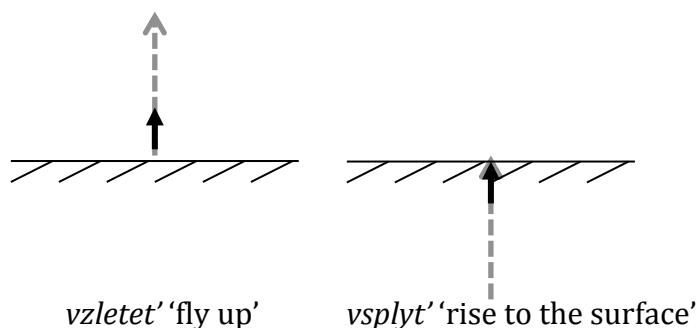


Figure 6: Image schemas of the verbs *vzletet'* 'fly up' and *vsplyt'* 'rise to the surface'.

In both cases we deal with locations close to the surface, which is crucial for VZ- as opposed to VOZ-.

7.5.3 Implications for semantics

I suggest that the proposed difference between VZ- and VOZ- in their prototypical spatial image schemas accounts for the difference in connotations and metaphorical effects of these prefixes.

The short length of the trajectory gives rise to negative connotations that often accompany verbs prefixed in VZ-, given that they often denote violation of a surface (*vzboronit'* 'furrow'), excitation of the original calm and undisturbed emotional state (*vzbesit'* 'madden'), or even damage and destruction (*vsporot'* 'rip open').

By contrast, the prefix VOZ- gives rise to abstract meanings assigning high status and bearing positive connotations (*vospet'* 'eulogize').

The short trajectory allows only a brief and often unexpected movement. Therefore, the low altitude encoded by VZ- explains why this prefix often carries semantic connotations of unexpectedness, abruptness, briefness, and intensity of an action (Zaliznjak & Šmelev 2000: 110). Thus, a remarkable number of connotations combined in a single marker VZ- is motivated by the spatial image schema encoded by this prefix. Moreover, the short trajectory makes it possible to observe both the starting point and the end-point of movement. I suggest that this motivates frequent and productive semelfactive use of this prefix (*vsmaxnut'* 'wave upwards once' < *maxnut'* 'wave'; *vzvizgnut'* 'scream once' < *vizžat'* 'scream').

By contrast, the trajectory of high altitude allows one to only see the point of departure. This is why VOZ- often focuses on the initial part of activity or a new state and contributes the ingressive reading, but not the semelfactive reading, as opposed to VZ- (*vossijat'* 'start shining' < *sijat'* 'shine'; *vozradovat'sja* 'start being glad' < *radovat'sja* 'be glad'). I return to this aspectual difference of VZ- and VOZ- and elaborate my argument in the discussion of 8.INGRESSIVE and 9.SEMELFACTIVE submeanings in 7.5.4.

7.5.4 Submeanings motivated by the prototype

Submeaning 2.VIOLATE A SURFACE, like the prototypical meaning 1.MOVE UPWARD, implies a very concrete spatial movement compatible with physical actions and activities. However, as opposed to the prototype, the upward movement is applied here not to the

entire object, but only to its uppermost part, the surface. In other words, the Trajector is the upper surface part of the Landmark (Lm). Therefore, I consider this submeaning to be a metonymical extension of the prototype 1.MOVE UPWARD. Figure 7 shows that the motion comes from below the surface and raises up its uppermost part.

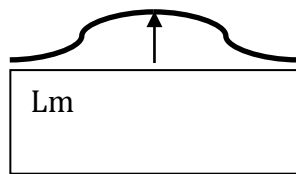


Figure 7: Image schema of submeaning 2.VIOLATE A SURFACE.

This meaning is expressed almost exclusively by the prefix VZ- and is attested in thirty-six standard verbs. By contrast, VOZ- in this meaning is attested only in one verb.

The base verbs that attach the prefix in this meaning describe various activities that are applied to different kinds of surfaces: ground, liquids, hair, skin, and other surfaces. The surface of the ground can be affected by plowing, digging, or scratching the soil, as in a whole series of verbs *vspaxat'* 'plow' (< *paxat'* 'plow'), *vskopat'* 'dig up' (< *kopat'* 'dig'), *vzryxlit'* 'make friable, loosen up' (< *ryxlit'* 'loosen'), and the verbs rarely attested in the corpus *vzboronovat'*, *vzboronit'*, and *vzborozdit'* all glossed as 'furrow' (< *boronovat'*, *boronit'*, *borozdit'* 'furrow'). Related are two factitive verbs *vsxolmit'(sja)* 'make hilly, raise ground into a hill' (< *xolm* 'hill') and *vzbugrit'(sja)* 'make hillocks' (< *bugor* 'hillock'). Whereas there are many verbs in this subgroup that are prefixed in VZ-, there is a single verb in VOZ-: *vozdelat'* 'till, cultivate (soil)' (< *delat'* 'do, make'), which is the only verb in VOZ- that represents the prefix meaning 2.VIOLATE A SURFACE.

Another substance with an upper surface relevant for the prefix VZ- is water: consider the derivatives *vskipet'* 'boil', *vskipjatit'(sja)* 'boil', *vzbalamutit'(sja)* 'stir up, trouble water', *vzboltat'(sja)* 'shake up (liquids)', *vspenit'(sja)* 'foam up'. The productivity of this subtype is evidenced by a number of marginal verbs attested in the corpus: *vzmutit'* 'make turbid', *vzrjabit'* 'disturb surface of water with wavelets', *vskolebat'(sja)* 'make waves', and *vspučinivat'sja* 'raise up' (< *pučina* 'bosom of the sea').

The "surface" of the human body is skin and thus we find the prefix VZ- in standard verbs like *vzdut'(sja)* 'swell' (< *dut'(sja)* 'blow') and *vspuxnut'* 'swell up' (< *puxnut'* 'swell'), and marginal verbs like *vspuzyrit'(sja)* 'blister up' (< *puzyrit'(sja)* 'blister'). Likewise, in *vspotet'* and *vzopret'* both denoting 'sweat' (< *potet'*, *pret'* 'sweat'), perspiration appears on the surface of the skin, and this facilitates the use of VZ-.

Hair or fur can be also conceptualized as a type of "surface". The prefix VZ- often applies to simplexes that denote tousling and bristling: *vz''erošit'(sja)*, *vzloxmatit'(sja)*, *vstrepat'*, *vstoporščit'(sja)* all denoting 'tousle up hair, ruffle, bristle'. The corpus attests a number of marginal verbs that arguably evidence productivity of this pattern: *vspušit'* 'fluff up', *vsxoxlit'* 'raise feathers', *vskuldyčivat'sja* 'bristle up', *vskosmatit'(sja)* 'make shaggy', *vz''eršit'(sja)* 'ruffle', and *vstopyrit'(sja)* 'make stick out, bristle'.

Other types of surfaces involve the human body: compare the standard verb *vspučit'(sja)* 'distend (of stomach)' and marginal verbs *vzgorbit'* and *vzgorbatit'* 'deform in the shape of a hump' (< *gorbit'*, *gorbatit'* 'hump').

Some verbs like *vzbit'(sja)* (< *bit'* 'beat') can apply to several types of surfaces: this verb describes shaking up pillows, fluffing up hair, whipping up cream. Similarly, many verbs discussed above can also describe raising of other surfaces than those which

they originally belong to. I illustrate this in (8), where the verb *vspučit'* 'distend (of stomach)' describes the parquet flooring in a house:

(8) *V moej gostinoj vspučilo parket, v spal'ne vzdulis' oboi.* [V. Soldatenko. Drugie opusy (2010)]

'The parquet flooring in my living room became **distended**, and in the bedroom the wallpaper became **bloated**.'

A separate subgroup includes verbs that denote damage or destruction of an object by violating its upper surface: *vsporot'(sja)* 'rip open' (< *porot'* 'rip, unstitch'), *vzrezat'(sja)* 'cut up the surface' (< *rezat'* 'cut'). Some verbs that belong here have lost a direct connection with the sense of a surface, namely *vzorvat'(sja)* 'explode' (< *rvat'(sja)* 'tear'), *vskryt'(sja)* 'open, unseal' (*kryt'* 'cover'), *vzlomat'(sja)* 'break' (< *lomat'* 'break'), *vzgret'* 'punish, scold' (< *gret'* 'heat'), *vzdrjučit'* 'beat with a stick' (< *drjučit'* 'thrash'). Marginal verbs that represent this use of VZ- include lexemes like *vzburavit'* 'drill up the surface' and *vzbučit'* 'make swell, beat up'.

Summing up, the submeaning 2.VIOLATE A SURFACE is very frequent and prominent for VZ- and applies to various kinds of surfaces. Productivity of this use for the prefix VZ- lends support to my claim that this prefix is specialized for short-scaled vertical movements. Because VOZ- is associated with a longer vertical path, it is not compatible with the concept of surface, which allows very short trajectory of a motion. This explains why the prefix VOZ- in the submeaning 2.VIOLATE A SURFACE is attested in only one verb.

Submeaning 3.AGITATE EMOTIONALLY is a metaphorical extension of the submeanings 1.MOVE UPWARD and 2.VIOLATE A SURFACE to the abstract domain of emotional states. The prefix VZ- carries this meaning in 18 standard verbs, and VOZ- is attested in 7 standard verbs.

Similar to how one can violate a surface of a substance like water, ground, skin, or hair, one can also disturb one's emotional serenity. The link between these two domains is evidenced by verbs that can be used in both concrete and figurative senses. A good example comes from the verb *vskipet'* which can mean both 'boil' and 'fly into a rage'. Likewise, some other verbs that describe turbulence of liquids have developed a specific metaphorical use based on mapping of the spatial schema of 2.VIOLATE A SURFACE onto the domain of emotional states: e.g. *vzbalamutit'(sja)* 'stir up, trouble water' > 'disturb, agitate a person'. Some verbs that developed this metaphorical use lost the initial spatial meaning. For example, the verb *vzvolnovat'* 'make worried' originates from the meaning 'disturb surface of water with waves', but can now refer only to emotional states, as opposed to its unprefixated base *volnovat'*, which retains both meanings 'make waves' and 'make worried'.

The shift to metaphorical use took place also in some verbs with VOZ-. For example, the verb *vozmutil'* 'perturb, outrage' can now only describe emotional states, whereas its counterpart in VZ-, *vzmutil'*, maintains the spatial meaning 'make turbid' based on the simplex *mutit'* 'trouble waters' (Evgen'eva 1999). The pair of these verbal counterparts in VZ- and VOZ- demonstrates the contrast between the two prefixes: the native VZ- refers to the concrete domain of spatial relations, whereas the loan VOZ- applies the same schema to the more abstract domain of emotional states.

Most verbs that manifest this use of VZ- bear a negative connotation of unexpected uncomfortable harassment caused by sudden intrusion: *vspoložit'(sja)* 'rouse, alarm' (< *položit'* 'rouse'), *vzbudoražit'* 'disarrange' (< *budoražit'* 'agitate, excite'),

vzbalamutit'(sja) 'disturb' (< *balamutit'(sja)* 'disturb'), *vzbesit'(sja)* 'infuriate, madden' (< *besit'(sja)* 'enrage, irritate'). However, there are exceptions like the verb *vzbodrit'(sja)* 'cheer up, encourage' (< *bodrit'(sja)* 'cheer up, encourage'), which lacks a negative connotation.

In addition to *vozmutil'* 'perturb', VOZ- expresses the submeaning 3.AGITATE EMOTIONALLY in verbs like *vozbudit'(sja)* 'excite, agitate' (< *budit'* 'wake'), *vosxilit'(sja)* 'delight' (< bound root attested in *poxilit'* 'hijack' and *xiščenje* 'thievery'), and *vosprjanut'* 'cheer up' (< bound root attested in *otprjanut'* 'spring back'). The use of VOZ- in these cases is arguably motivated by the prototype 1.MOVE UPWARD, which is reinterpreted in terms of orientation metaphors CONSCIOUS IS UP (e.g. *wake up*) and HAPPY IS UP (e.g. *I'm feeling up; spirits rose*) (Lakoff & Johnson 1980: 15; see Będkowska-Kopczyk 2012 for the discussion of parallel cases in Slovene). This metaphor has a clear physical basis: standing up is associated with being awake as opposed to lying down, which is associated with being asleep. This metaphor is apparently available for both prefixes VZ- and VOZ-. Emergence of an emotion is conceptualized in terms of upward motion. That is why the submeaning 3.AGITATE EMOTIONALLY can be expressed by both markers of upward path.

Submeaning 4.RESIST is a metaphorical extension of the prototype 1.MOVE UPWARD and is expressed by both prefixes VZ- and VOZ- in two and seven standard verbs respectively. In this meaning, the rising Trajector is interpreted as an obstacle or barrier on one's path. As a result, the upward motion is mapped to the domain of human relations and motivates a whole spectrum of confrontations, such as:

- disagreement in a conversation (*vozzazit'* 'raise an objection' < *razit'* 'strike')
- prohibition (*vospretit'* 'prohibit, forbid' < unclear base related to *pretit'* 'disgust'; *vozbranjat'(sja)* 'prohibit' < unclear base related to *branit'* 'scold, contradict')
- opposition (*vosprotivit'sja* 'oppose, resist' < *protivit'sja* 'oppose, resist'; *vosprepjatstvovat'* 'block, oppose' < *prepjatstvovat'* 'hinder')
- protest (*vz''erepenit'(sja)* 'bristle suddenly' < *erepenit'(sja)* 'bristle', *voztroptat'* 'start complaining against something' < *roptat'* 'complain')
- rebellion (*vzbuntovat'(sja)* 'incite to revolt' < *buntovat'(sja)* 'revolt, rebel', *vosstat'* 'revolt' (< *stat'* 'stand')

The semantic connection of 4.RESIST and 1.MOVE UPWARD can be seen in the verb *vosstat'* 'VOZ-stand' which has both the concrete spatial meaning 'rise' (e.g. 'rise in resurrection') and the metaphorical meaning 'rise in rebellion'. The concrete meaning 'rise' of this verb is now preserved only in idiomatic expressions like *vosstat' oto sna* 'rise from sleep', *vosstat' iz pepla* 'rise from ashes', and *vosstat' iz ruin* 'rise from ruins'. The verb *vosstat'* was in broader use in the 19th c. (Tabačenko 2011: 101) and could replace the verb *vstat'* 'get up, stand up' in contexts like (9):

- (9) *Vosstav poutru molčalivo, graf odevaetsja lenivo.* (A.S. Puškin "Graf Nulin", 1825)
'After having gotten up from bed silently in the morning, the count is lazily getting dressed.'

In contexts like (10), Modern Russian allows only *vstat* 'rise' (cf. Endresen & Plungian 2011), while *vosstat* is commonly used in the sense 'rise in rebellion'¹²³, abstracted from the physical motion that underlies its figural meaning:

- (10) [V] *Anglii vassaly vosstali protiv korolja*. [I. Pis'mennyj. Mjatežnaja pamjat' myšč // «Nauka i žizn' », 2006]
'In England the vassals **revolted** against the king.'

Submeaning 4.RESIST is also related to 3.AGITATE EMOTIONALLY. The verb *vz''erepenit'(sja)* 'bristle suddenly' (< *erepenit'(sja)* 'bristle') can be assigned both senses of the prefix. Moreover, many verbs of this group denote the inception of a protest and thus can be also interpreted as ingressesives with the prefixal meaning 8.INGRESSIVE: compare *vzbuntovat'(sja)* 'incite to revolt', *vosprotivit'sja* 'come to an opposition', and *vz''erepenit'(sja)* 'become very annoyed, start protesting', *vozroptat'* 'start complaining'.

In addition to standard Russian verbs discussed above, two marginal verbs are attested in the corpus: *vozbuxnut'* 'rebel against' (11) and *vz''erixorit'sja* (12) 'bristle up, protest against something'. They fully match the profile of 4.RESIST.

- (11) *Opasalsja, čto mastityj avtor vozbuxnet po oznakomlenii s publikaciej, no pozdnee ne vosposledovalo ni zvuka*. [M. Veller. Nožik Sereži Dovolatova (1997)]
'He was afraid of that the well-heeled author would **raise a claim** against the publication but not a single sound followed.'
- (12) *Tak i bojalsja, čto vz''erixoritsja, bešenij, isportit*. [A.I. Solženicy'n. Vse ravno (1993-1995)]
'[The captain] was afraid that he [technician], the crazy man, **would suddenly bristle up** and spoil it all.'

Summing up, submeaning 4.RESIST is attested for both prefixes, but VOZ- is more frequent. This submeaning is a metaphorical extension of the prototype 1.MOVE UPWARD. Standing up is interpreted as opposing a reverse force, thus gaining an additional sense of confrontation.

Submeaning 5. HIGH DOMINANT STATUS is another metaphorical extension of the prototype 1.MOVE UPWARD. The vertical axis of the spatial trajectory implied by the prototype is interpreted here as a hierarchical scale, and results in the concepts of high status, dominance, and control.

Two orientational metaphors are employed in this categorization, namely HIGH STATUS IS UP (*at the **peak** of one's career*) and HAVING CONTROL IS UP (*control **over** someone, being **on top** of the situation*) (Lakoff & Johnson 1980: 15).

Submeaning 5.HIGH DOMINANT STATUS is attested almost exclusively for the prefix VOZ-, which can be explained by its association with high vertical movement, as opposed to VZ-. In this light, 5.HIGH DOMINANT STATUS is the opposite of submeaning 2.VIOLATE A SURFACE, which is almost exclusively associated with VZ- because this prefix typically encodes a short upward trajectory.

¹²³ The sense 'rise in rebellion' can be traced back to the use of Old Russian verb *vъzstati* in 11th-14th cc. (Tabačenko 2011: 101).

There are 12 verbs that manifest submeaning 5.HIGH DOMINANT STATUS of the prefix. Some derivatives refer to taking the position of a leader (*vozglavit* ‘take over the leadership’ < *glava* ‘head, leader’) or a winner (*vostoržestvovat* ‘get to triumph’ < *toržestvovat* ‘celebrate a victory’). Transitive verbs denote elevating someone to a high position by praising, honoring, and glorifying: *vosxvalit* ‘glorify’ < *xvalit* ‘praise’, *vosslavit* ‘honor’ < *slavit* ‘honor’, and *vozveličit* ‘aggrandize, glorify’ < *velikij* ‘great’. The verb *vozomnit* means ‘get a false idea of one’s own importance’ and is formed from the base *mnit* ‘think, imagine’. The notions of influence and predominance are closely related to 5.HIGH DOMINANT STATUS, and we find them in the verbs *vozdejstvovat* ‘affect’ (< *dejstvovat* ‘act’) and *vozobladat* ‘prevail over’ (< *obladat* ‘possess’). The submeaning 5.HIGH DOMINANT STATUS is also implied in the imperfective verb *vossylat*, which designates sending something to an addressee of a higher status (to heavens, God, director) and formed from the simplex *slat* ‘send’. Likewise, the verb *vozzvat* ‘appeal to’ (< *zvat* ‘call’) implies a certain hierarchy: the one who asks for something has a lower status because the success of his or her request depends on the respond of the recipient.

The perfective verb *vozzvat* ‘appeal to’ prefixed in VOZ- has an imperfective partner verb *vzyvat* ‘appeal to’, which contains the prefix VZ-. This suggests that the prefix VZ- can also express the meaning 5.HIGH DOMINANT STATUS, being attested in a single verb. The fact that otherwise VZ- is not found in this meaning facilitates our analysis of this prefix in terms of a short trajectory not able to motivate such metaphorical extensions as the one discussed in this subsection.

Submeaning 6.BACK is mostly expressed by VOZ- (15 standard verbs), but is also attested for VZ- (3 standard verbs). I suggest that this meaning has two facets.

First BACK refers to a concrete spatial orientation that we observe in the verb *vozderžat*(*sja*) ‘restrain oneself, lit. ‘hold oneself back’ (< *deržat* ‘hold’) and the obsolete adverbial *vspjat* ‘backwards’ (< *vъz* ‘back’ + *pata* ‘heel, foot step’) (Vasmer 1971: v.1, 363). Here an actual backward movement takes place, as visualized in Figure 8:

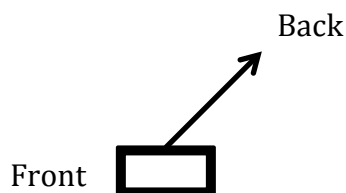


Figure 8: Image schema of spatial BACK.

The second facet of BACK translates the spatial relation into the temporal domain: here 6.BACK conceptualizes situations as objects¹²⁴ and refers to a return of something or someone to the state of affairs that took place earlier in time. This facet of the submeaning 6.BACK holds a presupposition that a Trajector has been separated from a Landmark and that it returns back to that Landmark in a concrete or a figural sense, as shown in Figure 9.

¹²⁴ EVENTS ARE OBJECTS metaphor is discussed by Lakoff & Johnson (1980: 30): “We use ontological metaphors to comprehend events, actions, activities, and states. Events and actions are conceptualized metaphorically as objects, activities as substances, states as containers.”

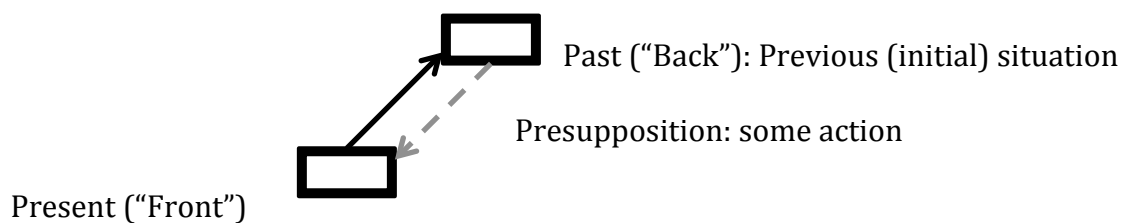


Figure 9: Image schema of temporal BACK.

The dotted grey arrow in Figure 9 designates an activity that took place as a presupposition. This activity brought about the present state of affairs (“Front” level: present). The black arrow designates the activity denoted by the prefixed verb: it goes back to the state of affairs that was there initially in the past (“Back” level). In this respect, the meaning 6.BACK is closely related to 8.INGRESSIVE because the BACK presupposes an initial situation that one comes back to as to a “beginning” of the present state of affairs. For this reason, Preobražensky (1964: 90) connects the verbs *vozvratit’* ‘return (to the initial state/place)’ (6.BACK) and *vozgoret’sja* ‘start burning’ (8.INGRESSIVE) as semantically related. Mahota (1993: 294) explicitly suggests that BACK should be interpreted as “motion back to the beginning”, or “turning back”. The image schema of 6.BACK in Figure 9 accounts for the verbs *vozvratit’* ‘give back’ (< *vorotit’* ‘turn’) and *vozvernut’(sja)* ‘return’ (< *vernut’* ‘return’) that denote a physical return of an object which was previously taken or borrowed. Example in (13) illustrates this use:

- (13) *Glavnoe, vozvernut’ obratno naši den’gi.* [A. Rostovskij. *Po zakonam volč’ej stai* (2000)] ‘Most importantly, to **return** our money back.’

Similarly, this meaning of the prefix can be observed in those verbs that refer to giving back someone’s due as in *vozdat’(sja)* ‘render’ (< *dat’* ‘give’), rewarding in response to a good deed as in *voznagradit’(sja)* ‘pay back’ (< *nagradit’* ‘reward’) and *vozblagodarit’* ‘return thanks’ (< *blagodarit’* ‘thank’), or compensation for spent resources, as in *vozmestit’(sja)* ‘compensate, refund’ (< *mesto* ‘place’¹²⁵, lit. ‘replace’). Moreover, it also accounts for bringing back a former union and reuniting separated parts, as in *vossoedinit’(sja)* ‘reunite’ (< *soedinit’(sja)* ‘unite’), or a response reaction gained by an object that is popular, desirable and in demand, as in *vostrebovat’(sja)* ‘reclaim, demand back, be in demand’ (< *trebovat’* ‘demand’).

In a similar way, VZ- and VOZ- are used in verbs denoting recollection, where the prefix in the meaning 6.BACK refers to bringing something forgotten back to mind through the work of memory: compare *vspomnit’(sja)* ‘recall’ (< *pomnit’(sja)* ‘remember’), a less frequently used verb *vospomnit’* ‘recall’ (< *pomnit’* ‘remember’), a standard nominalization *vospominanie* ‘recollection, remembrance’, *vspomjanut’(sja)* ‘recall’ (< *pomjanut’* ‘recollect, commemorate’). These examples are very important because they bring new light to the discussion of semantic differences between the two prefixes in question. Many linguists attribute the submeaning 6.BACK to the core semantic differences between VOZ- and VZ- in Modern Russian, claiming that this meaning can be expressed exclusively by VOZ-, but not by VZ- (Vinogradov et al. 1952: 580-581; Baranovskaja 1974: 123; Pilipenko 2001: 73; Efremova 2010: 543). The present study makes it possible to reformulate this claim in a more accurate way.

¹²⁵ According to Vasmer (1971: v.1, 334), *vozmestit’* ‘refund’ is more likely to be etymologically related to the noun *mesto* ‘place’ than to the verb *mstit’* ‘take revenge’.

Indeed, most verbs where the prefix contributes the submeaning 6.BACK are prefixed in VOZ-. Yet, as mentioned above, there are at least two verbs in VZ- that belong to this subgroup as well, namely *vspomnit'(sja)* and *vspomjanut'(sja)* 'recall', which are semantically compatible with the derivatives in VOZ-: *vospomnit'* 'recall' and *vospominanie* 'recollection'. Another verb that arguably manifests the meaning 6.BACK of the prefix VZ- is *vzyskat'(sja)*, which denotes 'force to pay the debt' (< *iskat'* 'search') and is semantically similar to the verbs in VOZ- *vozdats'(sja)* 'render, give someone his due' and *vozmestit'* 'compensate, refund'. Thus, the prefix VZ- can also express the submeaning 6.BACK, although rarely. I conclude that the two prefixes do overlap in this semantic zone.

There is another controversy that should be discussed in more detail. Some scholars who explored the use of VOZ- in Modern Russian make a fine-grained distinction between the meaning BACK and the meaning AGAIN. Possible definitions of the meaning AGAIN describe 'an action that produces something anew, again' (Baranovskaja 1974: 123) or 'performing an action again, anew' (Pilipenko 2001: 72; Efremova 2010: 543). The verbs that are claimed to represent this meaning of the prefix are very few, usually limited to the five lexemes listed in (14):

- | | |
|--|---------------------------------------|
| (14) <i>vozbnovit'(sja)</i> 'restore anew, re-establish' | < <i>obnovit'(sja)</i> 'renew' |
| <i>vosstanovit'(sja)</i> 'restore' | < <i>stanovit'(sja)</i> 'set, become' |
| <i>vossozdat'(sja)</i> 'restore anew' | < <i>sozdat'</i> 'create' |
| <i>vosproizvesti(s')</i> 'reproduce' | < <i>proizvesti(s')</i> 'produce' |
| <i>vozrodit'(sja)</i> 'revive' | < <i>rodit'(sja)</i> 'give birth' |

Before I turn to my account of these verbs, let me give a short summary of the historical background of submeanings BACK and AGAIN.

Diachronic studies show that the submeaning BACK is very old and is attested in Old Church Slavonic: compare the verb *vъz-dati* 'give back' (Xaburgajev 1974: 330), *vozderžat'sja* 'restrain oneself', literally 'hold oneself back' (< *deržat'* 'hold') and the adverbial *vspjat'* 'backwards' (from *vъz-patъ* < *vъz* 'back' + *patъ* 'foot') (Vasmer 1971: v.1, 363). Mahota (1993: 292) points out that both the meaning BACK and the meaning 'UP[WARDS]' are unquestionable in the semantics of the Old Church Slavonic VЪZ-. There is good evidence showing that both submeanings of this formant were present in the early Slavic translations from Greek, where VЪZ- primarily corresponds to ἀνα- (*ana*) denoting 'up[wards], increase, strengthening, repetition/improvement, backwards' (ibid: 292).

The meaning AGAIN, on the contrary, is not attested in Old Church Slavonic and is not found in other Slavic languages except some words in Belarusian and Bulgarian that were borrowed from Russian (Mahota 1993: 293). The reason is that, as Mahota 1993 convincingly showed, the prefix VOZ- began to be used in the meaning AGAIN only in the eighteenth century, when it was employed for translations of French and Latin words with the prefix RE- 'do anew, back'. Since the Slavic VOZ- and the Romance RE- share the submeaning BACK (compare Russian *vozmezdije* – Latin *recompensatio* 'retribution'; Russian *vozvratit'* – Latin *redhibeo, reverto* 'return'), it became possible to employ VOZ- for expression of the other submeaning of RE- as well, namely its repetitive meaning 'do AGAIN'. Mahota 1993 proposes that the meaning AGAIN of VOZ- is in fact a semantic loan from French/Latin secondary use of the prefix RE-. He showed that the use of VOZ- in the loan meaning AGAIN was productive but short-lived, limited to the period from the mid 18th c. up to the beginning of the 19th c. During this short period of several decades,

Russian introduced a set of neologisms which were morphological calques of French and Latin high-style abstract terms. Most of these words did not make it to Modern Russian: *vosperit'sja* 'be reledged (of birds)' (< French *se replumer*), *vossostavlenie* 'chemical recomposition', *voznačinat'* 'begin again', *voživotvorit'sja* 'resurrect', *vožjavit'sja* 'appear again'.¹²⁶ However, some of the new coinages from that time remained: *vozroždenie* 'renaissance' (Latin *renascor, regeneratio*), *vosproizvodit'* 'reproduce' (< French *reproductive*), *vossozidanie* 'reconstruction' (< Latin *reconstructio*), *vozobnovit'* 'renew'. Since the meaning AGAIN is not an etymological meaning of the prefix VOZ- but rather an artificial semantic loan that is no longer productive in Modern Russian, it is problematic to consider it a distinct submeaning of this prefix. Moreover, I propose that those five verbs in (14) that historically contain the prefix VOZ- in the meaning AGAIN (according to Mahota 1993) in Modern Russian can be analyzed in terms of the image schema BACK.

On the one hand, indeed, the verbs in (14) presuppose that something was gone, lost, destroyed, or forgotten and that a new round of a constructive activity restores something that was ruined or revives in the previous state of affairs: *vossozdat'(sja)* 'restore anew' (< *sozdat'* 'create'), *vosstanovit'(sja)* 'restore' (< *stanovit'(sja)* 'set, become'). On the other hand, I argue that the verbs *vozobnovit'(sja)* 're-establish' and *vozrodit'(sja)* 'revive' refer to re-establishment of an activity (collaboration, negotiations) or a state (tradition) that was previously stopped, and this re-establishment is conceptualized as coming back to something that took place before. In particular, the verbs *vozobnovit'(sja)* 're-establish' and *vozrodit'(sja)* 'revive' do not literally mean 'give birth again' or 'make new again' but rather 'come back to something that existed before it disappeared or was lost'. In a similar fashion, *vossozdat'(sja)* 'restore anew', *vosstanovit'(sja)* 'restore', and *vosproivesti(s')* 'reproduce' can also be viewed as returning back to the result of some activity achieved previously by an original round of activity. Figure 10 visualizes how the image schema of BACK can account for these verbs with semantics of AGAIN:

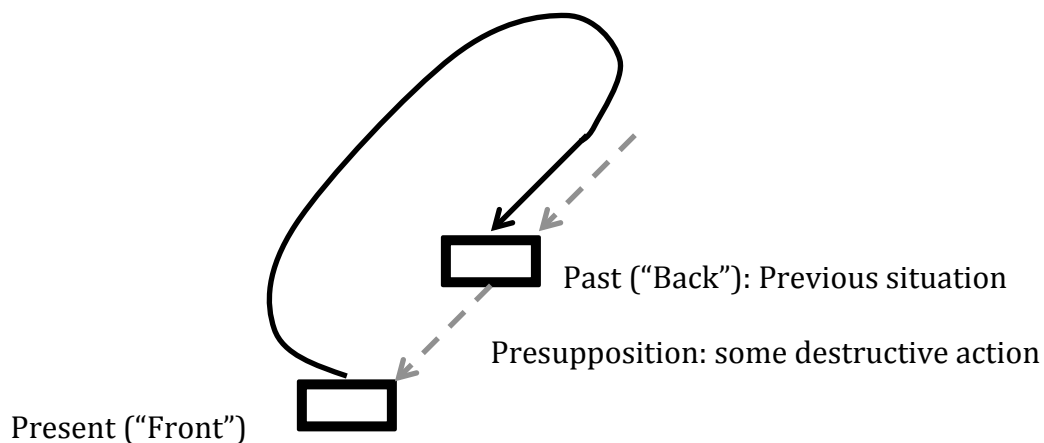


Figure 10: Image schema of temporal BACK that accounts for AGAIN.

Verbs like *vossozdat'* 'restore anew' refer to conducting a second, "new", round of the activity *sozdat'* 'create' (see the black arrow parallel to the dotted grey arrow in Figure 10). This "new" activity is performed in order to come back to the result (see the

¹²⁶ Mahota (1993: 301) provides an explicit quotation from Suxomplinov's "Istorija Rossijskoj Akademii" (1874-78) referring to Bogdanovič's discussion from 1789-92: "*Xristos vozjavilsja, t.e. vnov' javil suščnost' svoju*" 'Jesus Christ **resurrected/appeared again**, that is **showed** his nature **anew**'.

box of Past in Figure 10) achieved previously (stage 1: first dotted arrow) and subsequently lost (stage 2: second dotted arrow). Thus, I suggest that the semantics of the five verbs that potentially contain VOZ- in submeaning AGAIN can be accounted for by the image schema of VOZ- in submeaning BACK showed in Figure 10. In other words, the meaning BACK is present in these verbs along with AGAIN.

An interesting case is the verb *voskresnut'* 'resurrect, lit. come back to being alive, return'¹²⁷, where VOZ- arguably denotes BACK, but not AGAIN. Apparently, the root of this verb is not *krest* 'cross' but rather *kres* (from **krěsъ* and **krъsъ* 'coming alive' < **krěpso-*, **krъpso-*) (Vasmer 1971: v.2, 372). Similarly, the etymological dictionary by Preobražensky (1964: 382) lists the unprefixated *krěsiti* 'reanimate' and the dialectal noun *kresъ* 'resting; turn of time, solstice'. Accordingly, *Slovar' russkogo jazyka 11-17 vv.* (1981: v.8, 38) defines *krěsiti* as 'reanimate, revive, make alive': *Reče že imъ Olga: ljuba im estъ rěčъ vaša, uže mně muža ne krěsiti* 'Olga told them: I like your speech, but my husband cannot **revive**' from the Laurentian Chronicle (945). It is clear that in Modern Russian the verb *voskresnut'* 'resurrect' underwent deetymologization. The same holds for its causative correlate *voskresit'(sja)* 'make resurrect'. Yet Mahota (1993: 294) argues that the meaning of the prefix here is BACK.

In this discussion I mentioned that the meaning 6.BACK is closely related to 8.INGRESSIVE via the sense of returning to an initial state of affairs. At the same time, it is not clear how 6.BACK is related to 1.MOVE UPWARD. The relationship of these two meanings can be seen in the ambiguity of the verb *vosxodit'* 'VOZ-walk'. On the one hand, this verb can denote upward movement, as in (15), and on the other hand, it can describe tracing something back to earlier distant events in the past and thus exhibit the 6.BACK meaning of the prefix, according to the TIME IS SPACE metaphor (Haspelmath 1997), as in (16).

(15) [O]n *vosxodit* na prestol... [«Domovoj», 2002]
'He becomes a king (lit. he **goes up** onto the throne).'

(16) [P]odobnye slova mogut *vosxodit'* k očen' drevnim istokam [S. Starostin, G. Zelenko. U čelovečestva byl odin prajazyk, 2003]
'Such words might **date back** to very ancient sources.'

How UPWARD can turn into BACK/PAST is a puzzle. Here UPWARD is interpreted as 'earlier in time' and refers to a source or an origin. This temporal precedence in terms of a spatial vertical dimension might come from a model of a genetic tree (e.g. genetic tree of language families), where the preceding stages of history are located upward, and the descendants are situated downward. Similarly, when writing a text, one can refer to the preceding paragraph as located "above", and the following passage as coming "below". Šarić (2012: 206) suggests that since both UPWARD and BACK refer to embodied experience, their relationship can be explained in terms of different perspectives of body postures: when a person is standing, what is up is above the person's head. From another perspective what was up can be now viewed as being behind the person's head. Šarić 2012 observed the submeaning BACK in the semantics of the Croatian prefix and preposition UZ(-), related to Russian VZ- and VOZ-. She suggests that etymologically the

¹²⁷ Vasmer (1971: v.2, 372) lists the root morpheme *kres* as a source-root for *voskresenje* (Old Russian *вѣскрѣсеніе*, *вѣскрѣшеніе* with either *ъ* or *ѣ* in the root) 'resurrection' and *voskresiti* 'make resurrect'.

meaning BACK could be a semantic influence of the Greek prefix *ἀνα-* (*ana*) ‘up, back’ via the mediation of Old Church Slavonic translations.

Summing up this subsection, I make following conclusions: 1) The submeaning 6.BACK can be expressed by both VZ- (*vspomnit’* ‘recall’) and VOZ- (*vozvratit’* ‘return’), although it is more characteristic of the latter. Thus, the traditional view, according to which 6.BACK can be exclusively assigned to the prefix VOZ-, is disproved by the data; 2) One can distinguish between two facets of 6.BACK: spatial (*vozderžat’(sja)* ‘hold back’) and temporal (*vspomnit’* ‘recall’) based on the cognitive metaphors TIME IS SPACE and EVENTS ARE OBJECTS; 3) The submeaning 6.BACK is related to 1.MOVE UPWARD and 8.INGRESSIVE; 4) There is no need to recognize a separate submeaning AGAIN in verbs like *vossozdat’* ‘restore’ because they can be analyzed as representing the image schema 6.BACK.

Submeaning 7.GROW UP is attested for both prefixes: VZ- contributes this meaning in 6 standard verbs, whereas VOZ- does so in 5 standard verbs.

The process of growing implies increasing the height of an object, which is usually measured along a vertical scale. Therefore, growing is directly related to the upward movement. As one grows, one rises up from the original size, which is the point of departure. In other words, the submeaning 7.GROW UP is a metaphorical extension of the prototypical meaning 1.MOVE UPWARD that comes about via the common orientational metaphor MORE IS UP; LESS IS DOWN: “if you add more of a substance to a container or pile, the level goes up” (Lakoff & Johnson 1980: 15). This explains why the prefixes in question can extend the prototypical meaning 1.MOVE UPWARD to 7.GROW UP.

As shown in diachronic studies of derivational morphology, a spatial meaning can develop metaphorical extensions when a spatial marker is applied to bases of various non-spatial semantic classes (Nefed’ev 1995). Because both VZ- and VOZ- can attach to simplexes that denote feeding, cherishing, and raising, which alone do not imply the physical increase of an object’s height, upbringing, or maturation, I propose that this use of the two prefixes should be viewed as a separate submeaning 6.GROW UP, extended from the semantic prototype 1.MOVE UPWARD.

The semantic relation between the meanings 1.MOVE UPWARD and 7.GROW UP can be seen in the following example. A simplex base *rasti* ‘grow’ refers to concrete physical growth. When the prefix VZ- combines with this base, the prefix semantics can be interpreted as the prototypical 1.MOVE UPWARD, because growing makes its way up, especially in terms of the development of a living being: *vzrastit’* ‘foster’ (< *rastit’* ‘grow’), *vzrasti* ‘grow up’ (< *rasti* ‘grow’), as in (17). A parallel derivative in VOZ-, the verb *vozrasti* ‘increase’ (< *rasti* ‘grow’), goes beyond this concrete physical pattern and applies to objects that are inanimate (prices, volumes, numbers), as illustrated in (18).

(17) *Dar’ja moja tut <...> vzrosla, tut i sostaritsja.* [V. Astaf’ev. Oberton (1995-1996)]
‘My Darja **grew up** here and will grow old here.’

(18) *V rezul’tate nalogovaja nagruzka na malyj biznes vozrosla.* [«Bogatej», 2003.04.24] ‘As a result, the tax burden for small businesses **grew.**’

Other prefix+base combinations produce more abstract notions including raising a child via physical fostering with food and drink (the prefix VZ-), as well as psychological growth via ethical upbringing and emotional maturation (the prefix VOZ-). Interestingly, there is a division of labor among the two prefixes in question. The prefix VZ- combines

with bases that denote concrete physical impact and care: *kormit* ‘feed’ > *vskormit* ‘nurture, raise’, *poit* ‘give to drink’ > *vspoit* ‘foster’, *lelejat* ‘cherish’ > *vzlelejat* ‘foster’, *materet* ‘grow physically’ > *vzmateret* ‘grow up to physical maturity’. By contrast, the prefix VOZ- is observed in derivatives with more abstract meaning implying teaching certain ethical rules and behavioral patterns, as in *vospitat’(sja)* ‘bring up’, or both physical and emotional maturation, as in *vozmuzhat* ‘become mature, reach manhood’.

However, in spite of these subtle differences, in some verbs the conventionalized prefix can sometimes be replaced with its correlate without any semantic deformation. Compare the verb *vzlelejat* ‘nurture, foster’ (19) that has 74 attestations in the Modern Subcorpus of the RNC with its counterpart *vozelelejat* ‘cherish’ (20) that has three hits:

(19) *Ty menja vzlelejala i vskormila, ty načalo moego sčastja i slavy...* [Istorija odnogo pesennogo sjužeta // «Narodnoe tvorčestvo», 2003]
 ‘You **cherished** me and fostered me, you are the beginning of my happiness and glory’.

(20) *A proletariat razberetsja sam v tom, čto sozdano talantlivogo, vse, čto vozelelejano ljubov’ju...* [Ju. Elagin. Temnyj genij. (1998)]
 ‘And the proletariat will figure out itself what is created of talent and everything that was **cherished** with love.’

Summing up, the prefixed verbs that represent the submeaning 6.GROW UP of VZ- and VOZ- refer to both physical raising, fostering, growing up as well as upbringing and maturation. This submeaning is a zone where the two prefixes overlap and sometimes are even interchangeable with no serious difference in meaning.

Submeaning 8.INGRESSIVE of the prefixes VZ- and VOZ- is very old and can be traced back to their source prefix VʹBZ-, which is well attested in the ingressive meaning already in early East Slavic texts (Mahota 1993: 292). Illustrative examples in (21) and (22) are given from the Ostromir Gospel:

(21) *jako vsplačete se i vʹzrydaiete vy a mirъ vʹzradouetъ se* (John 16:20)
 ‘...that you **will [begin to] weep** and **[begin to] sob**, and the world **will rejoice**’.

(22) *ięzyki vʹzg[lag]liotъ novy...* (Mark 16:17)
 ‘... they **will [begin to] speak** in new languages...’

The 8.INGRESSIVE meaning is closely related to the prototype 1.MOVE UPWARD: vertical movement is associated with emergence of an object (e.g. *vozniknut* ‘uprise, emerge’ < *niknut* ‘lean down’, cf. Pilipenko 2001: 61). By rising over the line of surface or horizon, a Trajector becomes visible to an observer. In this light, uprising equals coming into existence. The spatial schema of an uprising Trajector, metaphorically mapped to the domain of events (EVENTS ARE OBJECTS metaphor described in Lakoff & Johnson 1980: 30), motivates the 8.INGRESSIVE meaning as the inception of a new state, a new activity, or an event.

The semantic link between 1.MOVE UPWARD and 8.INGRESSIVE is acknowledged by Preobražensky (1964: 90), who defines the overall semantics of the prefixes VZ- and VOZ- as “directed motion upward or to the beginning”, thus bringing together the sense of upward motion and the sense of the initial stage of a situation or an activity. Other

scholars including Isačenko¹²⁸ (2003/1965: 231), Baranovskaja (1974: 129), and Pilipenko (2001) pointed out that the 8.INGRESSIVE use of VZ- and VOZ- is semantically motivated and etymologically based upon the meaning 1.MOVE UPWARD.

Both prefixes VZ- and VOZ- can refer to the beginning phase of an action. However, it has not been clearly defined in the scholarly literature how the two prefixes are distributed across ingressive verbs. Most scholars agree that both prefixes feature certain expressivity when used in the 8.INGRESSIVE sense. However, the expressivity of VZ- does not equal that of VOZ-. Yet both expressive flavors become particularly apprehensible when contrasted with neutral garden-variety ingressive verbs prefixed in ZA- and PO-: e.g. *vozgoret'sja* vs. *zagoret'sja* 'start burning'; *vskričat'* 'cry out, exclaim' vs. *zakričat'* 'start shouting'; *vozljubit'* 'come to love' vs. *nevzljubit'* 'start disliking' vs. *poljubit'* 'start loving' (< *ljubit'* 'love').

The beginning of an action can be expressed in Russian in a number of ways, where each marker contributes a specific facet of ingressivity. The prefix ZA- is usually employed to mark the beginning of homogeneous situations, where the beginning, the middle and the final part of an activity do not differ from each other (so-called **inchoative** aktionsart) e.g. *zasvistet'* 'start whistling' (Zaliznjak & Šmelev 2000: 107). The prefix PO- is commonly associated with **ingressive** aktionsart: e.g. *poletet'* 'start flying', *pobežat'* 'start running' (Zaliznjak & Šmelev 2000: 109). Verbs in PO- often combine the beginning of a state with a result that marks entering a new state: *poljubit'* 'start loving', *počuvstvovat'* 'experience a feeling'. Sometimes, the ingressive meaning can be also expressed by the circumfixal construction RAZ-...-SJA: e.g. *razrydat'sja* 'burst into tears, start crying', *rasserdit'sja* 'fly into a rage, become angry', where the ingressive phase consists in a rapid accretion ("growth", the Russian term "razrastanie") of an emotion into a dominant feeling (Zaliznjak & Šmelev 2000: 111).

It was pointed out in many accounts that the prefix VZ- has its own profile of ingressivity which is very accurately defined in Zaliznjak & Šmelev (2000: 110) as **explosive** actionsart ("éksplozivnyj sposob dejstvija"). Zaliznjak and Šmelev (ibid: 110) describe it as a "special evocative expressive connotation that highlights abruptness and unpredictability of an action which often escapes control, and at the same time highlights the intensity of the performance".¹²⁹ We can observe this connotation of VZ- by comparing *vskričat'* 'cry out, exclaim' with the neutral marking of the beginning phase in *zakričat'* 'start crying'. Similarly, Isačenko (2003/1965: 230) concludes that the ingressive VZ- is very expressive and simultaneously denotes abruptness of action.

However, according to Baranovskaja (1974: 122), the sense of intensity is present in ingressive verbs prefixed in VOZ- as well: *vozlikovat'* 'rejoice', *vozradovat'sja* 'become glad'.

Moreover, both VZ- and VOZ- are claimed to be unproductive (Isačenko 2003/1965: 230) in ingressive verbs and thus might not deserve much attention. However, plentiful corpus attestations of new coinages disproves this view. In this light, it remains problematic how to differentiate between the two prefixes with respect to their expressivity and abstract grammatical semantics.

¹²⁸ Isačenko (2003/1965: 231) points to the link between the senses 1.MOVE UPWARD and 8.INGRESSIVE comparing a similar semantic development of prefixes in other languages: German *auf-schreien* 'exclaim, start shouting', English *speak up* 'start speaking'.

¹²⁹ The original wording is "osobyj ékspressivnyj ottенок, podčerkivajuščij vnezapnost' i nepredskazuemost' dejstvija, často ne poddajuščegosja kontrolju, i odnovenno ego intensivnost' (sr. *vskričat'* i *zakričat'*)" (Zaliznjak and Šmelev 2000: 110). (translated by myself – A.E.)

I propose that the difference between VZ- and VOZ- in this respect can be best understood and captured in terms of metaphorical mapping of their spatial image schemas (recall Sections 7.5.2 and 7.5.3). In particular, I argue that the grammatical dissemblance of VZ- and VOZ- is directly motivated by the difference within their spatial prototype, namely the length of the path they refer to. **The longer upward trajectory of VOZ- motivates gradual entry into a new state of affairs, while the short trajectory of VZ- justifies abruptness and unexpectedness of a rapid momentary phase, which often constitutes an event of its own.**

What makes this motivational link between spatial trajectories and temporal and aspectual connotations possible is the cognitive metaphor TIME IS SPACE (Haspelmath 1997), the tendency to structure and conceptualize time in terms of spatial relations. In this light, it becomes clear that the longer path takes longer time to get through, and the shorter path takes shorter time. So, it is natural that the prefix VOZ-, which designates the higher upward motion (as opposed to VZ-), is associated with INGRESSIVE aktionsart and thus holds a perspective over both the inception of a new state and what follows. The prefix VZ-, the short trajectory operator, on the contrary, naturally marks a momentaneous event and triggers SEMELFACTIVE interpretation. This semantic correspondence of variation in the spatial prototype on the one hand and the difference in Aktionsarten on the other hand is summarized in Figure 11.

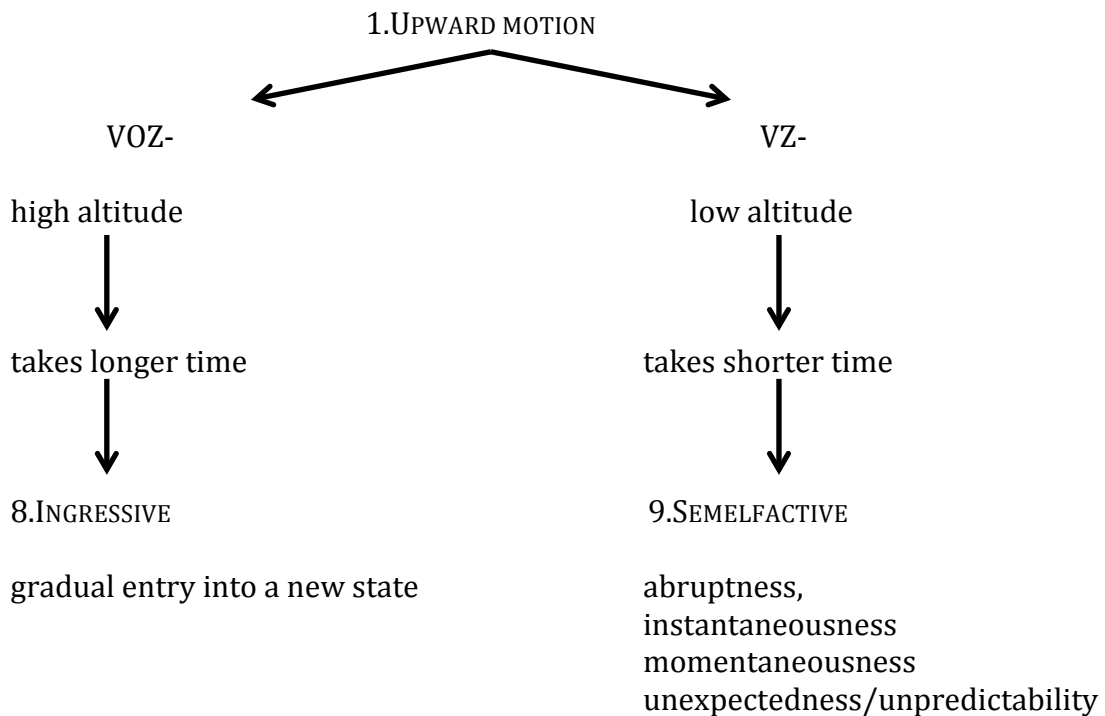


Figure 11: Spatial prototype variation motivating the difference in Aktionsarten.

Examination of the data reveals the following pattern: most ingressive verbs in my database denote inception of a new emotional state and bear the prefix VOZ-, while those verbs that are prefixed in VZ- refer to an abrupt production of sounds or basic bodily acts and should be rather interpreted as semelfactives (see next subsection on 9.SEMELFACTIVE).

This distinction accounts for the majority of verbs. Crucially, it accounts for INGRESSIVE vs. SEMELFACTIVE reading of many marginal verbs attested in the corpus. At the

same time, this distinction should be understood in terms of prototypical contrast between VZ- and VOZ-, which in some cases is blurred and diffuse. We cannot dismiss the minority of peripheral verbs that belong to the “grey” zone of this contrast and can be interpreted both ways (INGRESSIVE vs. SEMELFACTIVE) despite the profile determined by their prefix.

The 8.INGRESSIVE meaning can be expressed by both VZ- and VOZ-, but the prefix VOZ- is more prominent in this meaning: among standard ingressive verbs there are twenty-two prefixed in VOZ- and only three prefixed in VZ-. Most prototypical ingressive verbs with VOZ- denote entering an emotional state, emergence of a feeling (23) or a wish (24):

- (23) *vozradovat'(sja)* ‘start being glad’ (< *radovat'(sja)* ‘be glad’)
vozveselit'sja ‘start being glad’ (< *veselit'sja* ‘rejoice’)
voznenavidet'(sja) ‘start hating’ (< *nenavidet' ‘hate’*)
voznegodovat' ‘resent’ (< *negodovat' ‘resent’*)
vozljubit' ‘come to love’ (< *ljubit' ‘love’*)
vozgordit'sja ‘start being proud of’ (< *gordit'sja* ‘be proud of’)
vosčuvstvovat' ‘start feeling’ (< *čuvstvovat' ‘feel’*)

- (24) *vozželat'(sja)* ‘start wanting’ (< *želat' ‘want’*)
vozžažadat' ‘start wanting, get a desire’ (< *žažadat' ‘desire’*)
vosxotet' ‘start wanting’ (< *xotet' ‘want’*)
voznamerit'sja ‘take into one’s head to do smth’ (< *namerit'sja* ‘plan to do’)
vozmečtat' ‘start dreaming, wanting’ (< *mečtat' ‘dream’*)

Note that all these ingressive verbs imply that the inception of a new emotional state is followed by the designated emotional state, that is the emergence of a feeling continues with the process of experiencing the feeling, as I illustrate by the example in (25):

- (25) *Trudis' smirenno. I togda, byt' možet, vosčuvstvuesš' drugoe «ja»...* [Ju. Davydov. Zagovor, rodivšij myš' (1993)]
 ‘Work humbly. And then you might **come to the feeling** of the other “you” (alter ego).’

Ingressive meaning of this kind can be expressed by the prefix VZ- only marginally, profiling the abruptness of emergence of the new state. In verbs like *vzalkat' ‘start wanting’* (< *alkat' ‘want’*), *nevzljubit' ‘start disliking’* (< *ljubit' ‘love’*), and *vzrevnovat' ‘become jealous’* (< *revnovat' ‘feel jealousy’*) the effect brought in by the prefix can be compared with the job of a momentary action switch. No doubt, it makes it problematic to define whether these verbs denote the beginning of a new state or an immediate action.

Another type of ingressive verbs prefixed in VOZ- includes verbs that denote inflaming and subsequent burning and shining: *vossijat' ‘start shining’* (< *sijat' ‘shine’*), *vospylat' ‘catch fire, become inflamed with passion’* (< *pylat' ‘flame, burn’*), *vozžeč' ‘light, inflame’* (< *žeč' ‘burn’*), *vozgoret'sja ‘inflame’* (< *goret' ‘burn’*), and *vosplamenit'(sja) ‘inflame, set on fire’* (< *plamja* ‘fire’):

- (26) *Skoro, skoro vossijaet solnce pravdy nad zemlej russkoj.* [A. Alekseev. Sud'ba reformatora v epoxu kataklizmov // «Nauka i žizn'», 2007]

‘Soon, soon there **will shine** the sun of truth over the Russian land.’

Note that all verbs of burning can be used in figurative sense referring to an emergence of a passionate feeling. The same is true of the verb *vspylit*’, parallel to the above-mentioned *vospylat*’. This pair illustrates the contrast of VZ- vs. VOZ-. I suggest that the verb *vospylat*’ is ingressive and denotes ‘catch fire; become inflamed with passion’, whereas its counterpart *vspylit*’ means ‘suddenly become angry, fly into a rage’ and refers to a brief and sudden expression of one’s temper. In other words, attached to the same simplex base *pyl* ‘flame’ or *pylat*’ ‘be on fire’, the prefix VZ- turns the verb into a semelfactive. The use of this verb is illustrated in (27):

- (27) *Tumanov byl čelovek s xarakterom i srazu že vspylil – kak éto, kakaja-to pičuga delaet emu, izvestnomu režisseru, zamečanja!* [I. Arxipova. *Muzyka žizni* (1996)]
‘Tumanov was a man of character, and so he **immediately flew into a rage** – how come such a little bird dares to reprove him, a well-known director.’

Likewise, the verb *vspyxnut*’ ‘inflammate, flash, break out (of fire)’ refers to a sudden explosion and arguably represents semelfactive aktionsart rather than ingressive.

Among twenty-nine marginal ingressive verbs, there are twenty-one verbs prefixed in VOZ- and eight verbs prefixed in VZ-. Derivatives in VOZ- include verbs denoting a start of emotional state: *vossočuvstvovat*’ ‘start sympathizing’ (< *sočuvstvovat*’ ‘sympathize’), *vosskorbet*’ ‘start bewailing’ (< *skorbet*’ ‘mourn’), *vožžalet*’ ‘start feeling pity’ (< *žalet*’ ‘feeling pity’), *vozalkat*’ ‘start wanting’ (< *alkat*’ ‘want’), *vozrevnovat*’ ‘start being jealous’ (< *revnovat*’ ‘be jealous’). The latter verb is illustrated in (28):

- (28) *Zametiv, čto diakon pol’zuetsja ljubov’ju prixožan <...>, ona vozrevnovala i voznenavidela ego.* [Z. Maslennikova. *Žizn’ otca Aleksandra Menja.* (1992)]
‘Having noticed that the deacon is loved by the congregation, she **became jealous** and started hating him.’

Submeaning 9.SEMELFACTIVE should be recognized as one of the key submeanings of the prefixes VZ- and VOZ- (especially the former). Most of verbs in this group in addition to the prefix have also the suffix -NU-, which is a well-known semelfactive marker in Modern Russian (Dickey & Janda 2009; Makarova & Janda 2009; Kuznetsova & Makarova 2012): e.g. *vskriknut*’ ‘shout loud once’ (< *kriknut*’ ‘shout once’, *kričat*’ ‘shout’).

One might disagree with my statement that VZ- and VOZ- can express submeaning 9.SEMELFACTIVE and say that the semelfactive meaning in such verbs is contributed exclusively by the suffix -NU-, while the function of the prefix is somewhat different. Yet there are at least three arguments that support my analysis.

First, a prefix can also contribute semelfactive meaning, for example, the prefix S-, as in *svastat*’ ‘boast once’ (< *xvastat*’ ‘boast’) (Dickey & Janda 2009).

Second, the suffix -NU- and a prefix can both have the SEMELFACTIVE meaning and work together as a morphological construction. In this case, they attach to a simplex base simultaneously as a circumfix. For instance, the verb *vsplaknut*’ ‘shed a tear, weep a little’ contains both the prefix VZ- and the suffix -NU-, which attach to the verbal simplex base *plakat*’ ‘cry’, since there is no existing verb **plaknut*’. Similarly, *vzdremnut*’ ‘take a nap for a short while’ comes from *dremat*’ ‘drowse, nap’, because there is no **dremnut*’. The verb *vsrapnut*’ ‘snore once, take a short nap’ is formed from *xrapet*’ ‘snore while

sleeping', but not **xrapnut*'. In these verbs, the prefix and the suffix work together as a morphological construction VZ-...-NU with semelfactive semantics.

Third, sometimes there is no semelfactive suffix on semelfactive verbs: compare the verbs *vskričat* 'exclaim' (< *kričat* 'shout'), *vzvýt* 'set up a howl' (< *výt* 'howl, wail'), *vzlajat* 'bark once' (< *lajat* 'bark'), *vzroptat* 'express a complaint' (< *roptat* 'complain, grumble'). These verbs imply an action that does not last long. On the contrary, it takes a very short time and is usually unique in the context of a narrative. This is illustrated in (29):

(29) – *Dumaju, čto vy sdureli. – Sdureeš!* – *nervno vskričal* Gusakov. – *S vami sdureeš!* [V. Bykov. Boloto (2001)]

'I think that you've gone mad! – One can go mad! – Gusakov **shouted** nervously. – It's possible to go mad from dealing with you!'

Note that I analyze the verb *vskričat* 'exclaim' as semelfactive rather than ingressive, contrary to how it is usually interpreted. I suggest that the verb *vskričat* 'exclaim' does not mean 'start shouting', but rather 'shout once', and so it is usually accompanied with someone's reply in a narrative, as in (29).

Keeping in mind the three arguments given above, I conclude that 9.SEMELFACTIVE should be recognized as a distinct submeaning of the prefix VZ-. This holds also for the prefix VOZ-, although to a lesser degree. The next question is how this submeaning is motivated in the network of polysemy.

I follow Nessel (2013) in understanding semelfactivity not as a categorical (plus / minus) notion but rather as a radial network with more prototypical and more peripheral members. In recognizing semelfactive verbs, I use the four properties of semelfactives summarized by Nessel in terms of 1) uniformity, 2) instantaneousness, 3) non-resultativity, and 4) single occurrence.

In the examined data, the submeaning 9.SEMELFACTIVE is represented in a substantial group of verbs, namely 20 standard Russian verbal lexemes including 16 verbs prefixed in VZ- and 4 verbs prefixed in VOZ-. Among marginal verbs this submeaning is realized even more frequently, showing high productivity of the pattern with the prefix VZ-. The corpus attests forty-six VZ-verbs and at least six VOZ-verbs suggesting that Russian speakers easily use this pattern in order to coin novel semelfactive verbs on the fly.

Among semelfactive verbs prefixed in VZ- and VOZ- I distinguish among several subgroups, in particular verbs that denote a) sound and speech, 2) physical bodily acts, and 3) behavior. I will address each subgroup in turns.

The most populous subgroup in this inventory is sound and speech verbs. Here we find a lot of verbal lexemes prefixed in VZ- that denote short rapid sounds, often produced intensively and unexpectedly: *vzvizgnut* 'scream once' (< *vizgnut* 'scream once'), *vzvýt* 'set up a howl' (< *výt* 'howl, wail'), *vskričat* 'exclaim' (< *kričat* 'shout'), *vsxlipnut* 'sob once' (< *xlipat*, *xljupat* 'sniffle, sob'), *vzlajat* 'bark once' (< *lajat* 'bark'). The verb *vsxrapnut* can refer to both sound ('snore once') and behavior ('take a short nap') (< *xrapet* 'snore while sleeping'). The prefix VOZ- is also attested in semelfactive verbs of sound and speech: *voskliknut* 'exclaim' (< *klikat* 'call'), *vozglasit* 'exclaim' (< *glasit* 'read', *glas* 'voice'), *vozopit* 'shout loud once' (< *vopit* 'shout'), *vostrubit* 'announce in triumph' (< *trubit* 'trumpet, blow').

Another subgroup of semelfactive verbs includes lexemes that denote bodily acts. These are physical actions that can be performed repeatedly, for example *vzdoxnut* 'take

a breath, sigh once' (< *doxnut* 'take a breath once', *dyšat* 'breathe'), *vzdrognut* 'shiver once' (< *drognut* 'shake once with cold, fear', *drožat* 'shake'), *vzgljanut*(*sja*) 'look at once' (< *gljanut* 'take a look at', *gljadet* 'look at'), *vzbryknut*(*sja*) 'kick out' (< *brykat*'*sja* 'kick'), *vzmetnut*(*sja*) 'throw upward' (< *metnut* 'throw once', *metat* 'throw'), *vzmaxnut* 'wave upwards once' (< *maxnut* 'wave once', *maxat* 'wave').

- (30) *Dirižer **vzmaxnul** paločkaj. Orkestr zaigral "Sentjabr' v Pariže."* [S. Dovlatov. Inaja žizn' (1984)]
 'The conductor **waved** the wand. The orchester started playing "September in Paris".'

The third semelfactive group consists of verbs that refer to emotional and other behavior: *vzdumat*(*sja*) 'suddenly make an unexpected decision' (< *dumat* 'think'), *vzgrustnut*(*sja*) 'suddenly feel sad' (< *grustnyj* 'sad'), illustrated in (31), and *vzdremnut*(*sja*) 'take a nap for a short while' (< *dremat* 'take a nap').

- (31) *Radost' – na nule. Konstruktor čego-to vdruk **vzgrustnul**.* [V. Šukšin. Pečkilavočki. (1970-1972)]
 'Joy – on zero level. – All of a sudden the engineer **felt sad**.'

Another member of this behavior subgroup is the verb *vzrevnovat* 'suddenly feel jealousy for a short moment' (< *revnovat* 'feel jealousy'). This verb has twenty-one attestations in the corpus. A typical example of its use is given in (32):

- (32) *Pomnitsja, v kakoj-to moment ona daže **vzrevnovala**: Saška znal o Stase kuda bol'se, čem ona, zakonnaja žena.* [M. Djačenko, S. Djačenko. Magam možno vse (2001)]
 'I remember at some moment she even **turned jealous**: Saška knew much more about Stas than she, his lawful wife.'

The most productive semelfactive pattern in marginal verbs is represented by multiple coinages in VZ- that denote singularized animal sounds: *vsxrjuknut* 'give a grunt' (< *xrjuknut* 'give a grunt', *xrjukat* 'grunt'), *vspisknut* 'squeak shortly once' (< *pisknut* 'squeak once', *piščat* 'squeak'), *vzmjavknut* 'say meu once suddenly' (< *mjavkat* 'say meu (of a cat)'). Many marginal verbs denote human sounds like *vsxripnut* 'make a crackling sound, croak once suddenly' (< *xripet* 'make a crackling sound, croak'). Some marginal verbs of this pattern denote sounds produced by inanimate objects: *vzbrjaknut* 'clink shortly once' (< *brjaknut* 'clink, hit', *brjakat* 'clink, ding'), *vzgremet* 'growl once shortly' (< *gremet* 'rattle, growl').

7.6 Radial Category Profiling

The prefixes VZ- and VOZ- can be modelled as a single semantic network where they share all nine submeanings. However, it is problematic to claim that these prefixes are semantically identical. First, although VZ- and VOZ- share the spatial prototype 1.MOVE UPWARD, they differ in terms of altitude of the encoded trajectory. Second, the difference in the spatial schemas causes further divergence of VZ- and VOZ- in other submeanings of the network. In this section, I approach this issue quantitatively and compare the

distributions of the two prefixes across their sub-uses in terms of numbers of verbs that represent them.

Table 3 provides the numbers of standard verbs that feature submeanings of the two prefixes.

We can see that in the spatial prototype 1.MOVE UPWARD both prefixes are entrenched, but VZ- is attested in a larger number of verbs: 34 in VZ- as opposed to 14 in VOZ-. Likewise, the other spatial meaning 2.VIOLATE A SURFACE is almost exclusively manifested in verbs with VZ-.

In the group of metaphorical submeanings (3-7) the prefixes are distributed in the following way. The prefix VZ- predominates in 3.AGITATE EMOTIONALLY, whereas VOZ- takes over in 4.RESIST, 5.HIGH DOMINANT STATUS, and 6.BACK. The meaning 7.GROW UP is equally frequent for both prefixes.

Finally, VZ- and VOZ- are distributed across the abstract grammatical meanings: VZ- predominates in submeaning 9.SEMELFACTIVE, whereas VOZ- is more prominent than its alternative in submeaning 8.INGRESSIVE.

Type	Prefix submeaning	VZ-	VOZ-	VZ-/VOZ-	VZ- %	VOZ- %
Spatial submeanings	1.MOVE UPWARD	34	14	1	29%	15%
	2.VIOLATE A SURFACE	36	1	0	31%	1%
Metaphorical submeanings	3.AGITATE EMOTIONALLY	18	7	0	15%	8%
	4.RESIST	2	7	0	2%	8%
	5.HIGH DOMINANT STATUS	0	11	1	1%	12%
	6.BACK	3	15	0	3%	16%
Aspectual submeanings	7.GROW UP	6	5	0	5%	5%
	8.INGRESSIVE	3	22	0	3%	24%
	9.SEMELFACTIVE	16	4	0	14%	4%
	Unclear	0	5 ¹³⁰	0	0%	5%
Total:		118	91	2	100%	100%

Table 3: Distribution of standard verbs with VZ- and VOZ- (raw numbers and %).

The values in Table 3 are too sparse to run a statistical analysis, so I calculated the relative frequencies in terms of percentages.

Table 3 shows that most verbs in VZ- employ this prefix in the submeanings 2.VIOLATE A SURFACE (31%), 1.MOVE UPWARD (29%), 3.AGITATE EMOTIONALLY (15%) and 9.SEMELFACTIVE (14%). These uses of VZ- account for 89% of verbs formed by this prefix.

The prefix VOZ- invest in another set of submeanings: 8.INGRESSIVE (24%), 6.BACK (16%), 1.MOVE UPWARD (15%), and 5.HIGH DOMINANT STATUS (12%). This set of meanings comprises 67% of all attestations of VOZ- in standard verbs.

These sets of frequently attested submeanings constitute the Radial Category Profiles of the two prefixes. They are visualized in Figure 12. Note that the prominent uses of both VZ- and VOZ- include the prototypical meaning 1.MOVE UPWARD. This facilitates the close semantic relationship of VZ- and VOZ-.

¹³⁰ The verbs that are unclear in terms of prefix submeaning include: *vopol'zovat'sja* 'make use of' < *pol'zovat'sja* 'make use of'; *vostrebovat'(sja)* 'reclaim' < *trebovat'* 'demand'; *vozymet'* 'conceive (wish, intention)' < *imet'* 'possess', *vozmoc'* 'become able to perform something' < *moc'* 'be able'.

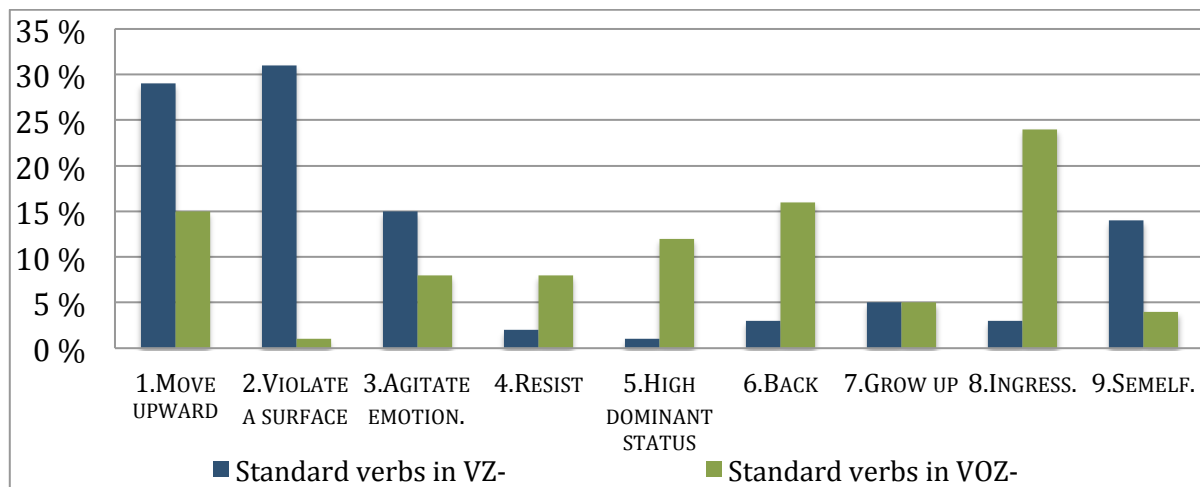


Figure 12: Radial Category Profiles of the prefixes VZ- and VOZ- (%).

Submeanings of the prefixes VZ- and VOZ- represent three groups, as shown in Table 3: submeanings 1 and 2 are spatial, submeanings 3-7 are metaphorical, and submeanings 8 and 9 are aspectual. Within these three semantic groups, the two prefixes show different kinds of behavior. In this sense, what predominates in the behavior of VZ- and VOZ-, similarity or divergence, largely depends on the domain of prefix behavior.

In spatial and metaphorical meanings, we observe both semantic overlap and differentiation between VZ- and VOZ- in terms of their predominance in attested verbs. The two prefixes are more similar in verbs where they carry the submeaning 1.MOVE UPWARD than in verbs that manifest the submeaning 2.VIOLATE A SURFACE. Likewise, VZ- and VOZ- are semantically more compatible in verbs where they designate the metaphorical meaning 7.GROW UP than in those metaphorical meanings that favor either VZ- (3.AGITATE EMOTIONALLY) or VOZ- (4.RESIST, 5.HIGH DOMINANT STATUS, 6.BACK).

In aspectual submeanings, we observe differentiation of the two prefixes: 9.SEMELFACTIVE strongly prioritizes VZ-, whereas 8.INGRESSIVE prominently favors VOZ-.

Summing up, the quantitative analysis of data in terms of Radial Category Profiling shows that VZ- and VOZ- have both distinct semantic profiles and a zone of semantic overlap.

7.7 VZ- and VOZ- in minimal pairs

As mentioned above, the two prefixes in question are not complementarily distributed across verbs. There is a group of verbs that form minimal pairs, the zone where VZ- and VOZ- overlap in their distribution. How large is this zone? In my database, there are 21 verbal bases that form derivatives with both prefixes. In other words, 42 verbs have a counterpart with the other prefix (VZ- or VOZ-) which yield 11% of the overall number of verbs attested with these prefixes in the corpus. These 21 minimal pairs of verbs that differ in prefix include thirteen pairs of standard verbs and eight pairs where one of the correlates is marginal and rarely attested in the corpus.

Many minimal pairs of verbs in VZ- and VOZ- have been discussed in detail in 7.6. In the present section, I focus on the variety of factors that determine the choice of the prefix in minimal pairs.

In the verbs *vzvesti* 'raise a hammer on a gun' vs. *vozvesti* 'bring to an elevated position, erect a building', the difference in prefix is the difference in the encoded altitude of the upward motion. The same type of contrast can be observed in the pair of

imperfectives *vsxodit'* vs. *vosxodit'* 'raise' discussed in 7.3: VZ- arguably refers to a short trajectory, whereas VOZ- is more appropriate for describing a motion directed to a higher destination.

In the pair *vstat'* 'stand up' vs. *vosstat'* 'rise in rebellion', the native VZ- carries the spatial meaning 1.MOVE UPWARD, whereas the Slavonic VOZ- refers to the metaphorical meaning 4.RESIST. Similarly, in the pair *vzmutit'* 'make turbid, stir up water' vs. *vozmutil'* 'agitate emotionally', VZ- denotes the spatial meaning 2.VIOLATE A SURFACE and VOZ- encodes the metaphorical meaning 3.AGITATE EMOTIONALLY. Thus, here the contrast lies in the opposition of two domains: concrete spatial relations vs. abstract non-spatial behavior.

In some cases like *vspylit'* 'fly into a rage' vs. *vospylat'* 'become inflamed with passion' and *nevzljubit'* 'start disliking' vs. *vozljubit'* 'come to love', the choice of the prefix might be determined by the evaluative connotations it carries: VZ- is often associated with negative and destructive situations, whereas VOZ- elevates the proposition designated by the verb to high status implying significance and esteem.

The minimal pair *vsplaknut'* 'shed a tear, weep a little' vs. *vosplakat'* 'start crying' exploits the aspectual difference between the two prefixes. The choice of VZ- is conditioned by the semelfactive reading of the target derivative, whereas VOZ- is chosen to designate ingressive aktionsart.

In the pairs *vspomnit'* – *vospomnit'* 'recall' and *vzjat'* – *vozymet'* 'take, obtain', the contrast between VZ- and VOZ- can be described in terms of register: the prefix VOZ- contributes a bookish flavor, whereas the derivatives in VZ- are stylistically neutral.

In some minimal pairs the verbs in VZ- and VOZ- have compatible semantics but differ in token frequency. For example, in the pair *vzdymat'* vs. *vozdymat'* 'raise' (822 vs. 14 attestations in the corpus respectively), the verb in VZ- is more entrenched in use. By contrast, the verb in VOZ- is more entrenched in the pair *vzrasti* vs. *vozrasti* 'grow, increase' (62 vs. 4,385 attestations, accordingly).

Summing up, minimal pairs of verbs prefixed in VZ- and VOZ- show that the presence of one prefix as opposed to its alternative can be due to different factors: encoded image schema of low vs. high altitude, the lexicalized use of the prefix in either primary (spatial) or secondary (metaphorical) meaning, aspectual or register differences between VZ- and VOZ-, as well as frequency effects.

7.8 Conclusions

The data presented in this chapter challenges the traditional concept of allomorphy in two respects.

First, the application of the semantic criterion to this data is problematic because the two prefixes are polysemous. On the one hand, both VZ- and VOZ- are attested in all nine submeanings. On the other hand, they have clearly distinct profiles within the shared network of submeanings. In other words, the two candidates for allomorphy are both semantically identical and semantically different.

Second, the data violates the second criterion for allomorphy, complementary distribution, by exhibiting an overlap in 21 minimal pairs of verbs including 13 pairs of standard verbs. Strictly speaking, this alone should preclude a monomorphemic account of these prefixes, because the traditional model of allomorphy does not allow any deviations. However, this option misses out the strong correlation between the two prefixes and their association with one another.

I suggest that it is more fruitful to analyze the status of VZ- and VOZ- in terms of Non-Standard Allomorphy, because such an account affords us a more balanced view over both similarities and differences of the two closely related formants. Both prefixes are well attested in Modern Russian verbs. Both of them share the spatial prototype of upward movement, a set of metaphorical extensions, and aspectual uses. Both prefixes are polysemous and share all nine submeanings. This facilitates their close semantic relationship and compatibility. At the same time, in each domain of use (spatial, metaphorical, and aspectual) the two prefixes have developed distinct profiles in terms of what kind of submeanings they express most frequently in attested verbs. Thus, the Russian language employs the partial formal difference of the two former phonological variants in order to designate subtle differences in meaning. At the same time, the formal similarity of VZ- and VOZ- makes it possible to associate the two prefixes with one another as two varieties of a single unit.

The case of the prefixes VZ- and VOZ- is anything but trivial. These prefixes share the spatial prototype of upward movement, which motivates all other submeanings, but at the same time they encode different altitudes of the trajectory. The native prefix VZ- implies a short upward trajectory, whereas the loan prefix VOZ- is associated with a path of higher altitude. I propose that this subtle difference in the spatial image schemas of the two prefixes motivates their measurably different Radial Category Profiles. The differences between VZ- and VOZ- are manifested in their metaphorical extensions, evaluative connotations, and aspectual uses.

Chapter 8

The prefixes VY- and IZ- ‘out of a container’

8.1 Introduction: What is special about this case

This chapter presents another pair of prefixes consisting of a native East Slavic prefix VY- and a loan Slavonic prefix IZ-. These prefixes co-existed and co-evolved over many centuries (at least since the 9th cent).

What makes this pair of prefixes special and different from similar pairs discussed in previous chapters, is the zone of their distributional overlap which is very large. Among 576 standard verbal lexemes formed by these prefixes, 223 (39%) lexemes have an alternative counterpart with the other prefix (e.g. *vytoptat’* – *istoptat’* ‘trample all over’ < *toptat’* ‘trample’), whereas 114 standard verbs in IZ- and 238 verbs in VY- do not allow such “prefix alternation”. This means that about 25% (112) of simplex bases can form a standard prefixed verb with either VY- and IZ-.

Judging from the high degree of overlap in their distribution, we expect the two prefixes in question to be semantically distinct from each other and represent candidates for two distinct morphemes rather than two allomorphs. However, in this chapter I will argue for the opposite kind of analysis, contrary to most common views of VY- and IZ- that see them as different morphemes.

What is highly problematic with such a clear-cut distinction is the overabundance of derivatives where VY- and IZ- contribute near-identical meaning. Compare pairs of interchangeable equivalent verbs like *vykupať’sja* – *iskupať’sja* ‘have a swim’ < *kupať’sja* ‘swim’, *vymotať* – *izmotať* ‘exhaust’ < *motat’* ‘shake’, *vymazat’* – *izmazat’* ‘smear up all over’ < *mazat’* ‘smear’, *vypravit’* – *ispravit’* ‘correct’ < *pravit’* ‘edit’. Other pairs of verbs, on the contrary, show a semantic contrast between the two prefixes, as in *vylomat’* (*dver’*) ‘break open’ (a door) vs. *izlomat’* ‘break an object in many places’, both formed from *lomat’* ‘break’.

The question of correlation between VY- and IZ- in terms of semantics is controversial and remains open. A number of recent studies bring attention either to VY- or IZ- in terms of cognitive linguistics accounts (Botvinik 2009; Pozolotina 2009) or semantic invariants (Dobrušina et al. 2001: 59-70). The diachronic distribution and development of VY- and IZ- in Russian is traced in Belozercev 1966 and Dadavaeva 1978. However, there is no comparative study that defines the modern status and relationship of VY- and IZ-.

The goal of this chapter is to provide a synchronic contrastive analysis of the two prefixes in order to find out how much they are similar in terms of semantics and what factors determine their distribution across verbs. The answer to these two questions would bear a theoretical implication about the status of these prefixes in Modern Russian – whether VY- and IZ- are distinct synonymous morphemes or suppletive allomorphs of a single morpheme.

The study I report on in this chapter was started as a joint project in collaboration with Tore Nessel and Laura A. Janda. Our major findings were reported in Nessel et al. 2009, 2011 and in a condensed version in Endresen et al. 2012: 266-

272; Janda et al. 2013: 61-66. These publications are based on the analysis of 277¹³¹ verbs with VY- and IZ- centered around the problem of semantically “empty” aspectual prefixes in Russian. By contrast, in this chapter I report on a study of 989 verbal lexemes which include both standard and marginal verbs attested in the Russian National Corpus. Apart from testing our semantic model against a larger dataset, I direct the discussion to address the criteria for allomorphy.

The chapter is structured as follows. In 8.2, I provide relevant background information about the diachronic origins of VY- and IZ-. The data is presented in 8.3, followed by analysis of semantics in 8.4 and analysis of distribution in 8.5. I put my findings together in the discussion section in 8.6 and summarize my conclusions in 8.7.

8.2 Formal similarity and historical relation

As opposed to the pairs of prefixes analyzed in previous chapters, VY- and IZ- have clearly different phonological shapes that are completely unrelated in terms of the active phonology and morphophonology of Modern Russian.

Moreover, the two prefixes have presumably never shared a single etymological source. Both prefixes have origins in Proto-Slavic, but their etymological sources are different. In the literature, IZ- is traditionally referred to as a Church Slavonic equivalent of the native Russian VY- (Berneker 1924: 440; Townsend 1968: 125; Miklosich 1970: 97, 397; Vasmer 1971; Dem’janov 2001: 336). Townsend (ibid: 125) suggests that “this accounts for the near identity of the meanings of the two prefixes and for the somewhat more abstract character of verbs in *iz-*.” However, the correlation of VY- and IZ- is not purely a matter of Russian vs. Slavonic equivalents.

Belozercev 1966 states that verbs in IZ- with the spatial or metaphorical meaning of outward movement like *izletet* ‘fly out’, *ispustit* ‘emit’, *iskupit* ‘redeem, expiate’, *istorgnut* ‘extract’ have been borrowed into Old Russian from Church Slavonic texts. Such verbs are highly idiomatic and are restricted to particular contexts with abstract nouns which refer to the notions of life, soul, death, guilt, etc. By contrast, there are verbs prefixed with “another” IZ-, native and widely attested in Old Russian, which expresses a non-spatial meaning of exhaustive completion of an activity: e.g. *izorvat* ‘tear up into many pieces’ < *rvat* ‘tear’. The picture is complicated even more due to a group of verbs with IZ- that existed in both Church Slavonic and Old Russian and therefore have two readings according to the origin of IZ-: e.g. in old texts the verb *iznyrjati* ‘IZ-dive’ could denote ‘resurface, come out’ with the Slavonic IZ- ‘OUT OF’ and ‘dive everywhere, dive a lot’ with the native to Russian IZ- ‘EXHAUSTIVE RESULT’. These findings are robust and are based on a thorough analysis of a large body of texts of different genres created in 11-17th cc. (Belozercev 1966).

Dadavaeva 1978 analyzed pairs of parallel verbs prefixed in VY- and IZ- in the context of the evolving norms of standard Russian of the second half of 18th century. Dadavaeva found that in some pairs, the verbs in IZ- gradually went out of use (*izletet* ‘fly out’, *issypat* ‘strew out’), whereas in other cases the counterparts in VY- and IZ-

¹³¹ In these studies it was important to work with a balanced, representative, and manageable dataset. According to the methodology that we adopted in these studies, we analyzed all Natural Perfectives and those Specialized Perfectives formed by VY- and IZ- that had over 100 attestations in the RNC.

developed semantic (e.g. *vyvesti* ‘lead out’ vs. *izvesti* ‘use up’) and/or collocational differences (e.g. *izzit’* ‘get rid of e.g. traditions’ vs. *vyžit’* ‘get rid of e.g. neighbors’).

To sum up, it would be wrong to claim that the use of the prefix IZ- can be explained exclusively by borrowing of this prefix from Church Slavonic. It is more accurate to consider that the modern use of IZ- in Russian integrates two different strata – loan and native. This explains the coexistence of obsolete verbs in IZ- with bookish flavor like *izvergnut’* ‘extrude’ on the one hand and evocative verbs in IZ- productive in colloquial speech like *ispereživat’sja* ‘exhaust oneself by being too worried’ or *izloxmatit’* ‘make very shaggy’ on the other hand. Today speakers apparently do not distinguish between the loan IZ- and the native IZ- and perceive all instances as a single prefix IZ-. This speculation is supported by the fact that even in thorough linguistic analyses of IZ- in Modern Russian, scholars do not normally connect different uses of this prefix with different historical origins. Moreover, all uses of IZ- are analyzed as belonging to a single morpheme and no reference to different origins of these uses is made (Švedova et al. 1980: v.1 § 862; Zaliznjak & Šmelev 2000: 91; Dobrušina et al. 2001: 59-64; Aver’janova 2008: 7).

Although VY- and IZ- are not cognate prefixes and are not similar phonologically, they are closely associated with one another due to the distributional patterns that were formed during the long history of their co-existence in Russian.

8.3 Data: Insights from numbers

This study is based on 989 verbal lexemes collected from the Russian National Corpus. This total number includes 579 (59%) verbs prefixed in VY- and 410 (41%) verbs prefixed in IZ- (including also IS- and IZO-). As in other case corpus studies presented in this dissertation, I use the Modern Subcorpus, which consists of texts created in 1950-2014 and contains over 123 million words. The database is available at <http://hdl.handle.net/10037.1/10078>.

In organizing this list of verbs, I follow the methodology described in detail in the previous case studies. I assign secondary imperfectives to their perfective partners and merge reflexive detransitivized verbs in *-sja* with their transitive counterparts. Therefore, I address verbal lexemes which can be manifested in a number of forms. A typical lexeme in the database looks like the verb *vyčistit’(sja)* ‘clean up’ which comprises a transitive perfective *vyčistit’_{PF.TR}*, a transitive secondary imperfective *vyčiščat’_{IPF.TR}*, a reflexive perfective *vyčistit’sja_{IPF.INTR}*, and a reflexive secondary imperfective *vyčiščat’sja_{IPF.INTR}*. Imperfective verbs are included in the database only if they lack a perfective partner attested in the corpus (*vyčixivat’* ‘sneeze smth out’) or in the language altogether (*izobilovat’* ‘abound’).

Both prefixes are attested in verbs of high and low token frequency. Therefore, in order to avoid any bias, in my analysis I distinguish between standard lexemes shared by most speakers and marginal coinages that might be unfamiliar to speakers. The threshold token frequency that works for this dataset is 9 corpus attestations. I consider verbs with nine and more attestations to be standard Russian verbs, and verbs with 8 and fewer attestations as marginal.

This distinction yields the pattern of prefix distribution shown in Table 1.¹³² At this point, we can make two important observations. First, among verbs prefixed in VY- there are more standard lexemes (374=65%) than among verbs in IZ- (202=49%). This suggests that VY- is more entrenched in the standard Modern Russian lexicon than IZ-. Although this difference itself is generally expected for the pair of Russian (native) and Slavonic (loan) prefixes, it is surprising that the difference is not dramatic.

Word Type	Token Frequency	Verbs in VY-	Verbs in IZ-	Total
Standard	≥ 9	374 (65%)	202 (49%)	576
Marginal	< 9	206 (35%)	207 (51%)	413
Total		580 (100%)	409 (100%)	989

Table 1: Standard and marginal verbs prefixed in VY- and IZ-.

The second observation concerns non-standard verbs. IZ- forms as many marginal verbs as VY- in terms of raw numbers (207 and 206 respectively), and has a much higher percentage rate in terms of relative distribution (51% and 35% respectively). One could wonder whether for IZ- the number of marginal verbs includes many obsolete verbs rather than novel coinages indicative of prefix productivity. Surprisingly, this is not the case. Among 207 marginal verbs in IZ- there are only 17 lexemes that can be considered obsolete and archaic: e.g. *izniknut* ‘disappear, die out’, *izronit* ‘drop a word or comment’, *izrinut* ‘extract’. By contrast, the vast majority of verbs prefixed with IZ- are newly formed coinages based on productive derivational patterns like *ispereživat’sja* ‘exhaust oneself by being too worried’ and *izèkonomit’sja* ‘exhaust oneself by saving money’ or *izuzorit* ‘decorate all over with patterns’ and *istatuirovat* ‘tattoo all over the place’. This suggests a non-trivial conclusion: IZ- does not lose out to VY- in terms of productivity. Moreover, both prefixes are highly productive and comparable in this regard. The next question is whether VY- and IZ- are productive in the same or different meanings, and I address this issue by means of semantic analysis (8.4.1) and profiling (8.4.2).

Table 2 addresses the aspectual properties of the two prefixes. It contains the numbers of Natural Perfectives and Specialized Perfectives¹³³ (Janda 2007) that these prefixes form. First, Table 2 indicates that IZ-, despite its partly Slavonic roots, is a strong perfectivizer which competes with the native counterpart VY- in forming Specialized Perfectives.

Prefix	Specialized Perfectives (SP)			Natural Perfectives (NP)		
	Standard	Marginal	Total	Standard	Marginal	Total
VY-	276	165	441	92	14	106
IZ-	162	201	363	36	3	39

Table 2: Standard and marginal verbs prefixed in VY- and IZ-.

¹³² This distributional difference between VY- and IZ- across standard and marginal verbs is statistically significant. A chi-squared test yields $X^2=22$; $df=1$, p -value = 2.272e-06. Cramer’s V equals 0.15, showing that the effect size of this difference is small.

¹³³ Here, Specialized Perfective is taken in a broad sense and understood as any perfective verb other than Natural Perfective. Therefore, in Table 2 Specialized Perfectives include perfective verbs that are formed from perfective simplexes like *vydat*_{PF} ‘hand out’ < *dat*_{PF} ‘give’.

Second, we observe that IZ- forms only 39 Natural Perfectives (i.e. aspectual pairs in the strict sense), whereas VY- creates a total of 106 Natural Perfectives. As we know from Janda et al. 2013 and Endresen et al. 2012, in Natural Perfectives the meaning of the verbal base overlaps with the meaning of the prefix, and this creates the effect of so called “empty” aspectual prefixes: e.g. VY- ‘OUT OF’ + *polot*_{IPF} ‘weed **out**’ yields *vypolot*_{PF} ‘weed out’. The fact that IZ- forms fewer aspectual pairs than VY- suggests that most of its uses are too specific or abstract to allow semantic overlap with the base.

In the analysis of polysemy presented in the next section I restrict myself to analyzable perfective verbs in VY- and IZ-, a total of 949 lexemes. This number includes 368 standard and 179 marginal lexemes in VY- as well as 198 standard and 204 marginal lexemes in IZ-. I exclude from my analysis all imperfective verbs. Among them there is the highly frequent imperfective verb *vygljadet* ‘look like’ which lacks a perfective counterpart arguably due to its loan origin. This verb is a morphological and semantic calque of the German *aussehen* ‘look like, lit. out-see’ that came to Russian in the 1830s (Vinogradov 1999: 118). I also exclude a productive pattern manifested exclusively by imperfective verbs¹³⁴ like *vydelyvat’sja*, *vykobenivat’sja*, *vykablučivat’(sja)*, where all such derivatives refer to showing off, doing something uncommon, extraordinary, or outrageous. I also put aside four imperfectives with IZ- that lack perfective partners: *izobilovat* ‘abound’ and three synonymous verbs *izmyvat’sja*, *izdevat’sja*, and *istjazat’(sja)* ‘torture, bully’.

8.4 Near-identical semantics

In the semantic analysis of VY- and IZ- I aim at answering the following questions: Do these prefixes differ in terms of their central submeaning – i.e. semantic prototype? In particular, do these prefixes have the same prototype or different prototypes? Overall, do these prefixes share all submeanings, the majority of submeanings or only a minor part of submeanings? How can we measure the degree of their semantic similarity and semantic difference?

8.4.1 Unwrapping polysemy

I propose that both prefixes in question can be accommodated within a single semantic model. This model is a radial network of twelve related submeanings organized around a central prototype, as shown in Figure 1. Submeanings are visualized as rectangles. Each submeaning is given a label and is illustrated by an example. The numbers in parentheses indicate the numbers of standard verbs where a given submeaning of the prefix is represented. This diagram includes all standard verbs prefixed in VY- and IZ-, that is 566 lexemes.

Figure 1 indicates that in Modern Russian the prefix IZ- inhabits a subset of submeanings attested for VY- (10 submeanings out of 12) and has therefore more restricted use than VY-. The ten submeanings where the two prefixes overlap are

¹³⁴ The only exception is the verb *vypendrivat’sja* which, unlike other verbs of this type, has a perfective counterpart *vypendrit’sja* ‘show off once’.

marked with color shading. This high degree of semantic overlap suggests that the two prefixes are very close to each other in terms of meaning.

The spatial prototypical image schema that motivates both concrete and metaphorical uses of both prefixes is labelled 1.OUT OF A CONTAINER: the Trajector moves out of a container, which is a bounded three-dimensional space, to a less bounded space. Interestingly, the observer can be located at the point of the Trajector's departure or arrival: compare *vyjti iz komnaty* 'leave the room' and *vyjti na scenu* 'come out on stage' (Nesset et al. 2011: 10). Similarly, for IZ- this is also possible: some verbs like *izgnat* 'exile, banish' and *istorgnut* 'extract' arguably focus on the object's removal from the presence of the observer, whereas other verbs like *izbrat* 'choose, select' and *izlovit* 'catch out' imply that the object becomes closer, or more available, to the observer at the end of the trajectory.

Although submeaning 1.OUT OF A CONTAINER is attested for both prefixes, there is a strong difference in terms of type frequency of lexemes. Under this meaning, VY-forms 122 verbs, whereas IZ- is attested only in 9 lexemes, some of which are rather obsolete like *istočit* 'effuse', *izrygnut* 'belch out', *izvérgnut'sja* 'push out', and *ispit* 'have a drink of'.

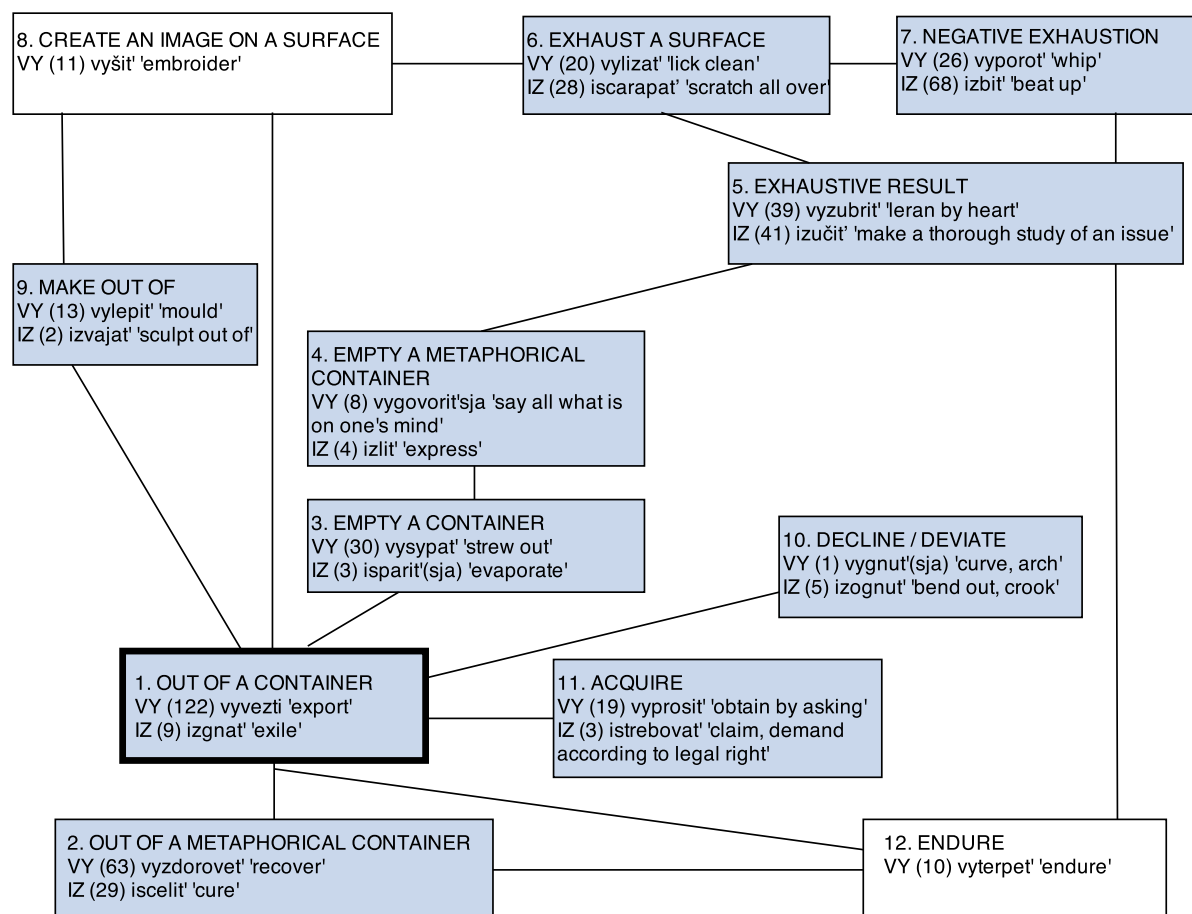


Figure 1: Radial category model for the prefixes VY- and IZ-: Semantic overlap in standard verbs.

For VY- this meaning is manifested in all verbs of determined motion like *vyplyt* 'swim out' and *vyletet* 'fly out', as well as verbs of concrete physical impact like *vypnut*

'kick out' and *vyputat'* 'disentangle', and factitive verbs like *vydvorit'* 'expel; kick away' (< *dvor* 'yard') and *vyvetrit'(sja)* 'air out; get rid of smell by airing the room' (< *veter* 'wind'). Apart from verbs already mentioned, IZ- in this meaning is observed in *izvleč'* and *iz"jat'*, both of which denote 'extract, take out of'.

Marginal verbs indicate that this concrete spatial use is still very relevant and productive for VY- in Modern Russian, as we can see in 51 novel coinages like *vyburit'* 'extort by drilling', *vylavirovat'* 'get out by maneuvering', *vydymit'* 'fume away', and *vymorgnut'* 'remove from one's eye by blinking':

- (1) *Mama smotrela na Mixas'ku, ej v glaz popala kakaja-to erundovina, i mama morgala-morgala, čtoby **vymorgnut'** ètu erundovinu.* [A. Lixanov. *Čistye kamuški* (1967)]
 'Mom was looking at Mixas'ka, she got some rubbish in her eye, and she was blinking and blinking, in order to blink this rubbish out.'

By contrast, the 9 marginal verbs in IZ- with the meaning 1.OUT OF A CONTAINER are not newly formed coinages but rather obsolete relics which are no longer in active use: *iznesti* 'carry out', *izrinut'* 'extract', *izletet'* 'fly out', *izdobyt'* 'obtain out of', *isxítit'* 'steal', and *iz(z)ut'* 'take off shoes'.

Submeaning 2. OUT OF A METAPHORICAL CONTAINER maps the spatial image schema of the prototype to the domain of states and situations that are conceptualized as metaphorical containers. This submeaning is frequently attested for both VY- and IZ-, with 63 and 29 standard lexemes accordingly. Many of metaphorical applications are available for both prefixes: compare "moving out" of the state of illness in *vyzdorovet'* 'recover' and *izlečit'* 'cure'; creating a new idea out of one's mind in *vydumat'* 'make up' and *izobresti* 'invent'; articulating or uttering a word in *vygovorit'* 'speak' and *izreč'* 'utter'; "stepping out" of the norm or standard behavior in *vytvorit'* 'act unexpectedly' and *izumit'* 'surprise'; spotting out an object in *vysledit'* 'track down' and *izyskat'* 'find'; causing something to emerge in a mature state in *vyrastit'* 'cultivate' and *vyňjančit'* 'bring up by dandling' and causing a more general change in *izmenit'* 'change'; taking out of an uncomfortable "canned" state of affairs in *vysvobodit'* 'befree', *vyručít'* 'rescue', and *izbavit'* 'liberate'. These examples suggest that the use of the two prefixes in the domain of states as metaphorical containers is very parallel and coherent. However, most of such parallel variants in VY- and IZ- differ in register or collocations. Register oppositions mostly contrapose the neutral VY- vs. the elevated IZ- (*vylečit'* vs. *izlečit'* 'cure') and colloquial or sub-standard VY- vs. standard IZ- (*vypravit'* vs. *ispravit'* 'correct'). A good example of difference in collocations comes from the verb *vykupit'* 'buy out; buy back', which describes commercial transaction, as opposed to its counterpart *iskupit'* 'expiate', which represents the same 'OUT-buy' but refers to abstract notions of guilt, sin, and mistake. Marginal verbs of this category 2. OUT OF A METAPHORICAL CONTAINER show higher productivity of VY- (38 coinages) as opposed to IZ- (only 6 coinages).

Submeanings 3.EMPTY A CONTAINER and 4.EMPTY A METAPHORICAL CONTAINER involve the same notion of a spatial or metaphorical container but additionally imply that the entire content of the container is removed, so that the container is fully emptied. Submeaning 3.EMPTY A CONTAINER of the prefix is realized in combinations with particular bases, most of which refer to procedures involving dry loose substances or

liquids. IZ- contributes this meaning in 3 verbs that refer to fluids: *izlit'sja* 'stream out', *isteč* 'bleed to death', and *isparit'(sja)* 'evaporate'. These verbs tend to develop figurative uses that refer to the flow of time (*isteč* 'expire (of time)'), emotions (*izlit'(sja)* 'express e.g. anger'), ideas (*izložít* 'lay out, present'), or possibilities (*isčerpát'(sja)* 'run out of'), where IZ- has submeaning 4.EMPTY A METAPHORICAL CONTAINER. Similarly, VY- refers to 3.EMPTY A CONTAINER in 30 verbs that describe evaporating, pouring, drinking, strewing, digging, and squeezing the entire content out of a physical container: *vykipet* 'boil away', *vyplesnut'(sja)* 'splash out'; *vypít'(sja)* 'drink up', *vylakat* 'lap all up', *vysypat* 'strew out', *vyvalít'(sja)* 'throw out', *vyryt* 'dig up', *vyžat* 'press out'. The same metaphorical mapping from 3.EMPTY A CONTAINER to 4.EMPTY A METAPHORICAL CONTAINER is available for VY-: the verb *vyplesnut'(sja)* can refer to splashing out liquids and expressing strong emotions. Metaphorical emptying is usually manifested by VY- in reflexive verbs that describe speech acts like *vyrugat'sja* 'swear', *vygovorit'sja* 'say all that is on one's mind', *vykričat'sja* 'empty oneself by shouting' or other exhaustive expressions of emotions as in *vyplakat'(sja)* 'cry using up all tears'. In these verbs with VY-, parallel to verbs with IZ-, the emptied metaphorical "container" is a human mind or a human body. Again, we see a strong parallelism in the semantics of the two prefixes. However, for VY- this kind of use is frequent and productive, whereas IZ- is attested in these meanings rarely (30 verbs vs. 3 verbs that employ VY- and IZ- of 3.EMPTY A CONTAINER respectively and 8 vs. 4 verbs in VY- and IZ- of 4.EMPTY A METAPHORICAL CONTAINER).

I suggest that these two submeanings that refer to exhaustive emptying a container form an intermediate node that connects the prototype to more abstract uses of VY- and IZ-. These abstract uses imply an exhaustive fulfillment of an activity and form a cluster of related submeanings 5.EXHAUSTIVE RESULT, 6.EXHAUST A SURFACE, and 7.NEGATIVE EXHAUSTION.

The most general sense 5.EXHAUSTIVE RESULT is attested for both prefixes in question. Here we observe activities that have a neutral or positive effect on an object: *vyzret* 'ripen', *vymesit* 'mix (a dough)', *vyučít* 'learn', *vyspat'sja* 'get a good night's sleep' in VY- and *ispeč* 'bake', *izmerít* 'measure', *iskupat* 'give a bath', *ispytat* 'experience' in IZ-. In standard lexemes, this use is equally frequent for both VY- and IZ- (40 and 41 verbs accordingly). However, the numbers of marginal coinages suggest that IZ- is much more productive than VY- in this submeaning. There are 85 marginal verbs in IZ- like *isšarit* 'search all over the place' and *isščelkat* 'use up all film on snapshots' and 31 marginal verbs in VY- like *vyšerstít* 'search intensely all over the place' and *vydyšat* 'use up all air by breathing'. Many marginal verbs in IZ- are coined on the basis of a highly productive pattern which combines this prefix with the reflexive marker *-sja* in a single construction: *izlenít'sja* 'become very lazy', *isxalturit'sja* 'do very careless work', *izmečtat'sja* 'become exhausted by dreaming a lot', *iznapominat'sja* 'exhaust oneself by reminding', *iznyt'sja* 'wear out oneself by whimpering and nagging', *izorat'sja* 'exhaust oneself by shouting', etc.

Submeaning 6.EXHAUST A SURFACE is manifested by the prefix VY- in the group of 20 standard verbs and by the prefix IZ- in 28 standard verbs. Again, marginal coinages show higher productivity of IZ-: 30 verbs in IZ- vs. 10 verbs in VY-. The verbs in VY- express various kinds of exhaustive treatment of the surface like *vybelít* 'bleach', *vydubit* 'tan a skin', *vylizat* 'lick clean', *vymostít* 'pave'. The verbs in IZ- more frequently refer to destructive impact as in *izborozdit* 'enridge' and *istoptat* 'trample

all over' but are not limited to such evaluative semantics, as we see in the neutral verb *isxodit'* 'walk all over the place'. And conversely, verbs in VY- can also refer to negative impact on the surface, as evidenced by the parallel synonymous variants *vypačkat'* – *ispačkat'* 'stain all over' and *vymazat'* – *izmazat'* 'smear all over'.

Submeaning 7.NEGATIVE EXHAUSTION is assigned to the prefix in those predicates that refer to activities with a negative impact. This submeaning stands out as more productive for IZ- in both standard and marginal verbs: IZ forms 68 standard and 61 marginal lexemes as opposed to VY- that is attested in 25 standard and 10 marginal verbs. Predicates in VY- often imply affecting an entire group of objects, as in *vymorit'* 'exterminate', *vyrezat'* 'massacre', *vyrodit'sja* 'die out, undergo degradation'. Verbs in IZ- include many lexemes that denote physical impact like *izranit'* 'wound all over', *iskusat'* 'bite all over', *isklevat'* 'peck a lot'. We can find synonymous variants like *ispepelit'* 'incinerate' and *vyžeč'* 'burn down', *izrugat'* and *vybranit'* 'scold', *isxlestat'* and *vyporot'* 'whip exhaustively', *izdergat'* 'exhaust by pulling or bothering' and *vymotat'* 'wear out by shaking'. However, the two prefixes imply different kinds of negative exhaustion. VY- implicates that an object is forced out of its unharmed well-being into a state of affairs where it is damaged or destroyed: compare *vymotat'* 'force out of a healthy state', *vyžeč'* 'empty the entire space by burning it down', *vybranit'* 'scold'. Destructive activity denoted by verbs in VY- can be applied to a single object, as in the verbs listed above, or to a whole group of homogeneous objects taken as a set, as in verbs like *vymorit'* 'exterminate', *vyrezat'* 'massacre', *vyrodit'sja* 'die out, undergo degradation', and *vyrubit'* 'cut down a forest'. By contrast, IZ- implies a different kind of exhaustiveness. IZ- denotes an exhaustive result of a destructive activity because it multiplies not objects affected by activity but rather completed acts, or portions, of activity. When IZ- attaches to verbs of physical impact like *lomat* 'break', *ranit'* 'wound', *kusat'* 'bite', *klevat'* 'peck', *dergat'* 'pull', it contributes the sense that activity described by the base is applied many times, affecting different spots of the same object: *izlomat'* means 'break in many places, all over', *izranit'* 'wound in many places', etc. At the same time, IZ- forms verbs of more general semantics like *isportit'*, *ispakostit'*, *isporabit'*, *ispoganit'* meaning 'spoil' as well as factitive verbs *izurodovat'* 'disfigure', *iskalečit'* 'cripple', and *izničtožit'* 'turn into nothing'. In these verbs the semantic contribution of IZ- resembles that of VY-: an object "moves out" of the state of welfare into a new disastrous state.

Classification of verbs within submeanings 5-7 of exhaustion is rather relative, because these submeanings form a continuum and are not mutually exclusive. All verbs classified for these submeanings could be alternatively assigned to the group of 5.EXHAUSTIVE RESULT, and many verbs simultaneously refer to 6.EXHAUST A SURFACE and 7.NEGATIVE EXHAUSTION. However, I apply this subdivision and distinguish between different kinds of exhaustion in order to provide a more granular comparison of the two prefixes. One of the outcomes of this analysis is that IZ- is indeed more associated with activities that carry out a negative destructive impact on the object than VY-. This semantic difference between VY- and IZ- becomes evident even in some pairs of parallel verbs. Compare, for example, *vysušit'* and *issušit'* both formed from the simplex *sušit'* 'dry'. The former variant 'VY-dry' carries a positive connotation and denotes 'dry (enough)', whereas its counterpart 'IZ-dry' has a negative connotation and means 'dry (more than enough), damage by drying'.

Submeaning 8.CREATE AN IMAGE ON A SURFACE is attested only for VY-. This use is directly related to the prototype 1.OUT OF A CONTAINER: like an object that comes out of a container (e.g. *vyjti na scenu* ‘come out on a stage’), an image appears on the surface (e.g. *vyšit’ uzor na skaterti* ‘embroider a pattern on a cloth’) and thus becomes visible and available for perception. We observe this use of the prefix in eleven verbs like *vyčertit’* ‘draw an image’, *vytkat’* ‘weave an image’, *vygravirovat’* ‘engrave’, and *vytatuirovat’* ‘tattoo an image’. Unlike VY-, IZ- is not attested in this meaning in standard or in marginal verbs.

Related to the image-on-surface use, submeaning 9.MAKE OUT OF implies emergence of a three-dimensional object out of a substance: *vyplavit’* ‘smelt’, *vykovat’* ‘forge’, *vyrezat’* ‘carve out’. Overall, there are 18 standard verbs with VY- and 2 standard verbs in IZ-. Among marginal coinages, there are 9 verbs in VY- and only 3 verbs in IZ-. Apart from the difference in frequency, we again witness semantic comparability of the two prefixes in pairs like *vylepit’* ‘mould’ and *izvajat’* ‘sculpt’ and more general terms like *vyrabotat’* ‘work out, manufacture’ and *izgotovit’* ‘produce’.

Unlike submeanings 8 and 9, submeaning 10.DECLINE/DEVIATE is very frequent and productive for IZ- and very rare for VY-. In standard verbs, it is manifested in 2 verbs with VY- and 11 verbs with IZ-: *vygnut’* ‘curve, arch’, *izognut’* ‘bend out, crook’, *izvit’* ‘wind, twist’, *iskrivit’* ‘bend, distort’, *izvratit’* ‘distort’. The initial natural unaffected shape of an object is conceptualized here as a metaphorical container which is abandoned as a result of deformation and distortion. Among marginal verbs, there are 8 verbs in IZ- and only one verb in VY- that demonstrate this use.

Submeaning 11.ACQUIRE refers to obtaining an object out of someone’s domain of possession. This use is frequent for VY- and is attested in 19 standard and 12 marginal verbs that denote different kinds of behavior from neutral asking (*vyprosit’*) and tiresome begging (*vykljančit’*) to obtaining by praying (*vymolit’*), doing a good service (*vyslužit’*), negotiating the price (*vytorgovat’*), winning (*vyigrat’*), suing in court (*vysudit’*), or direct forcing (*vynudit’*). IZ-, on the contrary, is attested in this meaning only in 3 standard and one marginal lexeme that are obsolete terms of jurisdiction: *istrebovat’* ‘demand according to legal right’, *isprosit’* ‘acquire through asking’, *isxlopotat’* and *isxodatajstvovat’* both meaning ‘obtain through efforts’.

The final submeaning 12.ENDURE can be expressed only by the prefix VY- and is manifested in 10 standard and 2 marginal verbs like *vyderžat’* ‘stand (up to), endure’, *vystojat’* ‘withstand; stand one’s ground’, *vyterpet’* ‘bear, endure’, *vyžit’* ‘survive’ and others. In these predicates, the notion of a container is applied to difficult circumstances that last during a period of time. A person who endures this state of affairs “comes out” of it as out of a metaphorical “container”.

Summing up, a very granular classification and comparison of senses carried by the two prefixes shows that VY- and IZ- are very similar semantically. Their polysemies can be modelled as a single radial network of twelve meanings, where IZ- inhabits a subset of ten meanings. Therefore, the zone of semantic overlap yields 83% of all meanings that corresponds to 545 (96%) of individual standard perfective verbs prefixed in VY- and IZ-. Moreover, both prefixes share the prototypical spatial image schema 1.OUT OF A CONTAINER, which motivates all concrete spatial and abstract non-spatial uses.

This analysis tested the findings reported in Nessel et al. 2011 against much larger dataset. First, the present study made it possible to refine the distribution of VY-

and IZ- in standard verbs. Second, we gain much information about the use of VY- and IZ- in marginal verbs including novel coinages and obsolete relics. Third, scaling down the token frequency threshold revealed an overlap of the two prefixes in one additional submeaning 3.EMPTY A CONTAINER.

8.4.2 Profiles within polysemy: Radial Category Profiling

The key outcome of the previous section is the striking finding of how similar and parallel the two prefixes are in their use: they refer to the same spatial image schema, express mostly the same submeanings, and develop the same abstract senses. Yet, they differ in terms of how frequently these shared submeanings are manifested in individual lexemes. In this section I provide a quantitative analysis of the distribution of verbs across prefixal subcategories proposed in the previous section. In doing so I will explore whether VY- and IZ- have distinct profiles within the semantic network that they share.

Similar to previous chapters, I adopt the methodology of Nessel et al. 2011 in assessing Radial Category Profiles of the prefixes in question. Table 3 contains the raw numbers of individual standard and marginal perfective verbs where the twelve submeanings of VY- and ten submeanings of IZ- are manifested.

Submeaning of the prefix	standard verbs in VY-	standard verbs in IZ-	marginal verbs in VY-	marginal verbs in IZ-
1.OUT OF A CONTAINER	122	9	51	9
2.OUT OF A METAPHORICAL CONTAINER	63	29	38	6
3.EMPTY A CONTAINER	30	3	9	0
4.EMPTY A METAPHORICAL CONTAINER	8	4	4	1
5.EXHAUSTIVE RESULT	40	41	31	85
6.EXHAUST A SURFACE	20	28	10	30
7.NEGATIVE EXHAUSTION	25	68	10	61
8.CREATE AN IMAGE ON A SURFACE	11	0	2	0
9.MAKE OUT OF	18	2	9	3
10.DECLINE/DEVIATE	2	11	1	8
11.ACQUIRE	19	3	12	1
12.ENDURE	10	0	2	0
Total number of verbs:	368	198	179	204

Table 3: Distribution of perfective verbs in VY- and IZ- across prefixal submeanings.

In order to test the distributional difference between VY- and IZ- statistically, I focus only on standard verbs and conflate submeanings into three groups – OUT OF, EXHAUST, and OTHER. The group OUT OF brings together six submeanings that are most closely related to the spatial image schema – 1.OUT OF A CONTAINER, 2.OUT OF A METAPHORICAL CONTAINER, 3.EMPTY A CONTAINER, 4.EMPTY A METAPHORICAL CONTAINER, 8.CREATE AN IMAGE ON A SURFACE, and 9.MAKE OUT OF. Second, I group together submeanings 5.EXHAUSTIVE RESULT, 6.EXHAUST A SURFACE, and 7.NEGATIVE EXHAUSTION, because they are more abstract

and remote from the spatial prototype. Finally, OTHER conflates the remaining submeanings 10.DECLINE/DEVIATE¹³⁵, 11.ACQUIRE, and 12.ENDURE.

Note that the distinction between the first two groups is supported syntactically by two different constructions. Spatial submeanings included in the first group are usually realized in verbs that combine with the preposition *iz* 'out' followed by a noun that names a container. Compare *vyjti iz komnaty* 'walk out of the room', *vysypat' pesok iz meška* 'strew the sand out of the sack', *vylepit' iz gliny* 'mold out of clay'. By contrast, abstract prefixal meanings of the second group EXHAUST are manifested in verbs that take a direct object (*izmerit' dlinu* 'measure the length', *vymazat' lico* 'smear the face') or are intransitive (*vyzret'* 'ripen'). This opposition in terms of syntactic constructions is a tendency observed in data. The system is flexible and many verbs can appear in both constructions depending on their reading – OUT OF or EXHAUST. Examples below illustrate the use of the verb *vymyt'(sja)* which denotes 'remove by washing' with VY-'1.OUT OF A CONTAINER' in (2) and means 'wash until it is clean' with VY-'6.EXHAUST A SURFACE' in (3):

- (2) *Mama raspuskaet mne kosički, čtoby **vymyt'** sol' iz volos.* [M. Šiškin. Pis'movnik. (2009)] 'Mom takes my hair out of braids in order to **wash** the salt **out** of it.'
- (3) *Pust' on lučše **vymoet** posudu, poka ty budeš' v magazine.* [«Daša», 2004] 'Let him rather **do the dishes** while you are in the store.'

In (2), the spatial reading of the prefix 1.OUT OF A CONTAINER is supported by the prepositional phrase *iz volos* 'out of the hair'. By contrast, the meaning 6.EXHAUST A SURFACE of the prefix in (3) is manifested in the verb which takes a direct object: *vymyt' posudu* 'wash the dishes'.

Similarly, many other verbs also allow for two readings of the prefix, and the spatial submeaning is usually syntactically supported by the preposition *iz* 'out'. Compare *vymesti musor iz komnaty* 'sweep trash out of the room' (1.OUT OF A CONTAINER) vs. *vymesti komnatu* 'sweep the room clean' (6.EXHAUST A SURFACE); *vykurit' iz gnezda* 'smoke out of the nest' (1.OUT OF A CONTAINER) vs. *vykurit' sigaretu* 'smoke up a cigarette' (5.EXHAUSTIVE RESULT); *vymarat' slovo iz teksta* 'delete the word out of text' (2.OUT OF A METAPHORICAL CONTAINER) vs. *vymarat' vse platje* 'stain the dress all over' (6.EXHAUST A SURFACE).

Table 4 and Diagrams 1 and 2 show relative frequency distributions of the two prefixes across semantic subgroups of meanings in standard and marginal verbs. These are radial category profiles of VY- and IZ-. We can see that the two prefixes have different centers of gravity: VY- is more frequently attested in spatial submeanings of the group 'OUT OF', whereas IZ- is very prominent in abstract submeanings of the group 'EXHAUST'.

¹³⁵ IZ- is more entrenched in this submeaning than VY-, which is an important fact in itself. I will ignore it for the sake of the present analysis and come back to it in 8.5.

Semantic group	Standard verbs				Marginal verbs			
	VY-	IZ-	VY- %	IZ %	VY-	IZ-	VY- %	IZ %
'OUT OF'	252	47	69%	24%	113	19	63%	9%
'EXHAUST'	85	137	23%	69%	51	176	29%	86%
'OTHER'	31	14	8%	7%	15	9	8%	5%
Total:	368	198	100%	100%	179	204	100%	100%

Table 4: Distribution of VY- and IZ- across semantic subgroups of submeanings.

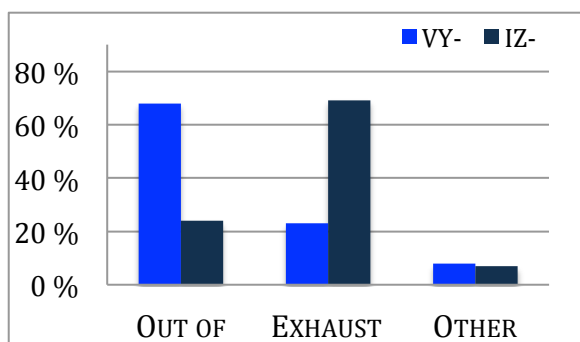


Diagram 1: Radial Category Profiles of VY- and IZ- (%) in standard verbs.

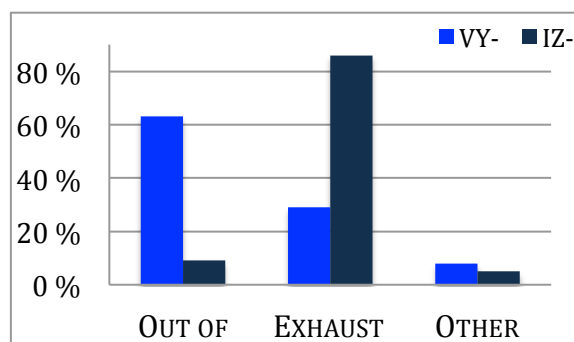


Diagram 2: Radial Category Profiles of VY- and IZ- (%) in marginal verbs.

Interestingly, in standard verbs, the two centers of gravity accommodate nearly 70% of uses for each of the two prefixes, whereas around 30% are found in remaining uses. Data on marginal verbs gives additional evidence for distributional difference between VY- and IZ-. Moreover, it shows that spatial uses of IZ- comprise only 9% of marginal verbs which is a much smaller rate than in standard verbs. Most of these verbs with low token frequency are obsolete remnants of Slavonic borrowings. Abstract uses of the type EXHAUST, on the contrary, show an even higher rate of productivity in marginal verbs prefixed in IZ- (86%) than in standard verbs in IZ- (69%). They have colloquial flavor and apparently correspond to the “other” IZ-, native to Russian (see section 8.5.3 for my account of this type of IZ-).

A chi-square test shows that the distributions of standard verbs in VY- and IZ- across prefixal submeanings are significantly different: X-squared = 119, df = 2, p-value < 2.2e-16, and the effect size is very large: Cramer’s V=0.46. The same statistical analysis of the distributions of marginal verbs indicates a highly significant difference (X-squared = 136, df = 2, p-value < 2.2e-16) with an even larger effect size: Cramer’s V=0.6.

Summing up, by means of Radial Category Profiling we measured the degree of semantic similarity and difference of the two prefixes in terms of type frequency of verbs attested for each submeaning. This analysis demonstrates that the two prefixes have distinct profiles within a shared network of polysemy. They are strongly associated with different parts of this radial network, and this semantic specialization coexists with a large number of parallel uses of VY- and IZ- within the same submeanings.

8.5 Distribution: Overlap and Contrastive uses

Now that we know that the overlap in the distribution of VY- and IZ- is extremely large, with 39% of lexemes in the database that have an alternative counterpart with the other prefix, a number of questions remain unsolved regarding the distribution of VY- and IZ-. Do minimal pairs of verbs show contrastive or free distribution of the prefixes? If both, what predominates? What factors condition the choice between VY- and IZ- in parallel verbs? Addressing these questions, in this section I focus on cases when a simplex verb can attach either of the two prefixes or only one of them. I will propose an account for both types of cases.

8.5.1 Verbs that can attach both prefixes

It is common knowledge that many pairs of parallel verbs in VY- and IZ- feature the same semantic contribution of the two prefixes with a slight difference in register. When contrasted, VY- gives a concrete spatial meaning, while verbs in IZ- refer to abstract notions and carry an elevated and bookish flavour. Compare *vylit' vodu* 'pour out water' vs. *izlit' dušu* 'express emotions, lit. pour out one's soul' (< *lit* 'pour'); *vygnat' iz doma na ulicu* 'kick out of home onto the street' vs. *izgnat' iz strany* 'exile, deport from the country' (< *gnat* 'chase'); *vybrat' knigu* 'choose a book' vs. *izbrat' prezidenta* 'elect a president' (< *brat* 'take'), *vybežat' iz doma* 'run outside' vs. *izbežat' vojny* 'avoid war' (< *bežat* 'run'), *vyključit' iz rozetki* 'remove a power plug from a socket' vs. *isključit' iz universiteta* 'expel from university'. Similarly, *vyžit'* 'VY-live' refers to forcing people to move out, whereas *izžit'* 'IZ-live' implies getting rid of traditions, stereotypes, etc. This opposition in terms of register and collocations is well explained in terms of Slavonic origins of the loan IZ- as opposed to the native prefix VY-. I conclude that in these minimal pairs of verbs both prefixes refer to roughly the same meaning OUT OF A [METAPHORICAL] CONTAINER (or the closely related meaning EMPTY A [METAPHORICAL] CONTAINER), therefore their use is semantically not contrastive. The choice between VY- and IZ- is governed by register and established types of uses and collocations, many of which, especially in case of IZ-, are idiomatic.

Another type of opposition between the two prefixes has been unnoticed in the literature. In many pairs of parallel verbs we can observe semantic contrast of VY- and IZ-. When these prefixes attach to the same simplex verb (usually a verb that denotes physical impact), VY- contributes the meaning 1-2.OUT OF A [METAPHORICAL] CONTAINER and refers to a single action, whereas IZ- provides the abstract meaning 5-7.EXHAUST and implies multiplication of an activity. Note that here we deal with the "other" IZ-, etymologically native to Russian, frequent and productive in colloquial speech and not carrying any bookish elevated flavor. Examples of such parallel contrastive pairs are numerous:

- (4) *vylomat'* 'break out' vs. *izlomat'* 'break many objects or one object in many places'
vybit 'beat out of' vs. *izbit'* 'beat up'
vyrezat' 'cut or carve out of' vs. *izrezat'* 'cut in different directions or places'
vyrvat' 'tear out' vs. *izorvat'* 'tear all through into many small pieces'
vydrat' 'flog out' vs. *izodrat'* 'flog all through into many small pieces'

vykopat' 'dig out' vs. *iskopat'* (e.g. *vsje pole*) 'dig all over the place'
vytoptat' 'trample grass off the place' vs. *istoptat'* 'trample all over the place'
vyčesat' 'comb out' vs. *isčasat'(sja)* 'scratch everywhere'
vyest' 'eat out, damage by corrosion' vs. *izjest'* 'damage an object in many places'
vypisat' 'copy out by writing' vs. *ispisat'* (*list* or *ručku*) 'use up all the space or ink while writing'
vyvedat' 'worm a secret out of someone' vs. *izvedat'* 'experience many things'
vyxodit' 'nurse someone out of illness' vs. *isxodit'* 'walk many trails or paths'

In all these verbs we can identify the following semantic opposition between VY- and IZ-. VY- typically refers to an event which consists of one single portion of activity, as we see in the verb *vyломat'* 'break out' formed from the base *lomat'* 'break' and VY- in the meaning 1.OUT OF A CONTAINER. By contrast, IZ- implies that an event is comprised of several acts, or portions, of activity (quantatively) or that the activity is extremely intense (qualitatively), as we see in the verb *izlomat'* 'break many objects or one object in many places'. This verb is formed from the same base *lomat'* 'break' and IZ- 7.NEGATIVE EXHAUSTION. In this light, I suggest that the meaning EXHAUST is a semantic effect which results from multiplication of an action that the simplex base refers to. Multiplication of the action is typically encoded in the prefix IZ-, and not in VY-. As a result, the verbs prefixed in IZ- often have a more expressive, or evocative, flavor and refer to more intensive activities than those described by parallel verbs prefixed in VY-. In this sense, IZ- is clearly a marked member in this pair of prefixes, both in terms of register and in terms of semantics.

Note, however, that in some cases the verb prefixed in VY- can have both spatial and exhaustive readings, and the parallel variant in IZ- does not exist. For example, *vypolot'* (*sornjak*) denotes 'pull out one weed', while *vypolot'* (*grjadku*) means 'pull out all weeds in the garden-bed'. Likewise, the verb *vytravit'* can refer to both 'chase someone out' and 'exterminate, poison all in the group'. Another good example comes from the verb *vyudit'* which can imply catching one fish or catching all the fish. There exist no counterparts in IZ- like **ispolot'*, **istravit'*, **izudit'*, probably because the meaning EXHAUST is already expressed by the verb in VY-. However, I argue that even when the prefix VY- has an exhaustive reading, it does not encode multiplication of an activity. Instead, VY- applies the activity named by the base to a more complex trajector, not a single object (e.g. one weed) but a set of homogeneous or similar objects (e.g. all weeds in a garden-bed).

This semantic contrast between VY- and IZ- is not an absolute rule but rather a tendency. In some cases, this contrast seems neutralized. In particular, in some pairs of parallel verbs, both prefixal counterparts can express exhaustive result of an activity, as illustrated in (5):

- (5) *vykupat'sja* – *iskupat'sja* 'have a swim'
vyučit' – *izučit'* 'learn a subject completely, make a thorough study of an issue'
vymoknut' – *izmoknut'* 'become soaking wet'
vyteret' – *isteret'* 'wear clothes threadbare'
vymotat' – *izmotat'* 'exhaust'
vyzolotit' – *izzolotit'* 'gild, cover with gold'
vymazat'(sja) – *izmazat'(sja)* 'smear up all over'

vypačkat'(sja) – ispačkat'(sja) ‘make dirty all over’

I hypothesize that the number of parallel uses of VY- and IZ- that show semantic contrast of the type illustrated in (4) is larger than the number of non-contrastive uses which differ in terms of register. This hypothesis is based on the fact that the use of IZ- in submeanings 5-7.EXHAUST has a much higher type frequency in standard verbs and productivity in marginal verbs, than the use of IZ- in submeanings 1-2.OUT OF A [METAPHORICAL] CONTAINER, as shown in 8.4.2. However, I leave calculating the proportions of semantic overlap and semantic contrast for future research.

8.5.2 Verbs that attach only one prefix

Many verbs with concrete spatial meaning prefixed in VY- with the sense 1.OUT OF A CONTAINER cannot attach IZ- instead: for example, *vydelit'* ‘single out’ – **izdelit'*, *vysledit'* ‘track down’ – **issledit'*, *vykrast'* ‘steal out of’ – **iskrast'*, *vyplatit'* ‘pay out, pay off’ – **isplatit'*, *vyselit'* ‘evict, force to move out’ – **isselit'*, *vyprovodit'* ‘show the door’ – **isprovodit'*, *vyperet'* ‘push out’ – **isperet'*, *vymanit'* ‘swindle out of’ – **izmanit'*, and others. Among such verbs there are many Natural Perfectives, where the meaning of the prefix and the semantics of the verbal base overlap (Nesset et al. 2011; Endresen et al. 2012: 266-272): *vydut'* ‘blow out of; produce by strong stream of air’ – **izdut'*, *vykristalizovat'sja* ‘crystallize’ – **iskristalizovat'sja*, *vykroit'* ‘cut out’ – **iskroit'*, *vylučit'* ‘chuck out’ – **izlučit'*, *vypučit'* ‘google (eyes)’ – **ispučit'*, *vytjanut'* ‘pull out’ – **istjanut'*, *vyrulit'* ‘drive out’ – **izrulit'*, *vyvalit'* ‘throw out’ – **izvalit'*, etc.

Among verbs that attach only one prefix (VY- or IZ-) there is a group of lexemes that refer to spatial motion. They deserve special attention in this discussion, because the two types of Russian verbs of motion differ dramatically in terms of which prefix of the two they prefer. Russian has a morphological distinction between two classes of verbs of motion: so called *unidirectional*¹³⁶ verbs like *plyt'* ‘swim or sail in one direction’ and *multidirectional* (or *non-directional*) verbs like *plavat'* ‘swim or sail in different directions or without a particular direction’.

Verbs of unidirectional motion like *plyt'* ‘swim’ almost exclusively attach VY-, whereas parallel derivatives in IZ- do not exist: *vyplyt'* – **isplyt'* ‘swim out’, *vyletet'* – **izletet'* ‘fly out’, *vynesti* – **iznesti* ‘carry out’, *vypolzti* – **ispolzti* ‘crawl out’, *vytaščit'* – **istaščit'* ‘drag out’. Exceptions are few: *vybežat'* – *izbežat'* from *bežat'* ‘run’, *vyjti* – *izojti* from *idti* ‘walk’, and *vygnat'* – *izgnat'* from *gnat'* ‘chase’. However, the verbs with IZ- in these pairs no longer refer to the spatial motion designated by their verbal bases. The verb *izbežat'* ‘IZ-run’ is not related to the physical activity of running and means ‘avoid an unwanted or unpleasant situation like a difficulty or problem’. The verb *izojti* ‘IZ-go’ has a highly restricted figurative use and is no longer related to the physical motion of walking: e.g. *izojti slezami* ‘cry one’s heart out’. The verb *izgnat'* ‘IZ-chase’ refers to exile, banishment, and loss of citizenship rather than to physically chasing someone out of a place (Nesset et. al 2011: 13). The use of IZ- in these verbs comes from the loan

¹³⁶ Terms *unidirectional/multidirectional* are used in Wade 1992; parallel terms *determinate/indeterminate* are employed in Forsyth 1970. For a detailed discussion of terminology see Nesset 2000, where the author argues for the terms *unidirectional/non-directional*.

Church Slavonic origin of this prefix. These verbs are remnants of a former larger group that faded away (lexemes like *izletet* 'fly out' went out of use).

On the contrary, verbs that denote multidirectional spatial motion like *plavat* 'swim, sail' normally attach the prefix IZ-, but not VY-. All such derivatives show the same type of meaning: *izplavat vse reki / ves' okean vdol' i poperek* 'swim or sail all over in all rivers; swim all over the ocean', *iz"ezdit vsju stranu* 'travel all over the country', *izletat vsju Indiju vdol' i poperek* 'fly all over India', *isxodit vse pole* 'walk all over the field', *iz"ezdit vsju stranu* 'travel all over the country'. Theoretically, one can easily form similar verbs like *ispolzat* e.g. *ves' pol* 'crawl all over the floor' from the multidirectional base verb *polzat* 'crawl'. In all these derivatives the prefix IZ- contributes the submeaning 6.EXHAUST A SURFACE and turns an intransitive base into a transitive derivative. The typical context of this use is illustrated in (6):

- (6) *Za tridcat' let, čto prorabotal Moxovec na Severe, on isxodil, izlazil, isplaval ego vdol' i poperek.* [V. Janelis. Gorjačij sever // «Texnika – molodeži», 1982]
'During the thirty years that Moxovec worked in the North he **has walked, climbed, and sailed it all over** in every direction.'

The same use of IZ- in submeaning 6.EXHAUST A SURFACE we observe in the verbs *izvozit* 'make dirty all over' and *iznosit* 'wear out clothes by means of intensive or frequent use' formed from the multidirectional transitive bases *vozit* 'drive, transport' and *nosit* 'carry, wear'.

The prefix VY- does not attach to such multidirectional motion verbs like *plavat* 'swim', *letat* 'fly', *begat* 'run', *polzat* 'crawl', *vodit* 'lead', *gonjat* 'chase'. There are only three multidirectional motion verbs that combine with VY-: *vyxodit bol'nogo* 'nurse someone out of the state of being ill', *vyezdit konja* 'ride a horse until it will easily follow commands', and *vynosit rebenka* 'complete the pregnancy successfully'. All three exceptions exhibit the same pattern: these verbs refer not to multidirectional movement in space but rather to getting out of a certain state – the state of being sick, being disobedient, being in the womb or being pregnant.

Why do verbs of multidirectional motion prefer IZ- and verbs of unidirectional motion attach VY-? I suggest that this is so because IZ- can encode multiplication of activity which is not possible for the prefix VY-. Note that the uses of exhaustive IZ- are not bookish but very colloquial and productive. These uses are associated with IZ- that is native to Russian. It is these uses that are semantically contrastive with VY-. Thus, I conclude that native IZ- is compatible with base verbs of multidirectional motion.

Summing up, the distribution of VY- and IZ- across classes of motion verbs, i.e. unidirectional and multidirectional bases, is near-complementary. Exceptions to this distributional pattern show the same pattern as we saw in the previous subsection:

- 1) semantic overlap of VY- with the loan Slavonic bookish IZ- (*vygnat* 'chase out of' vs. *izgnat* 'exile') and
- 2) semantic contrast between VY- and the native colloquial IZ- (*vyxodit* 'nurse "out" of illness' vs. *isxodit* 'walk all over the place').

On the basis of my observations I propose a possible explanation of these patterns of distribution in 8.5.3.

8.5.3 Possible explanation: two IZ- prefixes with different spatial image schemas

I propose that the cases of semantic overlap and semantic contrast in minimal pairs of verbs prefixed in VY- and IZ- can be explained in terms of the type of IZ-, or its historical origin. In cases where VY- and IZ- differ only in terms of register but show roughly the same spatial meaning, we are dealing with the IZ- that is a loan Slavonic prefix in Russian. In cases where we observe a semantic contrast of VY- and IZ-, we see the “other” IZ-, which is historically native to Russian.

The reason why these two IZ- prefixes have different semantic properties might be caused by the fact that their semantics is motivated by different prototypical image schemas. The loan IZ- has the same spatial image schema as VY- – 1.OUT OF A CONTAINER. The IZ- that is native to Russian might have a spatial image schema different from that of VY-. It is difficult to define though how it is different.

A helpful group of verbs in this regard includes those lexemes that feature submeaning 10.DECLINE/DEVIATE of the prefix. Recall that this submeaning is mostly represented by the prefix IZ- in verbs that describe distortion and deformation, both spatial, as in *izognut'(sja)* ‘bend out, crook’, *iskrivit'(sja)* ‘bend, distort’, *iskorežit'* ‘bend, warp’, *izvit'(sja)* ‘wind, twist’, and metaphorical, as in *iskoverkat'(sja)* ‘mispronounce or misspell’, *iskazit'(sja)* ‘misrepresent’ and *izvratit'(sja)* ‘pervert’. Reflexive verbs that belong here include *izvernút'sja* ‘turn oneself’, *izlovčit'sja* ‘twist oneself, contrive’, *isxítrít'sja* ‘manage; act cunningly’. This use of IZ- is productive and is attested in marginal verbs like *iskosobočit'sja* ‘twist oneself’ and *izzmeít'sja* ‘twist oneself like a snake’. We can find this meaning of IZ- also in nouns like *izgib* ‘curve’, *izlom* ‘crook, bend (of a river, road, eyebrow)’, *izvilina* ‘curvature, meander, brain gyrus’, and *izlučina* ‘winding bend of a river or a road’. All these words refer to the same distorted shape which consists of multiple bends or curves. This shape resembles a zigzag line found in nature in the shape of a snake, a river, or a road.

By contrast, the meaning 10.DECLINE/DEVIATE is attested for VY- in two verbs *vygnut'(sja)* ‘curve, arch’ and *vyprjamat'(sja)* ‘straighten’. However, the trajectory of deformation here is quite different – the shape of an arch, which implies only one curve.

I visualize the spatial image schemas of submeaning 10.DECLINE/DEVIATE for the prefixes VY- and IZ- in Figure 2. I suggest that VY- refers to the arch-shaped trajectory that consists of one curve bent outward (image to the left). The image to the right represents the trajectory of the prefix IZ- (the native one). This is a multicurved shape which resembles a zigzag line.

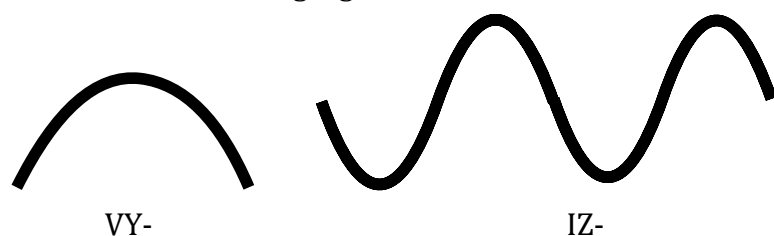


Figure 2: Image schemas of VY- and IZ- in submeaning 10.DECLINE/DEVIATE.

It is worth comparing two parallel verbs *vygnut'* ‘VY-bend’ and *izognut'* ‘IZ-bend’. This pair of spatial verbs sheds some light on what kind of spatial trajectories

VY- and IZ- imply when they combine with the base *gnut* ‘bend’. The verb *vygnut*’ refers to an arch-shaped posture bended outwards from the observer’s perspective (*vygnut*’ *spinu* ‘stretch out one’s back’, as opposed to *prognut*’ ‘cave in’), as shown in Figure 2.

By contrast, *izognut*’ can refer not only to arch-shaped objects but also to other shapes like the treble clef in music (cf. Nessel et al. 2011: 14), as evidenced by the example in (7):

- (7) *Vstrečalis’ nam dačniki s sobakami – s irlandskimi setterami ili borzými, izognutými, kak skripičnyj ključ.* [Ju. Koval’. *Kartofel’naja sobaka* (1972)]
 ‘We have met holiday visitors with dogs – Irish setters or borsois, **twisted like a treble clef.**’

Again, we see that even in this spatial use, the prefix VY- implies one curve, while IZ- can imply several curves. This also explains why the verb *izognut*’ ‘IZ-bend’ describes an activity with a higher degree of intensity than the verb in VY- *vygnut*’ ‘VY-bend’.

The prefix IZ- can also describe the shapes of a zigzag line, a climbing plant, or a spiral, as we see in the examples (8) and (9) of the verb *izvit*’(*sja*) ‘wind, twist’:

- (8) *Ja krug nego v’junom izovjus’, a <on> sidit da molčit* [G.]e. Nikolaeva. *Bitva v puti* (1959)]
 ‘I would **curl** around him like a **climbing plant**, but he would just remain sitting in silence.’
- (9) **Izvitye** formy bakterij – èto **spirali**. Naprimer, spirilly. [D. Makunin. *Mikroby. Ix imena.* // «Nauka i žizn’», 2008]
 ‘**Twisted** forms of bacteria are spirals. For example, spirillas.’

I suggest that the difference in the spatial image schemas illustrated in Figure 2 motivates semantic contrast of VY- and IZ- when they attach to the same base verbs. VY- is typically not compatible with multiplication of activity. Therefore, this prefix attaches to verbs of unidirectional motion. For native IZ-, multiplication is the key semantic property manifested in its productive uses. The exhaustive or intensive activity that this prefix normally refers to is a semantic effect of the multiplication of activity that IZ- implies. IZ- implies multiple curves of the trajectory in *izognut*’, and likewise, IZ- is semantically compatible with verbs of multidirectional motion. These verbs refer to motion exercised in several directions at each subportion of the motion. At the same time, IZ- is not compatible with verbs of unidirectional motion which denote one act of motion carried out in one particular direction.

Multiplication is a rather common phenomenon in the semantics of Russian prefixes. For example, we observe multiplication of activity in the use of the prefix PERE- in submeaning DISTRIBUTE (*perečitat’ vse knigi* ‘read all books’) and submeaning THOROUGH (*perepačkat’ vse plat’e* ‘stain the dress all over’). We find a similar effect in multiplication of landmarks in the use of the prefix O- in submeaning AFFECT A “CIRCLE” OF OBJECTS (*obzvonit’ vsech družej* ‘call all friends of one’s circle’).

To conclude, I propose that VY- is associated with a single portion of activity (single curve, single exit), whereas for IZ- it is possible to refer to a multiplied path, or

trajectory, subdivided into subparts, or subportions. We observe this difference between VY- and IZ- already in the domain of spatial relations – in verbs of motion and in the verbs *vygnut'* and *izognut'*. I further suggest that this idiosyncratic semantic property of the prefix IZ- to refer to multiple sub-trajectories serves to motivate its more abstract metaphoric uses of the type 5-7.EXHAUST.

8.5.4 A note on prosody and phonology

The two prefixes in question have very different prosodic properties. When VY- is attached to form a perfective verb, it always carries the stress, whereas IZ- is typically stressless. It is hard to tell whether these properties directly govern the choice between the two prefixes, especially in cases when semantic and register differences might be of greater importance. At the same time, the property of carrying the stress might restrict the number of verbal bases that the prefix VY- can attach to. In particular, the more syllables the base has, the more problematic it is to attach the prefix VY-. In derivatives like *vykristallizirovat'sja* 'crystallize', *vydifferencirovat'* 'extract by differentiating', or *vydressirovat'* 'train (animals)' VY- preserves this prosodic characteristic, but the verbs of such a length in VY- are in general rather rare.

Both VY- and IZ- contribute one syllable to the base. Exceptions concern those cases when IZ- is vocalized into IZO- in verbs like *izognut'* 'bend, twist' and thus contributes two syllables.

In addition, VY- and IZ- are different phonologically. VY- is a vowel-final prefix, whereas IZ- is consonant-final. IZ- undergoes allomorphic variation depending on the onset of the base: it has the default allomorph IZ-, the devoiced allomorph IS-, and the vocalized allomorph IZO-, similarly to the prefix RAZ- discussed in Chapter 3. VY-, on the contrary, always remains the same. It can attach to onsetless bases as well, as in *vyudit'* 'fish out'. It might be interesting to test the distribution of VY- and IZ- for phonological conditioning in terms of the initial phonemes of the base. Because suppletive allomorphs theoretically can be phonologically conditioned, one could test this possibility regarding VY- and IZ-. However, since VY- and IZ- are not phonologically related in Modern Russian, I suspect that phonology might not play a strong role in their distribution.

8.6 Discussion

How can we best capture the relationship between VY- and IZ- in Modern Russian? Can we describe it in terms of Non-Standard allomorphy, similar to what we saw in other pairs of Russian and Slavonic variants like S-/SO- and PERE-/PRE-? Or is there a stronger argument for classifying VY- and IZ- as distinct morphemes?

The relationship between the prefixes VY- and IZ- in Modern Russian is complex and ambivalent. To a large extent, this complexity is caused by the hybrid behavior of IZ-, the uses of which differ in terms of semantics, register, frequency, and productivity and correlate with two historical origins. The uses of IZ- that stem from the Slavonic origin are bookish, rather obsolete and unproductive, whereas those uses that originated from the strata native to Russian are colloquial, frequent, and productive. Semantically, the loan IZ- is very similar to VY- in submeanings 'OUT OF', and their relationship can be best described in terms of Non-Standard Allomorphy, where the

two prefixes are distributed by means of register and collocations. The native IZ- is semantically more distant from VY- in its ability to encode a multiplication of activity which might be motivated by a different spatial image schema (e.g. ZIGZAG). This difference often creates a semantic contrast between VY- and IZ- and suggests that we rather might be dealing with distinct morphemes that have different spatial prototypes.

However, the case is not so simple. The native IZ- is sometimes interchangeable with VY-: *iskupat'sja* and *vykupat'sja* 'take a swim', *izmazat'* and *vymazat'* 'smear all over', etc. Although postulating two IZ-morphemes for Modern Russian might have solved the problem of ambiguity, I argue that such an analysis would be counter-intuitive and very unnatural. In Figure 1, I have shown that all uses of IZ- and VY- are incorporated into a single radial network of polysemy, where all submeanings are interrelated. I assume that in Modern Russian all uses of IZ- belong to a single morpheme, which seems to have two centers of gravity that correlate with different origins.

Therefore, I conclude that IZ- and VY- represent a borderline case between Non-Standard Allomorphy and Non-Allomorphy. A strong argument for allomorphy comes from the near-identical semantics of the two prefixes. VY- and IZ- are closely and systematically connected and share most of their submeanings (10 out of 12). At the same time, some parameters of the data are challenging for an analysis of these prefixes in terms of allomorphs: a large distributional overlap (verbs that can attach both prefixes), cases of semantic contrast, and partial difference in spatial image schemas (Figure 2).

8.7 Conclusions

In this chapter, I presented a detailed semantic account of 989 verbs formed by the prefixes VY- and IZ-. The amount of data analyzed here exceeds all previous accounts, because it goes far beyond dictionary materials by including an abundance of novel coinages attested in the RNC. Taking into account token frequency made it possible to distinguish between standard and marginal verbs and thus look at both the core of the system as well as at its potential in terms of productive patterns of derivation.

For the first time, in this study I have brought together the hypothesis about different origins of the modern Russian IZ- advocated in Belozercev 1966, a comprehensive dataset of Modern Russian verbs, and a semantic account of prefix polysemy in terms of Cognitive Linguistics.

I found that the two origins of IZ- correlate with the hybrid properties of this prefix in Modern Russian and to some extent explain the multidimensional relationship that IZ- has with the prefix VY- (e.g. register differences under semantic overlap vs. semantic contrast). I conclude that the distribution of VY- and IZ- according to Russian vs. Slavonic lexical strata accounts only for a portion of the data. IZ- turns out to be a highly productive prefix that competes with VY- in forming new verbs and marking perfective aspect. Moreover, the two prefixes are found to be semantically comparable.

I suggest that more research is needed in order to clarify the relationship and status of VY- and IZ- in Modern Russian. At this point, many properties of the data look rather controversial, pointing in different directions and suggesting different accounts.

I assign VY- and IZ- a borderline status between Non-Standard Allomorphy and distinct but closely connected morphemes.

In future research, it would be fruitful to test whether the distribution of VY- and IZ- is governed by the semantic classes of base verbs and whether prosodic properties of the base (e.g. number of syllables or place of stress) correlate with different prosodic properties of VY- and IZ-.

Chapter 9

Rival prefixes in word-formation of Russian change-of-state (factitive) verbs: O- vs. U-

9.1 Introduction

Class-changing derivations like *large* > *enlarge* and *little* > *belittle* convert a word of one morphological class into another and may seem to lack any semantic content. However, if a language uses more than one formant for such a morphological operation, a question arises: what triggers the existence of several morphemes with identical morphological function? If they all serve the same purpose, what determines the choice among them?

I address these questions in the present study of Russian, where one can come across a series of close synonyms like *ob''-jasn-it'*, *u-jasn-it'*, *po-jasn-it'*, *vy-jasn-it'*, *pro-jasn-it'*, *raz''-jasn-it'*, and *iz''-jasn-it'*, all of which can be glossed as 'clarify' and are derived from the same adjectival base *jasn-yj* 'clear'. Such deadjectival change-of-state verbs constitute a productive and expanding part of Russian grammar, yet, this phenomenon remains largely overlooked.

In this chapter I report on a study of Russian change-of-state verbs and prefixes employed in their derivation.¹³⁷ I primarily focus on those change-of-state verbs that are formed from non-verbal bases and carry the factitive meaning 'make X be Y', denoting a qualitative change of an object. Following Townsend (1968: 143), I will henceforth call such verbs *factitive*, because I find this term more accurate regarding the data I examine.

I examine derivational patterns of both standard and novel factitive verbs and address three research questions: First, what determines the choice of the prefix? Second, what status do these rival verbalizing prefixes have in Modern Russian derivational morphology? And, third, should these prefixes be analyzed as allomorphs of a single morpheme-verbalizer or rather as distinct morphemes that just happen to perform the same function?

This chapter consists of two parts: the first part is based on corpus data, while the second part reports on an experiment. I start with a brief note on terminology (9.2) and provide some background information about factitive verbs in Russian regarding their productivity (9.3.1), affixes (9.3.2-3), and bases (9.3.4). In 9.3.5, I address the question of extensive prefix variation in factitive verbs and propose a major factor that governs the choice of the prefix. In 9.4, I focus on the competition between the two most productive patterns in O- and U- and propose an account that explains the choice between the two prefixes in corpus data. In 9.5, I test experimentally whether novel marginal verbs in O- and U- differ in terms of acceptability ratings assigned by native speakers. I put together my findings in 9.6 and discuss the implications of this study for my approach to allomorphy.

¹³⁷ Different parts of this study were presented at the seminar of the CLEAR group at the University of Tromsø in 2011, at SCLA-2013 in Zagreb, AATSEEL-2013 in Boston, ICLC-2013 at the University of Alberta, and SCLA-2014 at Harvard University. I am grateful to all audiences for fruitful feedback.

9.2 Terminology

The term *change-of-state (COS) verbs* might be better known in the literature than the term *factitive*, yet I adopt the latter because it is accurate with regard to the type of verbs I investigate in this chapter. In this study, under *factitive* verbs I understand those *change-of-state* verbs that are formed from non-verbal bases (not only adjectival, but also nominal, pronominal, adverbial, or phrasal) and bear the meaning ‘make/cause X be Y or Y-er’.

The term *change-of-state verbs* has a long tradition. Levin 1993 introduces this term in her book on classes of English verbs, where she offers a thorough classification of verbs based on their semantics and syntactic behavior¹³⁸. Levin describes *change-of-state verbs* as verbs “whose meaning involves causing a change of state” (1993: 240-248). Therefore, this group is large and includes a number of subtypes like *break* verbs (*break, crack*), *bend* verbs (*bend, fold*), *cooking* verbs (*bake, boil, fry*), verbs of *calibratable changes* of state (*climb, differ*), and, among others, deadjectival verbs (*brighten, freshen; humidify, intensify; Americanize, stabilize; accelerate, attenuate*). Thus, according to Levin 1993, the *change-of-state* (COS) verbal class includes deverbal, simplex, and deadjectival verbs.

In the present study, I adopt the term *factitive verbs* following Townsend (1968: 143), because referring to Russian deadjectival, denominal, and other types of verbalizations as *change-of-state verbs* would be imprecise. The notion of *change-of-state verbs* is too broad and includes *inchoatives* (verbs that mean ‘become X’, e.g. *fade*¹³⁹), which I leave beyond the scope of my study.¹⁴⁰

It is quite common to distinguish between *factitives* and *causatives* on the basis of the simplex stem: factitives are denominal and deadjectival (Tucker 1981; Hamp 1998; Townsend 1968: 143), while causatives are deverbal (Nichols 1993; Haspelmath 1993; Song 1996; Sanders & Sweetser 2009). One could also look at *factitives* even more restrictively and take them as deadjectival verbs as opposed to denominal verbalizations, as suggested in Heath 2008. However, Heath himself admits that “the distinctions between causative, factitive, and other transitive derivations are somewhat fuzzy, and there is much overlap in the respective morphologies” (Heath 2008: 326).

9.3 Factitive verbs in Russian: The world of possibilities

In this section, I characterize the overall picture and the extent of the factitive verb phenomenon in Modern Russian. My findings are based on corpus data and suggest that there are many more prefixes involved in formation of factitives than previously assumed: seventeen instead of five (Townsend 1968; Mel’čuk 1998). Moreover, I show

¹³⁸ According to Levin (1993: 1), the behavior of a verb is determined to a large extent by its meaning. The key syntactic property of COS verbs is that they participate in transitivity alternation: the same English verb like *clear* can have a causative (‘make clear’) or an inchoative (‘become clear’) reading depending on the context: *The wind cleared the sky* vs. *The sky cleared* (Levin 1993: 240; Hale & Keyser 2002: 99).

¹³⁹ The verb *fade* is mostly inchoative, but it is possible to force it into a transitive reading, as in *The sun faded the curtains*.

¹⁴⁰ Note that in Chapter 5 on O- and OB- I used the term *factitive* in a broader sense and referred to both factitive (in the narrow sense) and inchoative verbs.

that the variety of patterns is very rich, including not only adjectival bases, but also other types of non-verbal bases.

9.3.1 Productivity

Factitive verb formation is a prominent and productive phenomenon in Modern Russian. Factitive verbs are not only well-represented in the standard Russian lexicon (e.g. *odomašnit* 'domesticate', *uskorit* 'speed up', *razobščit* 'disjoin', *zanutnit* 'obscure'), but are also easily and often formed by speakers *ad hoc*. As attested in the Russian National Corpus, speakers of Russian continually coin novel factitives like *ukonkretit* 'concretize', *obytovit* 'make mundane', *rasteplit* 'warm up', *otknopit* 'unfasten', *obsmetanit* 'cover with sour cream', *podnovit* 'refresh' – verbs that fit perfectly into their context, are semantically clear and morphologically transparent, but nevertheless sometimes are only marginally acceptable for other speakers due to their innovative and non-conventional status.

Already back in the 1970s, Vinogradov (1972: 348) pointed out that the formation of new factitive verbs via a combination of prefixation and suffixation was an "increasingly prominent tendency" of Russian. The factitives that Vinogradov mentioned include denominal and deadjectival verbs like *uglubit* 'deepen', *razbazarit* 'waste by foolishness', *prizemlit'sja* 'land', *obobščestvit* 'socialize', *uplotnit* 'condense', *razkulačit* 'dispossess the kulaks', *skosobočit* 'make crooked', *zakavyčit* 'put inside quotation marks'. These verbs started out as innovations but were later conventionalized and became part of the standard lexicon. Other novel factitives have not made their way that far and remained marginal in the system.

A number of scholarly articles have brought attention to this kind of verbs, but were mostly limited to brief remarks on a handful of examples found in newspapers, fiction, or creative colloquial speech: *uproзраčit* 'make transparent' (Zemskaja 1992: 86); *prioxotit* 'make interested', *raspestrit* 'make colorful', *peretončit* 'make too thin' (Uluxanov 1996: 141-5). Scattered across various sources, numerous marginal factitives have remained on the borderline between words and non-words, where they have managed to escape a thorough linguistic investigation. Partly due to the fact that both standard and marginal factitive verbs have not been extensively collected, the extent of this phenomenon in Russian has been largely overlooked. For this reason, the central research question of this chapter, the choice of the prefix in such derivatives, has never been raised.

Apart from being productive and well attested, Russian factitive verbs show a rich variety of patterns. Few sources cite factitive prefixes other than O- and U-. The most detailed descriptions of the matter offered by Townsend (1968: 143) and Mel'čuk (1998: 384) add prefixes that are less commonly used in factitive derivation: ZA- (*za-trudn-i-t* 'complicate'), PO- (*po-niz-i-t* 'lower'), S(O)- (*s-bliz-i-t* 'bring together'), and RAZ- (*ras-šir-i-t* 'broaden'). At the same time, Uluxanov (1996: 132) mentions that any of the nineteen Russian verbal prefixes can potentially be used in a factitive verb, but does not offer a complete set of examples.

9.3.2 Morphological construction: Affixes involved in derivation

Factitives are usually recognized on the basis of semantics, bearing the meaning 'cause qualitative change of an object' (Mel'čuk 1998: 384). However, this is not helpful in distinguishing between a Russian factitive like *ob''-jasn-it* 'clarify' and a non-factitive

like *pere-delat* ‘redo’. For this study I adopt a combination of semantic and derivational criteria. I define factitive verbs as transitive verbs with the meaning ‘make X be Y or Y-er’ formed from a non-verbal base. In this light, it is crucial what kind of affixes are employed in the verbalization process.

In English, factitive verbs can be formed via conversion (*calm*_{ADJ} > *calm*_V), prefixation (*large* > *enlarge*, *little* > *belittle*), or suffixation (*white* > *whiten*, *simple* > *simplify*).

In Russian, when a factitive verb is formed, it employs a combination of affixes – a prefix and a suffix, as in adjective *jasn-yj* ‘clear’ > factitive verb *ob’-jasn-i-t’* ‘clarify’. Note that alongside the adjectival base *jasn-yj* ‘clear’, there is no existing verbal base **jasnit*’, so we have to admit that the two elements, the suffix *-i-* and the prefix *ob-*, attach to the stem *-jasn-* not one after another but simultaneously in a single act of derivation. In this sense, they form a single morpheme (cf. Zenskaja 1992: 85; Uluxanov 1996: 132; Nemčenko 1979: 39), or a prefix-suffix combination that could be referred to as a “circumfix” or a “morphological construction” (Booij 2010: 16).

The variation of prefixes is wide, but the suffix is mostly the same – the suffix *-i-*. This suffix is the major verbalizing operator in Russian (Uluxanov 1996: 132)¹⁴¹ and can in fact form factitive verbs even alone, without the help of the prefix: *belyj* ‘white’ > *belit* ‘whiten’, *suxoj* ‘dry’ > *sušit* ‘make dry’. Such unprefixated derivatives are mostly imperfective and are not always transitive: compare the intransitive derivative *kislit* ‘taste sour’ from the adjective *kislyj* ‘sour’. The attachment of the prefix typically makes the derived verb perfective (*ob’jasnit* ‘clarify’). Because unprefixated factitives are well-described in Rogoza 1992, I leave them outside the scope of this study.

9.3.3 Broad variation in prefixes

I suggest that the production of factitive verbs in Russian is a much broader and more diverse phenomenon than previously assumed. My findings show that deadjectival factitive verbs in Russian employ more than the five prefixes listed in Townsend 1968 and Mel’čuk 1998. Data from the Russian National Corpus provides examples for seventeen prefixes (out of the total nineteen Russian aspectual prefixes). In Table 1, I provide some examples that illustrate the use of seventeen Russian prefixes in factitive verbs.

Note that all these verbs are formed from adjectives listed in Table 1 and lack simplex verbal bases. This lends support to the argument that the prefix and the suffix form a morphological construction.

¹⁴¹ Alternative suffixes *-ova-* or *-irova-* are used in factitive verbs rather rarely: e.g. *radovat* ‘gladden’ < *rad* ‘glad’; *poëtizirovat* ‘poeticize’ < *poëtičnyj* ‘poetic’. In denominal verbs formed via a circumfix one can also rarely find the suffix *-a-*: *o-pojas-a-t’* ‘belt around’ (< *pojas* ‘belt’), *ob-uzd-a-t’* ‘restrain’ (< *uzda* ‘bridle’), *raz-gnev-a-t’* ‘make angry’ (< *gnev* ‘anger’) (Uluxanov 1996: 134). In this study I focus primarily on those verbs that are formed via the suffix *-i-* and leave other suffixes aside.

Prefix	Factitive verb	Gloss-V	Adjectival base	Gloss-ADJ
<i>o(b)-</i>	<i>odomašnit'</i>	'domesticate'	<i>domašnij</i>	'domestic'
<i>u-</i>	<i>umen'shit'</i>	'reduce'	<i>men'she</i>	'less'
<i>raz-</i>	<i>razobščit'</i>	'disjoin, disconnect'	<i>obščij</i>	'joint'
<i>za-</i>	<i>zanutnit'</i>	'obscure'	<i>mutnyj</i>	'muddy'
<i>s-</i>	<i>sokratit'</i>	'shorten; reduce'	<i>kratkij</i>	'short'
<i>pod-</i>	<i>podnovit'</i>	'refresh'	<i>novyj</i>	'new'
<i>pri-</i>	<i>priobščit'</i>	'join smb to'	<i>obščij</i>	'joint'
<i>pere-</i>	<i>pereogromit'</i>	'build too large'	<i>ogromnyj</i>	'huge'
<i>iz-</i>	<i>istoščit'</i>	'exhaust the supply'	<i>toščij</i>	'thin'
<i>pro-</i>	<i>proredit'</i>	'thin out'	<i>redkij</i>	'rare'
<i>na-</i>	<i>nasladit'(sja)</i>	'delight'	<i>sladkij</i>	'sweet'
<i>po-</i>	<i>pojasnit'</i>	'explain, clarify'	<i>jasnyj</i>	'clear'
<i>ot-</i>	<i>otčuždit'</i>	'estrangle; alienate'	<i>čuždyj</i>	'alien'
<i>vy-</i>	<i>vyjasnit'</i>	'figure out'	<i>jasnyj</i>	'clear'
<i>vz-</i>	<i>vzbystrit'</i>	'speed up'	<i>bystryj</i>	'fast'
<i>pre-</i>	<i>presytit'</i>	'sate; stuff'	<i>sytyj</i>	'full'
<i>voz-</i>	<i>vozveličit'</i>	'aggrandize; glorify'	<i>velikij</i>	'great'

Table 1: Prefixed deadjectival factitive verbs: selected examples.

9.3.4 Possible non-verbal bases

Factitive word-formation is not limited to derivation from adjectives. Townsend (1968: 144) mentions also nominal bases, as in *o-svobod-i-t'* 'liberate (expose to freedom), make free' formed from the noun *svoboda* 'freedom'. Another example is the verb *u-sil-i-t'* 'strengthen (or expose to strength, make strong)', formed from the noun *sila* 'strength'. Townsend makes the argument that "submitting' something to the noun is essentially the same as 'making' it the corresponding adjective, since the meaning 'expose to', or 'submit to', is very similar to the 'imposition of a quality' that we observe in deadjectival factitive verbs". This observation can be supported by numerous denominal factitive verbs that also employ a broad variety of prefixes, as shown in Table 2.

Prefix	Factitive verb	Gloss-Verb	Noun	Gloss-Noun
<i>o(b)-</i>	<i>obiletit'</i>	'give a ticket'	<i>bilet</i>	'ticket'
	<i>obsmetanit'</i>	'cover with sour cream'	<i>smetana</i>	'sour cream'
	<i>ofrancuzit'</i>	'make French in style'	<i>francuz</i>	'Frenchman'
<i>u-</i>	<i>udalit'</i>	'send off; cut away'	<i>dal'</i>	'distance'
	<i>udočerit'</i>	'adopt a female child'	<i>doč'</i>	'daughter'
	<i>uglubit'</i>	'deepen'	<i>glub'</i>	'depth'
<i>raz-</i>	<i>rassekretit'</i>	'reveal the secret'	<i>sekret</i>	'secret'
	<i>raskulačit'</i>	'eliminate kulaks'	<i>kulak</i>	'kulak'
	<i>raskošelit'</i>	'make smb spend money'	<i>košel'</i>	'purse'
<i>za-</i>	<i>zazemlit'</i>	'ground'	<i>zemlja</i>	'ground'
	<i>zakabalit'</i>	'enslave'	<i>kabala</i>	'bondage'
	<i>zabolotit'</i>	'waterlog'	<i>boloto</i>	'swamp'
<i>s-</i>	<i>sgrudit'</i>	'bring into a pile'	<i>gruda</i>	'pile'

	<i>sblizit'</i>	'bring close to each other'	<i>bliz'</i>	'nearness'
<i>pod-</i>	<i>podtenit'</i>	'shade slightly'	<i>ten'</i>	'shade'
	<i>podytožit'</i>	'sum up'	<i>itog</i>	'total'
<i>pri-</i>	<i>prizemlit'</i>	'land'	<i>zemlja</i>	'earth'
	<i>prigubit'</i>	'touch with lips (a cup)'	<i>guba</i>	'lip'
<i>pere-</i>	<i>peremestit'</i>	'transfer'	<i>mesto</i>	'place'
<i>iz-</i>	<i>iskorenit'</i>	'eradicate'	<i>koren'</i>	'root'
	<i>ispepelit'</i>	'incinerate'	<i>pepel</i>	'ashes'
	<i>izrešetit'</i>	'make a sieve out of smth'	<i>rešeto</i>	'sieve'
<i>pro-</i>	<i>provetrít'</i>	'air the room'	<i>veter</i>	'wind'
	<i>prosročít'</i>	'delay'	<i>srok</i>	'deadline'
<i>na-</i>	<i>naxoxlit'</i>	'ruffle up'	<i>xoxol(ok)</i>	'tuft'
<i>po-</i>	<i>povysít'</i>	'raise'	<i>vys'</i>	'height'
	<i>pomestít'</i>	'put in a place'	<i>mesto</i>	'place'
	<i>ponizít'</i>	'reduce'	<i>niz</i>	'bottom'
<i>ot-</i>	<i>otštorít'</i>	'remove a curtain'	<i>štora</i>	'curtain'
	<i>otsročít'</i>	'prolong'	<i>srok</i>	'deadline'
	<i>otknopít'</i>	'unfasten'	<i>knopka</i>	'press-button'
<i>vy-</i>	<i>vysvobodít'</i>	'make free'	<i>svoboda</i>	'freedom'
	<i>vydvorít'</i>	'evict'	<i>dvor</i>	'yard'
	<i>vyčlenít'</i>	'single out'	<i>člen</i>	'member of'
<i>vz-</i>	<i>vzbugrít'</i>	'make a hillock'	<i>bugor</i>	'hillock'
<i>pre-</i>	<i>prepojasat'</i>	'put a belt on'	<i>pojas</i>	'belt'
<i>voz-</i>	<i>vozglavit'</i>	'lead'	<i>glava</i>	'leader'
	<i>vozmestít'</i>	'compensate'	<i>mesto</i>	'place'
	<i>voskrylit'</i>	'fly upward'	<i>krylo</i>	'wing'
<i>v-</i>	<i>vmestít'</i>	'fit in'	<i>mesto</i>	'place'
	<i>vručít'</i>	'hand in'	<i>ruka</i>	'arm'
	<i>vnedrít'</i>	'embed'	<i>nedra</i>	'entrails'

Table 2: Denominal factitive verbs: illustrative examples.

In addition to adjectives and nouns, Russian morphology also recruits other types of non-verbal bases in order to form a factitive verb: pronouns, numerals, adverbs, and even noun phrases and prepositional phrases, as I show in Table 3.

Base type	Base	Gloss		Factitive V	Gloss
Pronoun	<i>svoj</i>	'one's own'	>	<i>prisvoit'</i>	'make one's own'
	<i>ničto</i>	'nothing'	>	<i>uničtožit'</i>	'destroy'
Numeral	<i>desjatero</i>	'tensome'	>	<i>udesjaterit'</i>	'increase ten times'
	<i>dvoje</i>	'twosome'	>	<i>udvoit'</i>	'double'
Adverb	<i>naružu</i>	'outwards'	>	<i>obnaružit'</i>	'reveal'
Phrase	<i>bez boli</i>	'without pain' (PP)	>	<i>obezbolit'</i>	'anaesthetize'
	<i>ravnyj ves</i>	'equal weight' (NP)	>	<i>uravnovesit'</i>	'balance'

Table 3: Various types of bases of factitive verbs in Modern Russian.

Note that all factitives listed Table 3 are also formed via the prefix-suffix morphological construction (circumfix), since they have no verbal base: there is no **svoit'*, **ničtožit'*, **desjaterit'*, **naružit'*, **bezbolit'*, or **ravnovesit'*.

Summing up, I have outlined the morphological system of factitive verbs in Modern Russian, focusing on the variety of affixes involved and types of non-verbal bases that they can be attached to. The question that I turn to in the next section is what organizes this variation.

9.3.5 The Spatial Motivation Hypothesis: What determines the choice of the prefix

Given that all prefixes in such verbs perform a factitivizing function, Townsend (1968: 143) suggests that “[t]he prefixes involved mean simply “factitive”; it is not necessary to look for another meaning”. Taking this approach, Townsend however does not explain what drives the choice of the prefix in such verbs.

Mel’čuk (1998: 384) views the choice of the factitive prefix as determined individually for each adjectival base¹⁴²: *utjaželit'* ‘make heavier’ but *oblegčit'* ‘lighten, simplify’; *rasširit'* ‘broaden’ but *suzit'* ‘narrow down’; *sostarit'* ‘age’ but *omolodit'* ‘rejuvenate’. This suggests that factitive derivation in Russian is far from regular and that even in morphologically transparent verbs factitive prefixes are stored as semantically bleached parts of factitive lexemes. This leads us to assume that the choice of the factitive prefix is nearly random and cannot be predicted. Yet this view cannot account for the production of novel factitive verbs, where speakers have to pick a prefix from the available inventory as they coin a verb.

I argue that the distribution of data across factitive prefixes reveals a part of Russian grammar that actively restricts the choice of a prefix and simultaneously is flexible enough to allow some variation. I propose a *Spatial Motivation Hypothesis* and formulate it as follows:

The Spatial Motivation Hypothesis

In the combinations of affixes employed in Russian factitive verbs, the suffix carries out the job of verbalization and turns a non-verbal base into a verb, whereas the prefix contributes the semantic interpretation of the base and specifies the derivative in terms of the perfective aspect. The semantic interpretation takes place as a result of semantic interaction between the meaning of the base and the meaning of the prefix. Furthermore, I suggest that in factitive verbs the prefixes retain their idiosyncratic spatial semantics which motivates their use in deverbal verbs. In other words, the choice of the prefix is motivated by the spatial image schema (the prototype).

In this light, I suggest that different prefixes employed in factitive verbs do not display an allomorphic relationship, but rather belong to distinct morphemes with different spatial prototypes. The use of these prefixes for non-verbal bases constitutes a sub-domain of their application in addition to their use in the domain of verbs.

¹⁴² The original text reads as follows: «v formal’nom otnošenii ruskij transformativ ves’ma dalek ot reguljarnosti: on obrazuetsja s pomošč’ju suffiksa-verbalizatora *-i-* i prefiksa, vybor kotorogo opredeljaetsja leksikografičeski» (Mel’čuk 1998: 384).

I propose that each prefix in these verbs provides its own semantic contribution based on its spatial prototype and image schema. In support of this argument I provide several illustrative examples with the prefixes PRI- and RAZ-.

The prefix PRI- has the prototypical spatial meaning ARRIVE manifested in derivatives like *priletet* 'arrive flying' formed from the simplex motion verb *letet* 'fly' (Endresen et al. 2012: 262; Makarova 2014: 134). A closely related spatial submeaning of PRI- is ATTACH, which additionally implies that the Trajector (the object which is attached) is smaller than the Landmark and becomes fixed in its final destination: *privjazat* 'tie to' < *vjazat* 'tie'.

Factitive verbs indicate the same spatial meaning ATTACH: compare the standard verbs *priobščit*(*sja*) 'make someone share something' formed from the adjective *obščij* 'shareable, common' and *prisvoit* 'make one's own' formed from the pronoun *svoj* 'one's own'. Exactly the same spatial use of the prefix PRI- is utilized in the marginal coinage *prisebjakat* 'make one's own' attested in the corpus and formed from the reflexive pronoun *sebja* 'self':

- (1) *Čušoj podarok prisebjakala, i brov'ju ne povela* [Ženščina + mužčina: Brak (forum) (2004)] 'She **made** someone else's present **her own** without the flicker of an eyelash.'

The prefix PRI- frequently forms factitive verbs from nouns, where it typically contributes the same spatial meaning ATTACH in derivatives like *prikarmanit* 'make one's own, lit. put into one's pocket' where the nominal base *karman* 'pocket' denotes the point of destination. Likewise, PRI- contributes the submeaning ARRIVE to nouns *zemlja* 'land, Earth' and *guba* 'lip' and creates factitives *prizemlit*(*sja*) 'land' and *prigubit* 'bring to lips and taste'. Metaphorical ARRIVE can be observed in the factitive verb *priskučit* 'lit. arrive at the state of boredom' from the noun *skuka* 'boredom', similarly to deverbal derivative *pripomnit*(*sja*) 'bring to one's mind from memory' from the verb *pomnit* 'remember'.

The nominal base can designate not only place of arrival, but also the means of attachment, as evidenced by the factitive verbs *primagnitit* 'attach via a magnet' and *priarkanit* 'noose' formed by combining the prefix PRI- in submeaning ATTACH with the bases *magnit* 'magnet' and *arkan* 'lasso'.

Summing up, factitive derivatives in PRI- demonstrate the same spatial meanings ARRIVE and ATTACH of this prefix as deverbal derivatives do. This suggests that we are dealing with the same prefix PRI- but in a different factitivizing function. Moreover, the choice of this prefix in factitive verbs is apparently determined by the spatial semantics that PRI- contributes to verbal simplexes. As I show below, the same generalizations hold for another prefix frequently used in factitive word-formation, the prefix RAZ-.

The prefix RAZ- contributes to verbal bases the semantics of dissociation, as opposed to PRI-, and has the prototypical spatial meaning APART (see Chapter 3), as in *raspilit* 'saw apart' < *pilit* 'saw'. In factitive verbs we observe the same spatial meaning APART of the prefix: compare the verb *razobščit* 'disjoin' formed from the adjective *obščij* 'shareable, common' and *razdvoit* 'split in two' from the numeral *dvoe* 'twosome'.

A related spatial submeaning of RAZ- manifested in verbs is SPREAD, which refers to a situation where the edges of an object move apart, but the internal cohesion of the object is maintained: *rasstelit* 'spread out (a cloth)'. In factitive verbs we find a parallel use of RAZ-: *rasklešit* 'make bell-bottomed, flared (of clothes)' formed from the analytic

adjective *kleš* 'flare (leg, skirt)' and *rasširit'* 'broaden' formed from the noun *šir'* 'breadth'.

The metaphorically interpreted meaning SPREAD is expressed by RAZ- in the factitive verb *razmnožit'* 'multiply, make many copies' formed from the numeral *mnogo* 'many'. The same meaning SPREAD of the prefix RAZ- is also manifested in verbs that describe affecting the surface, both deverbal like *razrisovat'* 'draw all over' (< verb *risovat'* 'draw') and denominal *razuzorit'* 'cover all over with patterns' (< noun *uzor* 'pattern').

Denominal factitives often represent another submeaning of RAZ-, UNDO, which is found in standard verbs like *raskulačit'* 'eliminate kulaks, lit. undo the status of kulak' (< *kulak* 'kulak') and *rassekretit'* 'reveal the secret' (< *sekret* 'secret') and novel coinages like *raskavyčit'* 'delete or lit. undo quotation marks' (< *kavyčki* 'quotation marks'). In parallel, the same use of RAZ- in the submeaning UNDO is very productive in deverbal derivatives: e.g. *razdumat'* 'change one's mind' (< *dumat'* 'think'), *rasšifrovat'* 'decipher' (< *šifrovat'* 'encipher').

Overall, we again observe that factitive verbs feature the same submeanings of the prefix RAZ- that are otherwise well attested in deverbal derivatives, including both spatial meanings like APART and SPREAD, and even more abstract meanings like UNDO. This parallelism suggests that we are dealing with the same morpheme RAZ-, which is multifunctional and can be attached to verbal bases to form causative verbs as well as to non-verbal bases to form factitive verbs.

Parallel examples of the kind that I provide for PRI- and RAZ- can be found for other prefixes. However, an extensive study of all such uses is beyond the scope of this chapter. My goal was to indicate the prominent pattern which appears from comparison of prefixed derivatives formed from verbal and non-verbal bases.

I conclude this section by asserting that the prefixes employed in factitive verbs constitute distinct morphemes that share a structural pattern (PREFIX-STEM-SUFFIX) rather than allomorphs of a single verbalizing morpheme. This conclusion is grounded in the observation that the prefixes in question express semantics that is closely compatible, if not identical, with their semantics in deverbal derivation. I have mostly focused on parallel deverbal and factitive uses of prefixes in their concrete spatial meanings, but there is also evidence of parallel uses of prefixes in metaphorical and abstract submeanings. The choice of the prefix in factitive verbs is thus governed by its spatial semantics manifested in deverbal derivation.

9.4 The corpus study of O- and U-

In this section, I explore whether the hypothesis outlined in 9.3.5 also applies to the prefixes O- and U-, which are the most frequently and productively used prefixes in formation of factitive verbs. Although widely recognized in the literature, these two verbalizing patterns have never been contrasted and analyzed as rival candidates in productive synchronic derivation.

The question that I address is: how do speakers distinguish between these two patterns and, in particular, what determines the choice of the prefix? At first glance, the spatial prototypes of these prefixes (MOVE AROUND vs. MOVE AWAY respectively) that motivate their use in deverbal derivation are not helpful in predicting the choice of the prefix in factitive verbs: compare *o-složn-i-t'* and *u-složn-i-t'*, both of which denote 'complicate' and are formed from the adjective *složnyj* 'complex'.

I suggest that in such uses the prefixes O- and U- have undergone semantic bleaching, so it is not immediately obvious how exactly their factitive uses are related to their spatial schemas. Because O- and U- form the majority of Russian factitive verbs and dominate over other prefixes used in this type of derivation, O- and U- (in combination with the suffix *-i-*) can be regarded as prominent grammaticalized verbalizers.

I present the first corpus-based contrastive study of the two patterns and show that they are not semantically identical. I examine a comprehensive set of verbs culled from the Russian National Corpus and explore why some adjectival bases attach only O-, other adjectives attach only U-, whereas a third group of adjectives can attach both prefixes. In my analysis I discover both a statistically robust difference in the distributions of O- and U- and a prominent degree of distributional overlap – the use of O- and U- with the same adjectival bases. In my proposal, I explain how the two morphological patterns differ semantically and show that it is possible to trace the grammaticalized semantics of O- and U- back to their spatial prototypes.

I explain how the data was collected in 9.4.1, lay out my proposal in 9.4.2, and support it with five arguments in five subsequent sections 9.4.2.1-5. In 9.4.3, I discuss how the use of O- and U- in factitive verbs is related to their overall semantic networks.

9.4.1 Data collection

In this study, I collected and analyzed 422 factitive verbs prefixed in O- and U-, including 155 factitive verbs formed from adjectives and lacking corresponding unprefixed verbs. The source of data is the Modern Subcorpus of the Russian National Corpus, which includes texts created 1950 – 2013 and in total contains over 113 million words. First, I extracted automatically all verbs that begin in *o* and *u* regardless of their derivational structure, and then manually identified factitive verbs according to semantic and structural criteria (Townsend 1968: 143). In order to avoid duplication of data, I counted perfective verbs and their secondary imperfective counterparts as single entries.

Collecting data from the corpus had a number of crucial advantages. Apart from lexemes well-established in the Russian lexicon and described in dictionaries, the corpus also provides data on newly formed factitive verbs which are marginal and not conventionalized. Whereas one could argue that in standard Russian verbs prefixes are “frozen” and there is no on-line choice of a prefix that is made when one speaks, by contrast, in *ad hoc* generated novel verbs speakers do have to choose a morphological pattern as they form a factitive verb. Moreover, the corpus provides not only the most complete set of relevant data but also frequencies of verbal lemmas that reflect how entrenched O- and U-factitives are in language use.

For the purposes of my argument, I focus on those deadjectival factitive verbs that lack any parallel possible verbal base. I explain my motivations and challenges in the following subsection.

9.4.1.1 Challenges of derivational analysis: Multiple motivations

Data collection is complicated by the existence of multiple motivations for many factitive verbs. Recall that we established circumfixal morphemes where no unprefixed verb is attested, as in *ob’jasnit’* ‘clarify’ < *jasnyj* ‘clear’. The unprefixed verb **jasnit’* does not exist in Modern Russian and therefore cannot serve as a base.

However, there also exist many verbs with factitive semantics ‘make X be Y or Y-er’ which in addition to an adjectival or nominal base are also associated with an existing simplex verb: e.g. *op’janit’* ‘intoxicate’ < *p’janyj* ‘intoxicated’ and *p’janit’* ‘intoxicate’. Should such multiply motivated verbs be analyzed as factitive or deverbal derivatives? Can they be both? If the strength of association with the two bases is different, is it possible to capture and measure this difference?

The problem of multiple motivations in Russian word-formation has been widely discussed in the scholarly literature (Uluxanov 1977: 37-55; Schupbach 1975: 112; Zemskaja 1975 for discussion). Polymotivation has been viewed as a pervasive characteristic of Russian derivational morphology (Katyšev 2005; Lopatin 1976 [2007]: 179) and cannot be simply ignored in the present analysis.

A valuable insight on deadjectival factitive verbs comes from Uluxanov (1977: 37), who shows that verbs like *očistit’* ‘clean’ simultaneously exist in two dimensions of structural relations. On the one hand, *o-čistit’* ‘clean’ is associated with a simplex verbal base *čistit’* ‘clean’ and, parallel to perfective prefixed verbs like *o-grabit’* ‘rob’, *o-putat’* ‘enlace’, *o-carapat’* ‘scratch’, it refers to completion of the activity named by the base verb. On the other hand, *o-čistit’* ‘clean’ is also associated with factitive verbs like *oporožnit’* ‘empty’, *ozdorovit’* ‘make healthy’, and *osložnit’* ‘complicate’ that lack an existing verbal base and are formed from the adjectives *porožnij* ‘empty’, *zdrovyj* ‘healthy’, and *složnyj* ‘complex’ respectively. As such, the verb *očistit’* ‘clean’ is associated with the adjective *čistyj* ‘clean’ and has a factitive meaning ‘make something clean’. Thus, in terms of derivational pattern, *očistit’* is also compatible with other factitive verbs formed via a combination of the prefix O- and the suffix -i-. Given the two possible interpretations, Uluxanov nevertheless argues that for verbs like *očistit’* ‘clean’ the verbal base is primary, while the adjectival base is secondary and indirect (“*oposredovannaja motivacija*”, cf. Uluxanov 1977: 36-7). Yet this does not justify ignoring verbs with multiple motivation links: they share with “pure” deadjectival factitives the final set of components and arguably the same reasons of why these components were chosen. Indeed, even if the prefix gets attached after the suffix (*čist-yj* ‘clean’ > *čist-i-t’* ‘clean’ > *o-čist-i-t’* ‘clean’) but not simultaneously (*porožn-ij* ‘empty’ > *o-porožn-i-t’* ‘empty’), there is still a reason why the prefix O- is chosen instead of one of the other eighteen possible perfective markers.

However, the coexisting verbal base suggests an alternative analysis of verbs like *očistit’* ‘clean’. For this reason, I leave such factitive verbs beyond the scope of the present study (instead, these verbs are analyzed in Chapter 5). Although the analysis presented in this chapter can be extended to multiply-motivated factitive verbs, for the sake of the present argument I leave them aside and focus on “pure” deadjectival factitives, i.e. those factitives that are motivated only by adjectives.

There is no clear-cut boundary between “pure” deadjectival factitives and factitives with multiple motivations. The existence of the verbal base has been checked for each verb in the Modern Subcorpus of the RNC and this revealed a number of factitives like *osčastlivit’* ‘make happy’, where a highly frequent adjective *sčastlivyj* ‘happy’ coexists with a few attestations of the simplex verb *sčastlivit’* ‘make happy’. The decision made for such intermediate cases was to count verbal bases with more than twenty corpus attestations as an existing alternative base. If the unprefixated verb was encountered in corpus less than twenty times, it was considered unlikely to serve as a base for that particular factitive. As a result, verbs like *osčastlivit’* (with 4 examples of

sčastlivit') and *okruglit'* (with 10 attestations of *kruglit'*)¹⁴³ were included in the analysis of “pure” deadjectival factitives.

In order to recognize a simplex verb as a possible base for a factitive, formal identity is not enough and sometimes misleading, as shown by Voronova (1985). Apparently, the factitive *ograničit'* ‘limit, set up boundaries’ has as its only base the noun *granica* ‘boundary’ and lacks a semantic link with the corresponding intransitive verb *graničit'* ‘share borders’. Similarly, *pokružít'* ‘do some turns’ is formed from the verb *kružít'* ‘do turns’, while *okružít'* ‘encircle’ is formed from the noun *krug* ‘circle’ (Voronova 1985: 155). I follow this logic in my analysis and argue that *udorožit'* ‘make more expensive’ is not formed on the basis of the intransitive and semantically different verb *dorožit'* ‘appreciate, esteem’ but rather from the adjective *dorogoj* ‘expensive’.

It has been argued that there are factitive verbs that are structurally motivated by an adjective but semantically associated with a noun: e.g. the verb *obvodnit'* means ‘fill with water’ and is therefore semantically motivated by the noun *voda* ‘water’, while structurally it is more compatible with the adjective *vodnyj* ‘watery’. The same goes for the verb *okrovavit'* ‘stain with blood’, semantically motivated by the noun *krov'* ‘blood’, but structurally by the adjective *krovavij* ‘bloody’ (Uluxanov 1977: 41). In these cases, priority was given to the structural relationship which is uncontroversial, while the semantic interpretation is more speculative.

In derivation of factitive verbs the adjectival base often gets truncated: e.g. *bagrjanyj* ‘purple’ > *obagrit'* ‘empurple’ (Tadžibaev 1985: 141). Partial phonological mismatch between a factitive and its base is not sufficient cause to dismiss the base. In recognizing adjectival bases I follow analyses described in the scholarly literature (Uluxanov 1977; Tadžibaev 1985; Voronova 1985).

9.4.1.2 Distribution of O- and U- across standard and marginal factitive verbs

The Russian National Corpus (www.ruscorpora.ru) attests many low frequency factitives like *omuzykalit'* ‘musicalize’ (< *muzykal'nyj* ‘musical’), *ukonkretit'* ‘concretize’ (< *konkretnyj* ‘concrete’), *ovnešnit'* ‘externalize’ (< *vnešnij* ‘external’), illustrated in (2) and (3).

- (2) *No tem ne menee ukonkretim texničeskie parametry – dlja ponimajuščix čitatelej.* [Komp'jutery budut novye (2003)]
'But let us **concretize** technical parameters – for advanced readers.'
- (3) *Fil'my o čudoviščax pozvoljajut zritelju èksteriorizirovat', "ovnešnit'" problemy v forme kinoobrazov, perenesti ix iz sebja na èkran.* [A. Kameneckij. SŠA kak ob"jekt psixoterapii. (2003)]
'Films about monsters enable a viewer to exteriorize, **externalize** problems in the form of movie images, transfer them from inside oneself to the screen.'

Although such novel verbs are usually clear in terms of their meaning and have a transparent and easily parsable morphological structure, they nevertheless have marginal status. At the same time, attestations of such verbs show how productive the morphological pattern is.

¹⁴³ The frequency values for factitive verbs, adjectival bases and possible verbal bases were extracted from the Modern Subcorpus of the RNC on December 19-27, 2012.

By contrast, standard conventionalized factitives in O- and U- (e.g. *odomašnit'* 'domesticate', *uskorit'* 'speed up') give evidence of how well-established the morphological construction is in the standard lexicon.

In order to account for differences in the status of factitive verbs, I have looked at how frequently each verb is attested in the corpus. Since the RNC is a large and well-balanced database of original Russian texts, I assume that it is representative in terms of how frequent, or entrenched, a particular lexeme is in language use.

Table 4 presents the distribution of deadjectival factitives (that lack parallel verbal bases) of different frequencies: i) highly frequent (verbs that have over 200 attestations in the Modern Subcorpus), ii) medium frequent (over 20 but less than 200 attestations) and iii) low frequent (20 attestations and less)¹⁴⁴.

Pattern	High freq (over 200)	Medium freq (over 20 below 200)	Low freq (20 and less, formed <i>ad hoc</i>)	Total
O- ... -IT'	21	21	45 = 68 %	87
U- ... -IT'	27	20	21 = 32%	68
Total	48	41	66 = 100%	155

Table 4. Data overview: Deadjectival factitive verbs that lack unprefixed verbal bases.

Crucially, the second and the third columns contain those factitive verbs that are well established in the Russian lexicon, while the fourth column summarizes *ad hoc* formations. The values presented in Table 4 are type frequencies, i.e. they correspond to the quantities of different verbal lexemes that are attested for each subgroup. For example, there are twenty-one verbs that are highly frequent and prefixed in O-, and there are twenty-seven highly frequent verbs that are prefixed in U-.

A statistical test shows that the difference in distribution of standard deadjectival factitives across the O- and U- patterns is not statistically significant¹⁴⁵, meaning that both patterns are almost equally frequently employed. By contrast, low frequency factitives that are novel and marginal are distributed across O- vs. U- in a statistically significant manner¹⁴⁶, namely O- is two times more frequent (productive) in novel factitive verbs than U-.

9.4.1.3 Distributional overlap in the use of O- and U- across bases

My findings suggest that the vast majority of adjectives that are involved can form a factitive verb by means of only one morphological pattern: 71 adjectives behave like *domašnij* 'domestic' and allow only the pattern with the prefix O-, while 52 adjectives take the other option and employ the pattern with the prefix U-, like *skoryj* 'fast'. In

¹⁴⁴ The thresholds 20 and 200 are not arbitrary and correspond to the ones calculated in the new frequency dictionary Lyashevskaya & Šaroff 2009, also based on the Modern Subcorpus of the RNC (1950-2007). These thresholds are employed in my analysis in order to break down the data into lexical groups of different status. Factitive verbs with corpus frequency 20 attestations and less are clearly marginal.

¹⁴⁵ The statistical analysis was carried out using open-source R software. The Fisher's Exact Test yields p-value = 0.5. This means that there is a fifty percent chance that one would get this distribution, in other words, no difference.

¹⁴⁶ The Fisher's Exact Test yields p-value = 0.01.

addition, there are 17 adjectival bases that are attested in the RNC with both morphological patterns.

This enables me to make a number of observations. First of all, this data suggests that adjectives have preferences for a particular factitivizing prefix: O- or U-. Second, there is no clear-cut boundary in the adjectival preferences for a particular prefix, instead there are transitional cases that for some reason allow both morphological patterns to be employed. Third, different token frequencies of the co-existing factitive counterparts formed from the same adjectival base (e.g. *osložnit'* vs. *usložnit'* 'complicate', both motivated by the adjective *složnyj* 'difficult') show that instead of a clear-cut boundary we have a continuum of gradient data and a scale of choices.

In order to model this gradience, I compare token frequencies¹⁴⁷ of factitive counterparts in the RNC, assuming that they reflect the entrenchment of these verbs in language use. Table 5 provides a complete list of the seventeen adjectives that can take both prefixes. In addition, for the sake of comparison, Table 5 also includes the adjectives *skoryj* 'fast' and *domašnij* 'home', which take only one of the two prefixes. The list is organized according to the relative entrenchment of O- vs. U- factitive counterparts, measured in terms of percentage (columns 6, 7) and ratio (column 8). Table 5 shows that the preferences of adjectival bases for O- vs. U- form a scale with the most polar cases at the extremes (where there is only one prefix attested) and a lot of intermediate cases in between (where one of the counterparts is more or less frequent and therefore more salient than the other).

There are pairs like *osložnit'* and *usložnit'* 'complicate' that are nearly equal in their frequency and salience (410 vs. 311 attestations respectively). By contrast, *osovremenit'* 'modernize' is much more frequent than *usovremenit'* 'make more modern' (59 vs. 1 attestations), while *uprostit'* 'simplify' is much more frequent than *oprostit'* 'simplify' (1,350 vs. 11 attestations accordingly). The difference in frequency is not so dramatic in pairs like *ožestočit'* 'harden' vs. *užestočit'* 'make regulations more strict' (686 vs. 210 attestations respectively) and *otjaželit'* 'make heavy' vs. *utjaželit'* 'make heavier' (34 vs. 93 attestations respectively).

Summing up, type and token frequencies of factitive verbs obtained from corpus reveal important aspects of data that should be taken into account, namely differences in the productivity of O- vs. U- in novel lexemes as well the gradient nature of preferences of the adjectival bases for O- vs. U-. The scale of preferences suggests that this is a probabilistic aspect of grammar. It can be best captured in terms of tendencies rather than strict exceptionless rules.

¹⁴⁷ Token frequency is the number of attestations of a particular factitive verb (taken as a set of all its finite and non-finite forms) in the corpus.

Adjective	Factitive in O-	Factitive in U-	Freq O-	Freq U-	% of O-	% of U-	Ratio
<i>skoryj</i> 'fast'	-----	<i>uskorit'</i> 'speed up'	0	2,008	0	100	0: 1
<i>tverdyj</i> 'firm, hard'	<i>otverdit'</i> 'harden'	<i>utverdit'</i> 'confirm, approve'	14	5,163	0.28	99.72	0.0027: 1
<i>plotnyj</i> 'solid, compact'	<i>oplotnit'</i> 'make compact'	<i>uplotnit'</i> 'condense, compress'	1	201	0.50	99.5	0.0049: 1
<i>prostoj</i> 'simple'	<i>oprostit'</i> 'simplify'	<i>uprostit'</i> 'simplify'	11	1,350	0.81	99.19	0.0081: 1
<i>vlažnyj</i> 'wet'	<i>ovlažnit'</i> 'make wet'	<i>uvlažnit'</i> 'moisten, wet'	3	159	1.25	98.75	0.0188: 1
<i>teplyj</i> 'warm'	<i>oteplit'</i> 'make warm'	<i>uteplit'</i> 'make warmer'	6	205	2.85	97.15	0.0292: 1
<i>korotkij</i> 'short'	<i>okorotit'</i> 'shorten, force to stop'	<i>ukorotit'</i> 'make shorter'	37	787	4.50	95.5	0.047: 1
<i>srednij</i> 'middle, medium'	<i>osrednit'</i> 'make median (in numbers, statistics)'	<i>usrednit'</i> 'make median (in statistics), make mediocre'	27	266	9.21	90.78	0.1194: 1
<i>tjaželyj</i> 'heavy'	<i>otjaželit'</i> 'make heavy'	<i>utjaželit'</i> 'make heavier'	34	93	27.00	73.00	0.36: 1
<i>ser'eznyj</i> 'serious'	<i>oser'eznit'</i> 'make serious'	<i>user'eznit'</i> 'make more serious'	1	1	50.00	50.0	1: 1
<i>složnyj</i> 'difficult'	<i>osložnit'</i> 'complicate'	<i>usložnit'</i> 'complicate'	410	311	57.00	43.0	1: 0.758
<i>žestkij</i> 'hard'	<i>ožestočit'</i> 'harden'	<i>užestočit'</i> 'make rules stricter'	686	210	77.00	23.0	1: 0.3
<i>edinyj</i> 'single, joint'	<i>ob'edinit'</i> 'combine, unite'	<i>uedinit'</i> 'isolate'	4,669	305	94.00	6.0	1: 0.0653
<i>strannyj</i> 'strange'	<i>ostran(n)it'</i> 'defamiliarize'	<i>ustrannit'</i> 'defamiliarize'	20	1	95.00	5.0	1: 0.05
<i>dobryj</i> 'good'	<i>odobrit'</i> 'approve'	<i>udobrit'</i> 'fertilize soil'	2742	108	96.00	4.0	1: 0.0393
<i>jasnyj</i> 'clear'	<i>ob'jasnit'</i> 'clarify'	<i>ujasnit'</i> 'understand, make clear(er) to oneself'	18,149	541	97.00	3.0	1: 0.0298
<i>sovremennyj</i> 'modern'	<i>osovremenit'</i> 'modernize'	<i>usovremenit'</i> 'make more modern'	59	1	98.40	1.6	1: 0.0169
<i>bednyj</i> 'poor'	<i>obednit'</i> 'impoverish'	<i>ubednit'</i> 'belittle, disparage'	199	1	99.50	0.5	1: 0.005
<i>domašnij</i> 'home'	<i>odomašnit'</i> 'domesticate'	-----	53	0	100	0	1: 0

Table 5: Scale of proportions of factitive counterparts according to their token frequencies (includes seventeen adjectives that can take both O- and U-).

9.4.2 The Scalarity Hypothesis: The difference in the factitive use of O- and U-

In word-formation of Russian factitive verbs, O- and U- are rival competing prefixes which have an identical function: both of them are parts of a verbalizing prefix-suffix morphological construction which transforms a non-verbal base into a verb. I propose that while sharing this general structural function, the two prefixes exhibit a subtle semantic difference, as described in *the Scalarity Hypothesis*:

The Scalarity Hypothesis

In factitive verbs, the prefix O- typically encodes non-scalar changes of states, whereas the prefix U- typically encodes scalar changes of states.

Given that there is a choice between two options, O- and U-, why do some adjectives like *domašnij* ‘home’ form a factitive verb only by means of the prefix O- (*o-domašn-it* ‘domesticate’), whereas other adjectives like *skoryj* ‘fast’ are compatible only with the prefix U- (*u-skor-it* ‘speed up’)? I suggest that the reason for such distributional preferences is that two prefixes make different semantic contributions to an adjectival base.

In factitive verbs, the prefix O- bears the meaning ‘make X be Y’ and suggests the simple imposition of a new property on an object. By contrast, the factitive prefix U- usually has the meaning ‘make X be Y-er’ and suggests a scalar change in the degree of a property. Thus, U- appears to be semantically more specific and therefore marked.

Returning to our examples, the factitive verb in O-, *odomašnit* ‘domesticate’, refers to the imposition of a property, which is arguably not gradable: one can either be domesticated or not. The adjective *domašnij* ‘home’ in Russian is relational and does not imply any scale. In particular, this adjective normally does not have forms of comparative and superlative degrees.¹⁴⁸

O-... -IT’	‘make X be Y ’	<i>o-domašn-it</i> ‘domesticate’	< <i>domašnij</i> ‘home’
U-...-IT’	‘make X be Y-er ’	<i>u-skor-it</i> ‘speed up’	< <i>skoryj</i> ‘fast’

According to my argument, the prefix O- does not normally refer to scalar changes and rather means ‘make X be Y’. Therefore, the non-scalar O- is compatible with the non-scalar adjectival base.

By contrast, the factitive verb *uskorit* ‘speed up’ refers to the increase of speed, which is scalar. Speed itself is a gradable property, and the adjective *skoryj* ‘fast’ is qualitative. It has the comparative form *skoree* ‘faster’ and the superlative degree form *skorejšij* ‘the fastest’. In creating a factitive verb, such a scalar adjective is most compatible with the prefix U-, which typically expresses scalar semantics ‘make X be Y-er’.

I suggest that the semantic opposition of the factitive O- and U- can be visualized as shown in Figure 1. Figure 1 depicts a scale of gradation of a property subdivided into three stages: zero (-), positive (+), and comparative (++) degrees. It shows that the two prefixes are specialized on different parts of this scale. O- takes the transition from the null degree of a property to the positive degree: (-) >> (+). The prefix U-, by contrast, expresses the transition from the positive degree of the property to its higher degree (+) >> (++) . Thus the scale is distributed across the two prefixes as visualized in Figure 1.

¹⁴⁸ Theoretically, one can create the forms *domašnee* ‘more home’ and *samyj domašnij* or *domašnejšij* ‘the most home’ and thus construe *domašnij* as a qualitative adjective that refers to a gradable property. However, this is potentially possible to do with any adjective, because the borderline between relational and qualitative adjectives is fuzzy. Still, the comparative and superlative degrees of *domašnij* would deviate from the conventionalized use of this adjective and could only appear in creative or figurative speech.

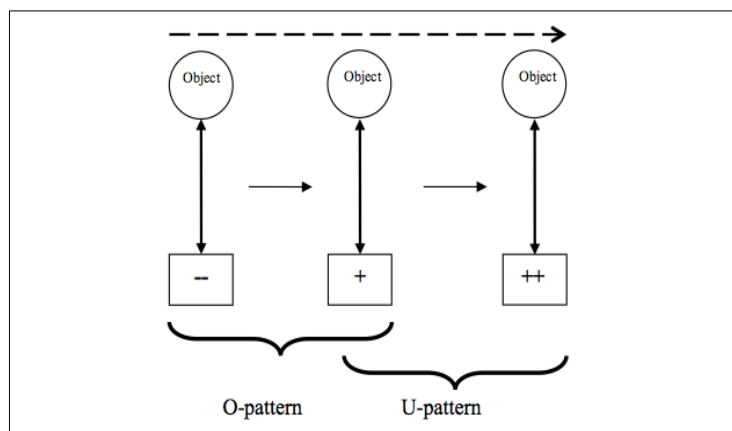


Figure 1: Semantic opposition of the prefixes O- and U- in deadjectival factitive verbs.

Indeed, many deadjectival verbs prefixed in U- refer to an increase in the degree of a property. Usually, this property is not new to an object, as it was characterized with this property earlier. Examples are numerous and I list only a representative sample: *učastit'* 'make more frequent, quicken', *utjaželit'* 'make heavier', *udlinit'* 'lengthen, prolong', *uveličit'* 'increase, extend', *udeševit'* 'make cheaper', *užestočit'* 'make regulations more strict', *ustrožit'* 'make stricter', *ukorotit'* 'make shorter', *ukrupnit'* 'make larger, enlarge', *uskorit'* 'speed up', *ubystrit'* 'increase the speed', *usložnit'* 'make more complicated', *usugubit'* 'increase, aggravate', *uteplit'* 'make warmer', *utončit'* 'make thinner', and others.

Novel factitives also follow this semantic pattern and frequently feature the prefix U- referring to scalar changes of states: *uvkusnit'* 'make more delicious', *udorožit'* 'make more expensive', *ukonkretit'* 'make more concrete', *ukrasivit'* 'make more beautiful', *uprozačit'* 'make more transparent', *user'eznit'* 'make more serious', and *usovremenit'* 'make more modern'.

By contrast, the factitive use of O- is neutral and normally does not imply any scalarity. Semantically, O- carries a property from being absent to being present and simply refers to the imposition of a new property on an object. The object was arguably not characterized with this property before (recall that U- applies not to new properties, but to properties that are already present). This semantic contribution of O- can be seen in many factitive verbs: *oglupit'* 'stupefy', *oglušit'* 'deafen', *ogorčit'* 'make upset', *ogrubit'* 'roughen, coarsen', *ožestočit'* 'harden', *ozdorovit'* 'make healthy', *okislit'* 'oxidize', *okruglit'* 'make round', *očerstvit'* 'make unsympathetic, soulless', *otjaželit'* 'make heavy', *otrezvit'* 'make sober', *oščastlivit'* 'make happy', *odomašnit'* 'domesticate', *osovremenit'* 'make modern', *osvetlit'* 'lighten', *osvežit'* 'freshen', *obrusit'* 'Russianize', *oržavit'* 'corrode', *opresnit'* 'make fresh (about water)', *opošlit'* 'vulgarize, banalize', *oporožnit'* 'make empty', *obostrit'* 'sharpen', *obosobit'* 'separate', *obobščit'* 'generalize', *obnaličit'* 'convert into cash', *obnažit'* 'nude', *obmirščit'* 'secularize', *obagrit'* 'empurple', *obogatit'* 'enrich', *obednit'* 'impoverish', *obrxatit'* 'make pregnant', and others.

We find the same pattern in O- in novel factitive verbs like *ovnešnit'* 'externalize', *ovlažnit'* 'make wet', and *omeždunarodit'* 'internationalize'.

It is worth mentioning that this proposal concerning the choice of the prefix in Russian factitives is compatible with the *incorporation theory* proposed for English deadjectival change-of-state verbs by Hale and Keyser (2002: 100). They argue that the derivation of verbs like *clarify* from adjectives involves the process of *incorporation*,

whereby the adjective moves from its external position of the adjectival complement located outside the verb – into the verb: e.g. *make the point clear* > *clarify the point*. Therefore, according to Hale and Keyser (2002: 149), deadjectival change-of-state (or *factitive*, in my terminology) verbs can be decomposed into light heads and adjectival subcomponents: *clarify* = *make clear*. This account was further elaborated into the idea that the internal adjectival subcomponents of factitive verbs maintain their semantic properties and even remain syntactically active (Levinson 2010) and can be modified by degree expressions like *entirely* and *completely* (Rappaport Hovav & Levin 2002: 2).¹⁴⁹ The discussion was mostly centered around telicity issues of factitive English verbs depending on whether they are formed from gradable or non-gradable adjectives. In my study I show that the gradability (or scalarity) of change implied by the adjectival base of a factitive verb is a factor in determining what grammatical marker of perfectivity is used: the marker O- or the marker U-.

The story about the factitive use of O- and U- would be incomplete if I did not discuss the cases of distributional overlap, shown in the 17 minimal pairs of verbs in Table 5. Recall that those are the cases where the same adjectival base can form a factitive verb with both prefixes.

Some minimal pairs show semantic contrast, because the derivatives in O- and U- have developed distinct meanings. For example, the verbs *odobrit* ‘approve’ and *udobrit* ‘fertilize soil’ both formed from the adjective *dobryj* ‘good, kind’ are not interchangeable and refer to entirely different domains. At the same time, *the Scalarity Hypothesis* is supported by these verbs, because the factitive in O- denotes ‘approve’ and does not imply any internal change of the object’s nature or integrity. The counterpart in U-, by contrast, refers to fertilization, which is a change of the soil’s properties.

Verbs in some minimal pairs with O- and U- are interchangeable in some contexts. Usually, such verbs have more compositional semantics, which can be clearly decomposed into the contribution of the prefix and the contribution of the base. Consider the pair of verbs *osložnit*’ and *usložnit*’, where both factitives mean ‘complicate’ and are formed from the scalar adjective *složnyj* ‘complex’. I suggest that in these verbs O- and U- overlap not only in terms of distribution, but also partially in terms of semantics, because the contrast in terms of scalarity of change is neutralized. They have near-identical semantics in the corpus examples (4) and (5), which show a context that allows the use of both factitive counterparts:

- (4) *Odnovremenno sistemnyj blok i monitor perevezti nereal’no. Značit, pridetsja pokupat’ v Moskve. Začem **osložnjat**’ sebe žizn’?* [Pis’mo materi k dočeri (2004)]
 ‘It is not possible to transport both the computer case and the display at the same time. This means that we will have to buy [one of them] in Moscow. Why **complicate** life?’
- (5) *Začem **usložnjat**’ sebe žizn’, berja za edinicu dliny djujmy, futy i mili?* [Je. Rubin. Pan ili propal. Žizneopisanie (1999 – 2000)]
 ‘Why **complicate** life by using such units of length such as inches, feet, and miles?’

¹⁴⁹ I come back to degree modifiers in Section 9.4.2.4 (Argument 4) of the present Chapter.

Note that in (4) and (5) we are dealing with imperfective counterparts of the verbs *osložnit'* and *usložnit'* 'complicate'. Both perfective verbs appear in this context too, although less frequently, and show the same interchangeability as we see in (4) and (5).

However, it is possible that in other contexts the contrast between *osložnit'* and *usložnit'* can be activated, and the variant in O- would refer to complicating something that was easy, whereas the verb in U- would imply complicating something that was rather complex already in the beginning.

Summing up, the pairs of parallel verbs show both semantic contrast and semantic overlap. However, this does not undermine *the Scalarity Hypothesis* presented in this section.

In favor of this proposal I offer a number of arguments: i) the use of comparatives as bases; ii) correlation between prefixes and adjectival classes; iii) restrictions on the choice of the prefix for different non-adjectival bases; iv) analysis of the syntactic context of factitive verbs, and v) exploration of possible phonological factors.

9.4.2.1 Argument 1: Comparatives employ U-

I argue that factitive verbs in U- are semantically motivated by comparative forms of adjectives rather than by positive forms. However, formally this is hard to prove, because the palatalization of the stem-final consonant in the factitive verb can be due to a) the influence of the following suffix *-i-* or b) due to the comparative suffix of the base stem: e.g. the change [g > ž] in *ustrožit'* 'make stricter' can result either from *u-strog-i-t'* < *strogij* 'strict' or from *u-strož-i-t'* < *strože* 'stricter'.

We know that inchoative verbs in *-e-* are semantically motivated by comparative forms of adjectives (Nikitevič 2006: 224) without being formally based on the comparative form. For example, the inchoative verb *tolstet'* 'grow fatter' is semantically motivated by the comparative *tolšče* 'fatter', but formally associated with the positive form *tolstij* 'fat'. So, this possibility is potentially open for deadjectival factitives in U- as well.

My observation about the scalarity of the factitive pattern with U- is supported by cases when a factitive is uncontroversially derived from the comparative form of an adjective, namely in case of suppletive comparative forms¹⁵⁰ (Uluxanov 1977: 55; Nikitevič 2006: 224; Klobukova 1985: 85), as shown in Table 6. For example, the verb *ulučšit'* is formed from the comparative *lučše* 'better', but not from the positive form *xorošij* 'good'.¹⁵¹

¹⁵⁰ Motivation of factitive verbs by suppletive comparative forms is well attested crosslinguistically. For Slavic parallels compare Czech verbs *zlepšit* 'enhance, improve' (< *lepší* 'better_{COMP}', which is the comparative form of *dobrý* 'good'), *zhoršit* 'aggravate' (< *horší* 'worse_{COMP}', the comparative from *zlý* 'bad'). For Germanic parallels consider German verbs (*ver*)*bessern* 'improve' (< *besser* 'better_{COMP}'), *vermehren* 'multiply' (< *mehr* 'more_{COMP}') (cf. Uluxanov 1977: 57). In German, factitive verbs can also have the non-suppletive comparative form with a comparative suffix as a base: *verschlechtern* 'worsen' (< *schlechter* 'worse_{COMP}'), *verkleinern* 'reduce' (< *kleiner* 'smaller_{COMP}'), *vergrößern* 'enlarge' (< *größer* 'larger_{COMP}'), as well as the unprefixed verbs *mildern* 'mitigate', *nähern* 'come closer', *schmälern* 'belittle, narrow down' (Fleischer 1971: 292).

¹⁵¹ An additional example comes from the factitive verb *uveličit'* 'increase, extend'. Uluxanov (1977: 57) argues that formally this verb is motivated by the positive form *velikij* 'great', but semantically – by the comparative form *bol'she* 'larger'.

Verb	Gloss	Adj. comparative	Gloss	Adj. positive	Gloss
<i>umen'sit'</i>	'reduce'	<i>men'se</i>	'smaller'	<i>malyj</i>	'small'
<i>ulučsit'</i>	'improve'	<i>lučše</i>	'better'	<i>xorošij</i>	'good'
<i>uxudšit'</i>	'make worse'	<i>xuže</i>	'worse'	<i>ploxoj</i>	'bad'

Table 6: Factitive verbs with comparative adjectival bases as opposed to adjectives in positive degree.

Crucially, suppletive comparative forms of adjectives serve as bases for factitive verbs in U-, while attachment of the prefix O- to unambiguous comparative bases is unattested. This fact supports my proposal about the contrast between O- and U- in terms of scalar changes of states: adjectival forms of comparative degree imply a higher degree of a property on the property scale and are semantically compatible with U-, which also encodes scalarity. By contrast, comparatives avoid the prefix O-, which is not compatible with the expression of degree.

9.4.2.2 Argument 2: Adjectival classes

In this subsection I report on the significant correlation that I found between the choice of the prefix (O- vs. U-) and the two classes of adjectives.

In the scholarly literature on Russian, it is common to distinguish between two classes of adjectives: *relational* adjectives denote a property that is not gradable (e.g. *evropejskij* 'European'), whereas *qualitative* adjectives refer to a property that is gradable (e.g. *tixij* 'quiet') and therefore they typically have forms of comparative and superlative degree and are often used with degree modifiers (Nørgård-Sørensen 2011; Dixon & Aikhenvald 2004). My proposal in terms of *the Scalarity Hypothesis* predicts that the two types of adjectives should correlate with the choice of the factitivizing prefix: U- is expected to be strongly associated with qualitative adjectives, whereas O- is likely to be chosen by relational adjectives.

My account presented below is based on 155 factitive verbs in O- and U- that have clear adjectival bases and lack unprefixated verbal counterparts. In order to minimize the degree of subjectivity in this study, I collected and analyzed the tags for adjectives (qualitative vs. relational) provided by an independent source – the Russian National Corpus.

Figure 2 summarizes the overall distribution of the two prefixes across adjectival classes. It shows that the opposition of qualitative and relational adjectives correlates with their preferences for O- vs. U-. The difference between the distributions for O- and U- is statistically significant.¹⁵² Note that qualitative adjectives are more numerous and of higher frequency in my database. This might reflect the density of qualitative adjectives in the corpus overall.

Figure 2 shows that the prefix O- can combine with both relational and qualitative adjectival bases but prefers qualitative ones. The prefix U- is more selective with respect to adjectival classes: the relative proportion of qualitative vs. relational adjectival bases in case of U- is different, and the relational adjectives are apparently dispreferred. Recall that the prefix U- requires adjectives that refer to properties, which can be interpreted in terms of gradation and scalarity. This requirement is met in case of

¹⁵² Fisher's Exact Test yields p-value = 0.0096 (95 percent confidence interval: 0.17 to 0.83; sample estimates: odds ratio 0.38).

qualitative adjectives, but some relational adjectives can also allow for the qualitative scalar interpretation.

Figure 2 shows that although the distribution of adjectival classes across the two prefixes is not complementary, relational

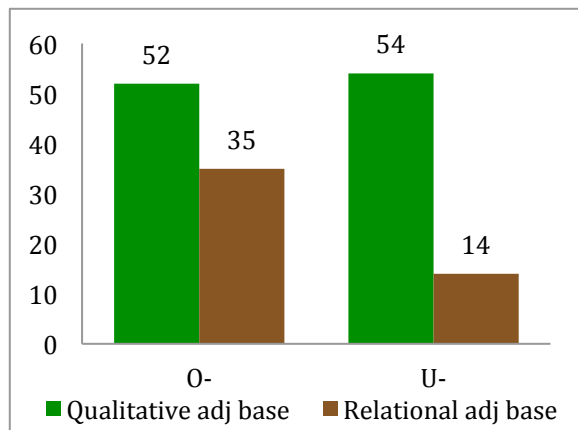


Figure 2: Overall distribution of the prefixes O- and U- across adjectival bases of different types.

adjectives less often combine with the prefix U- than with O-, and this finding supports the prediction of *the Scalarity Hypothesis*.

Note that qualitative adjectival bases are attested with both O- and U-, which might suggest that O- is degree-neutral.

Below I break down the overall distribution of 155 factitive verbs shown in Figure 2 into two subsets. Figure 3 visualizes the distribution of standard factitive verbs (with high and medium corpus frequency), whereas Figure 4 addresses novel formations (with low corpus frequency).

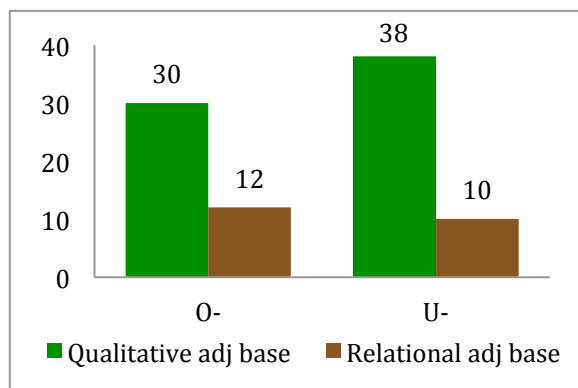


Figure 3: Standard factitive verbs.

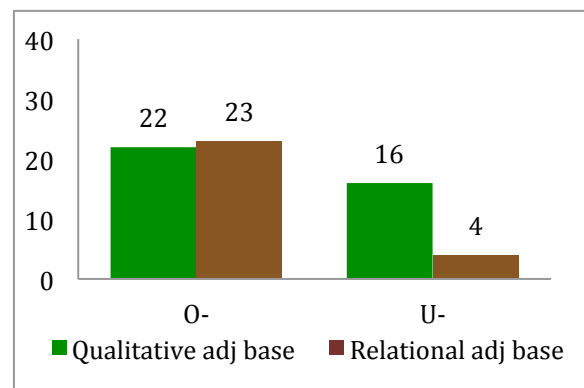


Figure 4: Marginal factitive verbs.

As Figure 3 shows, in standard factitive verbs, both prefixes prefer qualitative adjectives over relational adjectives and show similar pattern of preferences. Here, the difference between O- and U- is statistically non-significant¹⁵³. This suggests that the distribution of prefixes across adjectival bases in the set of standard verbs might be random.

However, in low frequency factitives which represent marginal novel coinages attested in the corpus (Figure 4), the prefix U- has a strong preference for qualitative adjectival bases over relational bases which supports *the Scalarity Hypothesis*. Given the distributions of verbs with O- and U-, the correlation between U- and qualitative adjectival class is statistically significant¹⁵⁴.

¹⁵³ Fisher's Exact Test yields p-value = 0.46 (95 percent confidence interval: 0.22 to 1.9; sample estimates: odds ratio 0.66).

¹⁵⁴ Fisher's Exact Test yields p-value = 0.02 (95 percent confidence interval: 0.05 to 0.9; sample estimates: odds ratio 0.24).

Summing up, O- is compatible with both qualitative and relational adjectives but prefers the former. In novel factitives, the pattern U- has a significant preference for qualitative adjectival bases over relational bases. This supports our claim that U- is associated with scalar changes. The prefix O-, on the contrary, appears to be neutral in terms of scalarity. The findings show that O- is frequently attested with scalar (qualitative) adjectives suggesting that its use is less restrictive than that of U-.

9.4.2.3 Argument 3: Beyond adjectival bases

In this section, I look at factitive verbs derived from non-adjectival bases. Table 7 provides an inventory of possible types of bases and summarizes the distribution of 422 factitive verbs in O- and U- across these types.

Base type	O- verbs	U- verbs	Examples of O- and U-
Only adjective	87	68	<i>odomašnit</i> 'domesticate', <i>uskorit</i> 'speed up'
Adjective and verb	30	10	<i>omračit</i> 'darken, gloom', <i>ukrepiť</i> 'fortify'
Only noun	17	58	<i>obiletit</i> 'sell a ticket', <i>usilit</i> 'strengthen'
Noun and verb	17	49	<i>obvinit</i> 'blame', <i>ustydit</i> 'shame'
Pronoun	1	3	<i>osvoit</i> 'master', <i>uničtožit</i> 'destroy'
Numeral	0	8	<i>udesjaterit</i> 'increase ten times'
Noun phrase	6	3	<i>oblagorazumit</i> 'make prudent', <i>uravnovesit</i> 'balance'
Preposition phrase	64	1	<i>obezglavit</i> 'behead', <i>ubezzakonit</i> 'make illegal'
Total	222	200	

Table 7: Factitive verbs in O- and U- beyond adjectival bases.

There are two important observations that we can make about this data.

First, only the prefix U- can be employed in order to form a factitive verb from a numeral. Note that these factitives always refer to an increase of number: e.g. *u-desjaterit* 'increase ten times' < *desjatero* 'tensome'. The prefix O- is not attested in factitive verbs that have numerals as bases.

Second, it is the prefix O- that can almost exclusively form factitive verbs from prepositional phrases with the preposition *bez* 'without': e.g. *o-bez-glav-i-t* 'behead' < *bez golovy* 'without a head'. Such verbs indicate a change from the state of having a property to the state where the property is lost. In other words, such verbs refer to presence vs. absence of a property that the prefix O- is associated with.

To sum up, the restrictions for the prefix set by numeral and phrasal bases conform to the claim of *the Scalarity Hypothesis* and support the semantic opposition between the prefixes O- and U- in factitive verbs.

9.4.2.4 Argument 4: Degree modifiers

One could expect that the semantic contrast between O- and U- is most evident in minimal pairs of factitive verbs. Yet, often it is rather difficult to tease them apart since they are often near-synonyms and show little difference in argument preferences and typical collocations.

However, I find that lexical markers of scalar properties and scalar changes often accompany factitive verbs in U-. Usually, these are quantificational expressions and

adverbs, or so-called degree modifiers, which emphasize an increase in the degree of a property. Analysis of corpus attestations suggests that degree modifiers are usually used in the close context of verbs in U-, and rarely used with the correlating factitives in O-. Consider the following examples from the Russian National Corpus:

- (6) *V starinu k noge katoržnika prikovyvali cep' s tjaželoj girej, čtoby otjaželit' ego šag i sdelat' nesposobnym k pobegu.* [Ja.I. Perel'man. *Mežplanetnye putešestvija* (1915)]
 'In the old days, they used to attach a chain with a heavy weight to the prisoner's foot, in order to **weigh down** his gait and make him unable to escape.'
- (7) *Sostojanie – inogda do togo tjaželoje, čto ešče by utjaželit' – i uže nel'zja žit'.* [V. Erofeev. *Proza iz žurnala «Veče»* (1973)]
 'The situation is sometimes so difficult that if one were to **make it even more difficult**, then life would be impossible.'

Examples (6) and (7) illustrate the use of the factitive verbs *otjaželit'* 'make heavy' and *utjaželit'* 'make heavier'. In (6), the factitive verb is prefixed in O- and is degree-neutral, because it refers to an assignment of a new quality (heavy weight) which takes place, when a heavy chain becomes physically attached to a prisoner's foot.

By contrast, in (7), the factitive verb prefixed in U- expresses an increase of the quality which is already present in the situation – the metaphorical "heaviness", or complexity of the situation. The verb *utjaželit'* denotes an increase of heaviness and therefore is accompanied with lexical markers of degree: the adverb *ešče* 'even more' and the expression *i uže nel'zja žit'* 'then life would be impossible'.

The same contrast between the verbs in O- and U- holds for examples (8) and (9). Here, we find the degree-neutral verb *osovremenit'* 'modernize' and the degree-marked verb *usovremenit'* 'make more modern' accompanied with the comparative form of an adjective *polučše* 'better'.

- (8) *Ideja «osovremenit'» Dostoevskogo uže provociruet ulybku.* [V. Polupanov. «Dauny» v «xaose» (2001)]
 'The idea of "**modernizing**" Dostoevsky provokes a smile.'
- (9) *Nam nužno specèffekty polučše sdelat', usovremenit'* (from www.google.com).
 'We should make the special effects a little better, **more modern**.'

Examples (10), (11), and (12) follow the same pattern. The verb in U- *usložnit'* 'complicate, make more complicated' is accompanied with degree modifiers *predel'no* 'to the utmost degree' (11) and *raza v dva* 'about two times' (12). In (10), where the verb in O- *osložnit'* 'complicate' is used, such markers of degree are absent.

- (10) *Da i drugie analizy podtverždali predpoloženie o diabete, osložnennom tjaželoj infekciej.* [D. Šraer-Petrov. *Oxota na «ryžego djavola»*, 2009]
 'The other tests also confirmed the supposition that it was diabetes **complicated** by a serious infection.'
- (11) *Ja namerenno predel'no usložnil buduščim ispolniteljam zadaču.* [B. Rudenko. *Spravka o naličii intellekta*, 2006]

'I intentionally **made** the task **more complicated** to the utmost degree for future workers.'

- (12) *Dlja rešenija «problemy» predlagaetsja raza v dva usložnit' otpravku knig.* [kollektivnyj forum. Kniga žalob i predloženij. (2004-2006)]
 'In order to solve the "problem", it is recommended that one **makes** the sending of books about two times more complicated.'

9.4.2.5 Argument 5: Phonology

I tested whether the choice between O- and U- in factitive verbs is conditioned by phonological or prosodic properties of the base. I tagged my data for the following factors: onset cluster of the base (OnsetCluster: yes & no); place of articulation of the base onset (OnsetPlace: dental, labial, palatal, velar); manner of articulation of the base onset (OnsetManner: stop, fricative, affricate, sonorant); number of syllables in the base and in the prefixed derivative (BaseNumberOfSyll; VerbNumberOfSyll: 1, 2, 3, 4, 5); and place of stress in the derivative verb (StressVerbStemOrTheme: stem or thematic vowel).

For the statistical analysis I used the combined model of Classification Trees (CART) and Random Forests which can account for multifactorial phenomena and complex interactions of variables (cf. Chapter 3, 5, and the present Chapter 9 Section 9.5.3.6.3). The result is shown in Figures 5 and 6.



Figure 5: Ctree for the corpus data.

Figure 6: Variable importance scale.

Figure 5 shows that the distribution of O- and U- is conditioned by a single factor. Figure 6 indicates that this factor is the adjectival class of the base (qualitative vs. relational, according to the tag assigned in the Russian National Corpus). Other factors are phonological and prosodic and they have no significant impact on the choice of the prefix (O- vs. U-) in deadjectival factitive verbs. In other words, the statistical analysis supports the claim of *the Scalarity Hypothesis*.

To sum up, I have argued that the factitive verbs in O- and U- make different semantic contributions. I discussed a number of empirical arguments that support my claim. The next section addresses the question of how the factitive uses of O- and U- are

related to their spatial meanings and how this relationship can explain their use in those factitive verbs that have not been captured yet in the proposed opposition of O- and U-.

9.4.3 Factitive use of O- and U- and their spatial meanings

I suggest that the contrast between O- and U- in factitive verbs is partially motivated by the different spatial prototypes of these prefixes, MOVE AROUND vs. MOVE AWAY respectively.

Factitive O- implies metaphorical “encircling” or “envelopment” of an object with a new property, while factitive U- indicates metaphorical “removal” of an object from the initial state into another state. This is a metaphorical “removal” of an object away from the reference point, the Landmark. The Landmark is the standard, or normal, degree of a property. The Trajector moves AWAY from it towards an increased degree of the same property. In this sense, the property scale encoded by the prefix U- is a path, and the Trajector moves along this path. In this sense, the Cognitive Linguistic approach to prefixal semantics provides additional support for my proposal. It enables me to establish the conceptual relation between the factitive uses of O- and U- on the one hand and their spatial meanings on the other hand.

I take a step further and compare the factitive uses of O- and U- with their uses in deverbal derivation. I suggest that the uses of O- and U- in factitive verbs are incorporated into the semantic networks of O- and U- established on the basis of deverbal verbs. In particular, I found that those factitive verbs that have not received an explanation for the choice of the prefix in *the Scalarity Hypothesis* can be accounted for from the perspective of prefix polysemy.

Figure 7 presents a radial category model of polysemy of the prefix U- established in Endresen et al. 2012 and Janda et al. 2013: 31, 58. Recall that the cognitive approach models the polysemy of a prefix as a radial network of interrelated meanings organized around a spatial prototype. We observe that many factitive verbs feature the same submeanings of U- as the ones attested in deverbal verbs.

The prototypical meaning 1.MOVE AWAY of the prefix U- is manifested in deverbal derivatives like *ubežat* ‘run away’ (< *bežat* ‘run’) and *ukrast* ‘steal’ (< *krast* ‘steal’) and in the factitive verb *uprazdnit* ‘eliminate’ (< *prazdnj* ‘indolent’).

Submeaning 2.MOVE DOWNWARDS is expressed by U- in deverbal verbs *ukatat* ‘make smooth by rolling’ (< *katat* ‘roll’) and *utrambovat* ‘press down in order to make smooth’ (< *trambovat* ‘press’) and the factitive verb *uplotnit* ‘condense’ from the adjective *plotnj* ‘dense’.

Submeaning 3.CONTROL of the prefix can be found in deverbal verbs like *uladit* ‘arrange’ (< *ladit* ‘get on’) and *uregulirovat* ‘regulate’ (< *regulirovat* ‘regulate’) and in the factitive verbs *usmirit* and *uspokoit* both meaning ‘calm down’ formed from the adjectives *smirnj* ‘quiet’ and *spokojnj* ‘calm’ respectively.

Submeaning 4.REDUCE is contributed by the prefix U- in deverbal verbs like *usoxnut* ‘dry out’ (< *soxnut* ‘dry’) and in the factitive verb *uničtožit* ‘turn into nothing’ formed on the basis of the pronoun *ničto* ‘nothing’.

Submeaning 5.HARM is encoded in the prefix U- in deverbal verbs like *ubit* ‘kill’ (< *bit* ‘beat’) and in the factitive verb *umertvit* ‘murder’ formed from the adjective *mertvj* ‘dead’.

Submeaning 6.PERCEIVE is attested for U- in deverbal verbs like *ugljadet* ‘spot with the eyes’ (< *gljadet* ‘look at’) and *uznat* ‘recognize, find out’ (< *znat* ‘know’) and in the factitive verb *usvoit* ‘grasp’ from the pronoun *svoj* ‘one’s own’.

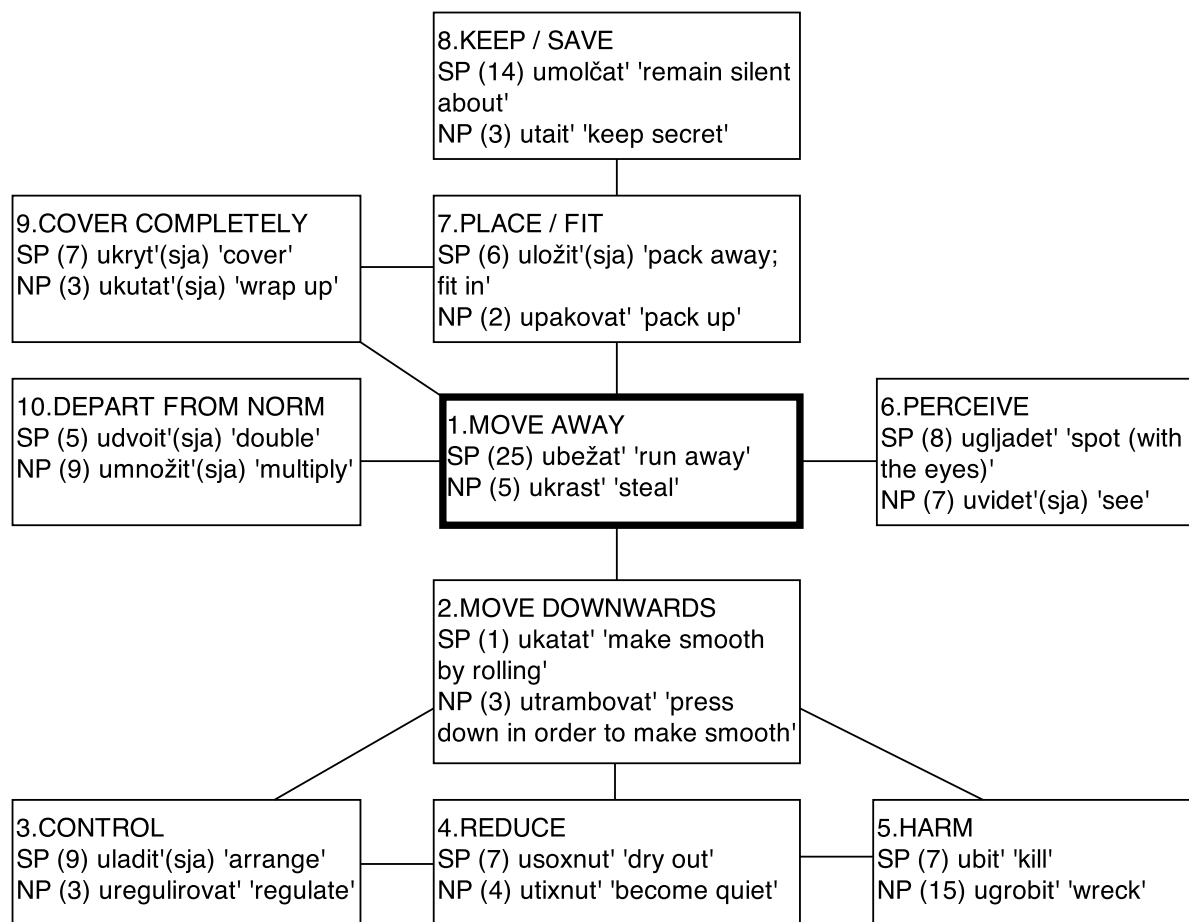


Figure 7: Radial category of meanings of the prefix U-.

Submeaning 9.COVER COMPLETELY of the prefix is identical in deverbal verbs like *usypat* 'cover by strewing' (< *sypat* 'strew') and in the factitive verb *ubelit* 'whiten' formed from the adjective *belyj* 'white'.

Submeaning 10.DEPART FROM NORM is the same in deverbal verbs like *učudit* 'act in an unexpected way' (< *čudit* 'act in an unexpected way') and in factitive verbs like *ukrupnit* 'enlarge' (< *krupnyj* 'large').

In addition, I suggest that the scalarity encoded in U- in factitive verbs is related to the spatial meanings of this prefix and especially to the vertical axis implied in the submeanings 2.MOVE DOWNWARDS, 4.REDUCE, and 3.CONTROL. The vertical axis motivates a change in the degree of the given property, which becomes "deeper" in the sense of being more entrenched and more intense. By contrast, the spatial prototypical meaning of O-, 1.MOVE AROUND, does not imply any vertical dimension. This spatial meaning motivates the spatially oriented meanings 11.AFFECT A SURFACE and 12.ENVELOP as well as the factitive meaning 15.IMPOSE A NEW FEATURE.

I have described the semantic network of the polysemous prefix O- in Chapter 5 and have addressed in detail the question of factitive verbs incorporated in the polysemy of this prefix. Recall that there are many factitive verbs in O- that manifest the same submeanings of this prefix as the ones attested for O- in deverbal verbs.

Submeaning 10.SURROUND of the prefix is found in both deverbal verbs like *očertit* 'draw around' (< *čertit* 'draw') and factitive verbs like *okajmit* 'decorate with edging around' (< *kajma* 'edging').

Submeaning 11.AFFECT A SURFACE is expressed by O- in deverbal verbs like *okleit'* 'cover with something glued' (< *kleit'* 'glue') and in factitive verbs like *ogolit'* 'make nude' (< *golyj'* 'nude').

Submeaning 12.ENVELOP of the prefix O- can be found in deverbal verbs like *okutat'* 'wrap up all over' (< *kutat'* 'wrap') and in factitive verbs like *očexlit'* 'put into a case' (< *čexol'* 'case').

Finally, submeaning 15.IMPOSE/ACQUIRE A NEW QUALITY is dominated by factitive verbs, many of which have a parallel verbal unprefixated base. For example, the verb *op'janit'* 'make intoxicated' is associated with both the adjective *p'janyj'* 'drunk' and the verbal simplex *p'janit'* 'make intoxicated' and thus represents both factitive and deverbal derivation at the same time.

To sum up, I have demonstrated that many uses of O- and U- in factitive verbs are semantically compatible with how these prefixes are used in deverbal verbs. I conclude that factitive verbs are in fact well incorporated in the overall semantic networks of these prefixes and in this sense belong to two distinct verbal aspectual morphemes O- and U- which have different spatial prototypes. Moreover, I suggest that the different semantic contributions encoded by O- and U- in factitive verbs are partially motivated by their different spatial prototypes, MOVE AROUND vs. MOVE AWAY, respectively. This finding supports *the Spatial Motivation Hypothesis* proposed in 9.3.5 which claims that the choice of the prefix in factitive verbs is motivated by idiosyncratic semantics of the prefix based on its spatial image schema.

In this light, the unified account of the factitive O- and U- in terms of Non-Standard allomorphs of a single morpheme-verbalizer is not appropriate. Rather, O- and U- are sub-uses of rival affixes that compete in formation of factitive verbs. The factitivizing function is a sub-use of the two distinct deverbal morphemes O- and U-, that is a peripheral domain of their application where the two morphemes overlap.

9.4.4 Conclusions for the corpus study

In response to the three research questions that I raised in the beginning of this chapter I make the following conclusions. First, I suggest that the choice of the prefix in factitive verbs is determined by the semantics of the prefix, in particular the spatial image schema of the prefix (cf. *the Spatial Motivation Hypothesis* in 9.3.5). Second, I define the status of these rival verbalizing prefixes as sub-uses of discrete morphemes which are peripheral with regard to deverbal derivation. I analyze the participation of these prefixes in factitive word-formation as a peripheral function of rival distinct morphemes. In this light, these prefixes are not allomorphs of a single verbalizing morpheme, because they have different spatial prototypes.

I propose that the factitive uses of the prefixes O- and U- belong to distinct morphemes that overlap in the function of verbalizer in the periphery of their semantics. In factitive verbs O- and U- have different semantic contributions which I account for in terms of *the Scalarity Hypothesis*. In addition, I suggest that it is possible to trace the grammaticalized factitive semantics of O- and U- back to their spatial prototypes, MOVE AROUND vs. MOVE AWAY respectively. Therefore, I conclude that a unified account of the two prefixes in terms of allomorphs of a single morpheme is problematic. Rather, within the verbalizing morphological construction the two prefixes are rival candidates that belong to distinct morphemes.

In factitives, O- is the default choice and refers to simple imposition of a new quality on an object, while U- is semantically more marked and refers to scalar changes.

In standard factitive verbs, O- is less restrictive than U- in terms of types of possible bases, and O- is more productive than U- in novel coinages.

9.5 The experimental study

In this section, I report on an experimental study¹⁵⁵, which explored the use of the two Russian prefixes O- or U- in novel marginal words. The experiment elicited acceptability judgements of native speakers regarding standard, marginal, and nonce factitive verbs formed by the two prefixes. I will first present three research questions that were addressed in the experiment (9.5.1), then describe the experimental design (9.5.2), and turn to the analysis of experimental results via different statistical techniques (9.5.3).

9.5.1 Research questions

9.5.1.1 O- vs. U-

Recall that the corpus study showed that marginal factitive verbs prefixed in O- are two times more numerous than those prefixed in U-. This gives support to the conclusion that in factitive verbs the O-pattern (e.g. *opoxabit* 'profane') is more productive than the U-pattern (e.g. *usovremenit* 'modernize'). In addition, the semantic analysis suggests that the O-pattern is conceptually the simplest of the two rival alternatives and thus should be considered the default.

If so, one might expect that those marginal factitives that are formed by the O-pattern should overall be more acceptable to native speakers than the factitives formed via the U-pattern which is marked and semantically more particular. In this experiment we wanted to find out whether the two derivational patterns are significantly different with regard to their relative naturalness to Russian speakers. We can thus formulate the following research question:

Does the type of the prefix (O- vs. U-) make a difference in how marginal words are perceived by native speakers? If so, does the productivity of the prefix play any role? In particular, does the more productive prefix (O-) form more acceptable novel marginal verbs than the less productive one (U-)?

The hypothesis behind this question was the following:

Novel marginal verbs formed by the prefix O-, the default and relatively more productive verbalizing prefix as opposed to U-, should be judged as more natural and acceptable than marginal verbs in U-.

In other words, we expected the speakers to assign higher acceptability scores to marginal verbs prefixed in O- and lower scores – to derivatives in U-.

¹⁵⁵ This experimental study is a joint project done in collaboration with Laura A. Janda. The results were reported at the International Cognitive Linguistic Conference (ICLC-2013) hold on 23.06.-28.06.2013 at the University of Alberta, Canada; and at the 13th Annual Conference of the Slavic Cognitive Linguistics Association (SCLC-2014) 1.02.-17.02.2014 at Harvard University, USA. We are grateful to both audiences for fruitful discussions.

9.5.1.2 Novel marginal vs. Standard vs. Nonce verbs

We explored the relative degree of naturalness, or acceptability, of the two factitive patterns by exposing subjects to stimuli of three different kinds: standard words, marginal words, and nonce words. These three classes of words form a scale in terms of token frequency of attestation (from thousands of attestations of standard words to zero attestations of nonce words), recognizability, and acceptability.

Marginal verbs fill the gap between the actual and the impossible in a language. On the one hand, marginal verbs exist because they are attested – they have been generated and uttered, recorded, or written down. On the other hand, marginal verbs do not exist, because most speakers might have never heard them. In this light, marginal verbs can be placed in between standard words and nonce words. Under standard words we understand those that are conventionalized: standard words are familiar to the entire language community or, at least, to a large part of it. The other extreme is nonce words – words that do not exist. Although they might conform to the phonological rules of a language, they are not associated with any meaning.

Marginal words lie in between these two categories. They have been created by a speaker or some speakers and can be understood by other speakers. Normally, marginal verbs are structurally analyzable and semantically transparent, primarily because they are generated on the basis of a productive morphological pattern. This is crucially important for the present study, where we examine the two most productive derivational patterns in the formation of factitive verbs – the morphological constructions in O- and U-. We assume that an average Russian speaker can easily identify the meaning of such marginal factitive verbs like *oser'eznit'* 'make serious' because of their transparent semantic compositionality. In order to deduce the meaning 'make something serious' for the marginal verb *oser'eznit'*, it is enough to know the adjective *ser'eznyj* 'serious' on the one hand and, on the other hand, standard verbs like *obogatit'* 'enrich', *okruglit'* 'round, express in round numbers', *osvežit'* 'refresh' associated with the base adjectives *bogatyj* 'rich', *kruglyj* 'round', and *svežij* 'fresh'.

By definition, marginal verbs are not established in the standard lexicon. Rather, such verbs constitute spontaneous creations generated by speakers on the fly, on a certain occasion. Usually, marginal verbs are attested at least once in the corpus or elsewhere. For the purposes of this study we selected those marginal factitive verbs that have a minimal number of attestations in the Russian National Corpus – from one to eight occurrences (see the details of the experimental design in Section 9.5.2).

Note that marginal verbs used in the experiment are characterized with marginal status due to their novelty, and not because they are obsolete. Therefore, these marginal verbs reflect the productivity of two derivational patterns with O- and U-. In this light, novel marginal verbs are similar to what is sometimes called "possible words" (Bauer 2001, 2012), "new coinages", and "potential words" (Aronoff 1983; "*potencial'nye slova*" in the Russian linguistic tradition cf. Zemskaja 1972).

Because marginal verbs share some properties with each of the two other classes, the second research question that we addressed was:

Are marginal factitive verbs of the two rival patterns (O- and U-) perceived more like standard verbs or more like nonce verbs?

Regarding the perceptability of marginal factitive verbs, we formulated four hypotheses that were tested in the experiment. These hypotheses can be schematically presented as

shown in Figure 8: either marginal and standard verbs are perceived the same way (Hypothesis 1), or marginal and nonce words are perceived the same way (Hypothesis 2), or each of these three categories is perceived differently (Hypothesis 3), or they all form one continuum (Hypothesis 4).

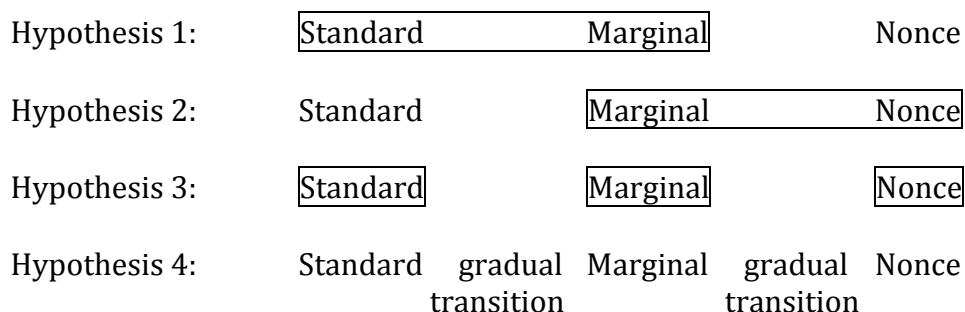


Figure 8: Four hypotheses tested in the experimental study.

Within each of these four hypotheses we looked for statistically robust differences in perception of marginal factitive verbs prefixed in O- and in U-.

9.5.1.3 Children vs. Adults

The third question that we addressed in the experiment is whether the speakers' leniency regarding marginal verbs of these two patterns (O- and U-) changes with age. In particular, we were interested in two age groups of speakers – school age children and adults. We expected that children would demonstrate more lenient acceptability judgements because their vocabulary and linguistic standards undergo a process of development and shaping. Teenagers might be more liberal and open to unfamiliar words than adults, whose linguistic standards and preferences have already stabilized. In this regard, adults are expected to give more conservative judgments and be generally less willing to accept marginal words.

We take the age of twenty-five as an approximate threshold for adulthood, because by this age most adults in Russia complete their education, enter the job market, and also outgrow colloquialisms typical for youth. On the other hand, we are particularly interested in fourteen-to-seventeen-year-olds who are usually at the peak of implementing their youth slang and are arguably very open to linguistic innovations.

9.5.1.4 A note on perspective: Competence & Performance

In the experiment, we tested the actual acceptability of novel marginal factitive verbs found in the corpus. Therefore, the objective of this study was to combine the crucial perspectives – speakers' competence and performance – which are usually thought of as conceptually distinct facets of language and are usually examined separately from each other. As Haspelmath (2002: 112) pointed out, newly coined words and rule productivity are mostly studied from the perspective of linguistic performance. However, productivity of derivational patterns and probabilities in grammar refer not only to linguistic performance but also to competence. In this sense, these phenomena pose a serious problem for the rigid distinction of competence and performance that we find in the generative tradition.

In this study, we combined these two perspectives: from the perspective of performance, we looked at novel marginal factitive verbs generated by native speakers and attested in a corpus. From the perspective of competence, we wanted to find out how these novel marginal words are perceived by other native speakers.

9.5.2 Experimental design: Elicitation of acceptability judgements

The experiment was designed as a score-assignment test. Each subject was presented with a total of sixty sentences and a rating system. Each sentence contained an underlined factitive verb, as illustrated in (13):

- (13) *Davno pora kak-to opriličit' naše obščenie bolee mjagkimi vyraženiĭami.*
'It's high time we made our interaction respectable by using gentler expressions.'

The task was to evaluate the factitive verb according to a scale of acceptability judgements¹⁵⁶ accompanied with scores from one to five points. The scores were given in parallel to the following evaluative statements:

- 5 points – *Èto soveršenno normal'noe slovo russkogo jazyka.*
'This is an absolutely normal Russian word.'
- 4 points – *Èto slovo normal'noe, no ego malo ispol'zujut.*
'This word is normal, but it is rarely used.'
- 3 points – *Èto slovo zvučit stranno, no, možet byt', ego kto-to ispol'zuet.*
'This word sounds strange, but someone might use it.'
- 2 points – *Èto slovo zvučit stranno, i ego vrjad li kto-to ispol'zuet.*
'This word sounds strange and it is unlikely that anyone uses it.'
- 1 point – *Ètogo slova v russkom jazyke net.*
'This word does not exist in the Russian language.'

Thus, subjects had a multiple-choice task and had to choose the combination of a statement and a score which described best their intuition about an underlined verb. Keeping in mind the purposes of this study and the cultural background of participants, we can name several arguments in favor of this methodology.

First, the rating scale was subdivided into five points because the five level system is very familiar for all Russian speakers due to the grading scale in schools which also consists of five grades, where grade «1» corresponds to the worse performance, while the top grade «5» corresponds to the best performance. This gradation is reflected in our scale of points from 1 to 5, where one point should be assigned to a word that does not exist in Russian, and the highest score of five points should be assigned to a perfectly normal Russian word. This also explains why we prefer to use the five point system of gradation instead of common three-point (Collins et al. 2009) and seven-point scales (Bermel & Knittle 2012a,b). Additionally, because of the school evaluation system

¹⁵⁶ In using the term “acceptability judgements” instead of “grammaticality judgements” we follow the careful argumentation provided in Bermel & Knittle 2012a,b.

we used the scale from 5 to 1 instead of other options like +2, +1, 0, -1, -2. On the other hand, the statement ‘This word sounds strange, but someone might use it’ corresponds best to the score «3», being half way from 1 to 5, rather than to the score «0», which suggests certain neutrality and does not comply with our purposes. It is worth mentioning that we chose the descending scale from 5 down to 1 over the ascending alternative (from 1 to 5) because the former corresponds to the iconic gradation “the higher (spatially) – the better”.

Second, instead of exposing subjects to a Likert type scale which gives explicit descriptions like “perfectly normal” and “unacceptable” only to the top and the bottom ends of the scale (Bermel & Knittle 2012a,b; Dąbrowska 2010), we assign an evaluative statement to each of the five scores. We suggest that the combination of numeric scores and descriptive judgements makes it possible to have better control over subjects’ intuitions. Otherwise, subjects would have to create their own descriptions of the five scores, and this would remain opaque in terms of outcome. The forced-choice system that we used provided a uniform set of descriptions that each subject relied on. In our experiment this methodology helped to shield results from unwanted additional opaque variables and to collect more robust data. The way the evaluative statements are formulated is very important in this case as well. Note that we use an impersonal construction in *èto slovo malo ispol’zujut* and the indefinite pronoun in *možet byt’, ego kto-to ispol’zuet*. This wording was meant to invite subjects to think generally, having in mind the whole language community. Hopefully, this wording helped subjects not to bias themselves by thinking of particular individuals or a single category of people like acquaintances, friends, family members, or colleagues who might or might not use a stimulus word. Lastly, the evaluative statements are formulated in such a way that they maximally correspond to an ordinal scale with approximately comparable intervals between each two statements.

Third, for this study we employed a five score scale instead of the method of magnitude estimation (Bard et al. 1996, Cowart 1997, Sorace & Keller 2005) because, as it was argued in Dąbrowska (2010), using a scale is much more natural, intuitive, and easier for subjects. In our study this was important especially for school-aged participants.

The stimuli were presented in a semi-random order and the order was the same for all participants. The first two warm-up sentences contained standard factitive verbs, while the third sentence already introduced a marginal verb. We made sure that in each questionnaire there was no sequence of more than two adjacent sentences that introduced the same prefix or the same type of stimulus. This was done in order to prevent subjects from developing a uniform strategy that could bias their judgements.

9.5.2.1 Stimuli

In order to test all four hypotheses with regard to the two derivational patterns of factitive verbs, we designed the experiment so that we could collect speakers’ responses triggered by all three relevant categories of words. Each questionnaire exposed the subjects to three groups of stimuli – standard, marginal, and nonce factitive verbs prefixed in O- and U-. In order to limit a questionnaire to a manageable size, we used twenty stimuli in each group, where ten verbs were prefixed in O-, and the other ten verbs – in U-. The prefix and the word-type conditions yield the total of sixty stimuli, where standard and nonce factitives were two groups of controls and distractors, whereas the twenty marginal verbs were the tested experimental items.

There is at least one more reason for why it was important to counterbalance the set of marginal verbs with other kinds of verbs. As Schütze (1996: 184) fairly points out, a survey contains a bias “if there are substantially more grammatical sentences in the test materials than ungrammatical sentences or vice versa”. The unwanted possibility here is that subjects can “get into a yea-saying or nay-saying mode or will come to expect deviance”. In order to avoid this effect, the numbers of stimuli for different conditions should be “roughly equal”. In our study, we provided the subjects with two opposite extremes that the marginal verbs could be compared to – they were expected to be less acceptable than standard verbs but more acceptable than nonce verbs.

For this study we chose those marginal factitive verbs that were all morphologically transparent and semantically analyzable, like standard verbs. The marginal verbs differed from the standard verbs only because the former were not conventionalized and therefore mostly unfamiliar to an average speaker and could sound odd. On the other hand, being odd made them close to nonce verbs, the other extreme of the scale. However, the marginal verbs still made sense and had a familiar base and word-formation pattern, while nonce words, by contrast, could not be associated with any existing simplex bases.

In order to exclude other possible variables from the experimental conditions, all standard and marginal factitives chosen for the experiment were deadjectival and had a clear adjectival base. None of the factitive verbs had a parallel simplex verbal base: e.g. *ob”jasnit’* ‘clarify’ < *jasnyj* ‘clear’, not **jasnit’* ‘make clear’. Nonce verbs were shaped like factitives by means of verbal morphology: they contained the same prefixes O- and U- and the verbalizing suffix *-i-*, but no recognizable root.

Moreover, the mode of presentation of all three types of stimuli was made uniform in terms of context. All stimuli were presented as perfective infinitives in a sentence which was borrowed or based upon a real sentence attested in the Russian National Corpus. We made sure that the contexts chosen for standard and marginal factitive verbs were typical, neutral in register, and maximally supported the factitive meaning of the verb. The contexts for standard and marginal factitive verbs were directly extracted from the corpus and were shortened in some cases. In a few cases a better context for a marginal factitive was found via the search engines www.yandex.ru and www.google.ru. The contexts of nonce factitive verbs were created so that they would support the change-of-state meaning of the nonce verb. All contexts of verbal stimuli used in experiment are listed in Appendix 6 and translated into English.

Table 8 presents a list of the standard factitives used in the experiment.

O-factitive	Gloss	Freq	U-factitive	Gloss	Freq
<i>ob”jasnit’</i>	clarify	18,149	<i>utočnit’</i>	define more precisely	2,860
<i>oblegčit’</i>	simplify, lighten	1,802	<i>umen’šit’</i>	reduce	2,010
<i>oslabit’</i>	weaken, loosen	1,401	<i>uskorit’</i>	speed up	2,008
<i>okruglit’</i>	express in round numbers	939	<i>ulučšit’</i>	improve	1,899
<i>obogatit’</i>	enrich	800	<i>uprostit’</i>	simplify	1,350
<i>ožestočit’</i>	harden, obdurate	686	<i>ukorotit’</i>	make shorter	787
<i>osložnit’</i>	complicate	410	<i>usložnit’</i>	complicate	311
<i>ogolit’</i>	denude	387	<i>uteplit’</i>	make warmer	205
<i>osčastlivit’</i>	make happy	343	<i>uplotnit’</i>	compress	201
<i>osvežit’</i>	freshen	280	<i>uxudšit’</i>	make worse	199

Table 8: Standard factitive verbs used in experiment (control group 1).

All verbs in Table 8 have high token frequencies in the corpus. The verbs are given in descending order of their token frequencies. The frequencies here are overall numbers of all attestations found in the Modern Subcorpus of the RNC, which includes the texts created in 1950-2012.

Table 9 provides a list of all marginal verbal stimuli employed in the experiment. When choosing these verbs we used two criteria – minimal token frequency in the corpus and transparency of the word’s derivational structure, in particular a clear semantic and structural association link with a base that speakers can easily rely on. In Table 9 the verbs are listed in increasing order of token frequencies – from one to eight corpus attestations.

O-factitive	Gloss	Freq	U-factitive	Gloss	Freq
<i>omeždunarodit'</i>	internationalize	1	<i>uvkusnit'</i>	make tastier	1
<i>opoxabit'</i>	profane, pollute	1	<i>umedlit'</i>	make slower	1
<i>opriličit'</i>	make decent	1	<i>ukrasivit'</i>	make prettier	1
<i>oser'ėznit'</i>	make serious	1	<i>user'ėznit'</i>	make more serious	1
<i>ostekljanit'</i>	make glassy	1	<i>ukonkretit'</i>	make more concrete	1
<i>oržavit'</i>	corrode	2	<i>usovremenit'</i>	make more modern	1
<i>osurovit'</i>	make rigorous	2	<i>ustrožit'</i>	make stricter	3
<i>obytovit'</i>	vulgarize	3	<i>ucelomudrit'</i>	make more innocent	3
<i>ovnešnit'</i>	externalize	4	<i>uprozračit'</i>	make more transparent	4
<i>omuzykalit'</i>	musicalize	4	<i>udorožit'</i>	make more expensive	8

Table 9: Marginal factitive verbs used in experiment (tested group).

Table 10 provides a list of all nonce factitive verbs used in the experiment. These nonce verbs were adopted from the previously conducted psycholinguistic experiment described in Chapter 5, as well as in Baydimirova 2010, Endresen 2013, and Baayen et al. 2013.

O-factitive	U-factitive
<i>osurit'</i>	<i>usaglit'</i>
<i>otovit'</i>	<i>utulit'</i>
<i>oduktit'</i>	<i>udamlit'</i>
<i>ogabit'</i>	<i>uguzvit'</i>
<i>okočlit'</i>	<i>ukampit'</i>
<i>ošaklit'</i>	<i>ušadrit'</i>
<i>očavit'</i>	<i>učopit'</i>
<i>oblusit'</i>	<i>uloprit'</i>
<i>obnomit'</i>	<i>unokrit'</i>
<i>obmomlit'</i>	<i>umarvit'</i>

These nonce verbs were created manually. They satisfy well-formedness constraints of Russian phonotactics and sound native-like to an average speaker.

Table 10 demonstrates that each nonce factitive in O- had a parallel nonce factitive in U- which contains the same consonant of the base (*s, t, d, g*, etc.) but the base itself is not identical: e.g. *osurit'* and *usaglit'*, *otovit'* and *utulit'*, etc. This is done in order to balance the set of nonce stimuli.

Table 10: Nonce factitive verbs used in experiment (control group 2).

9.5.2.2 Administration

The experiment was administered as a questionnaire with no limits on time. An average time for completion of a questionnaire was twenty minutes.

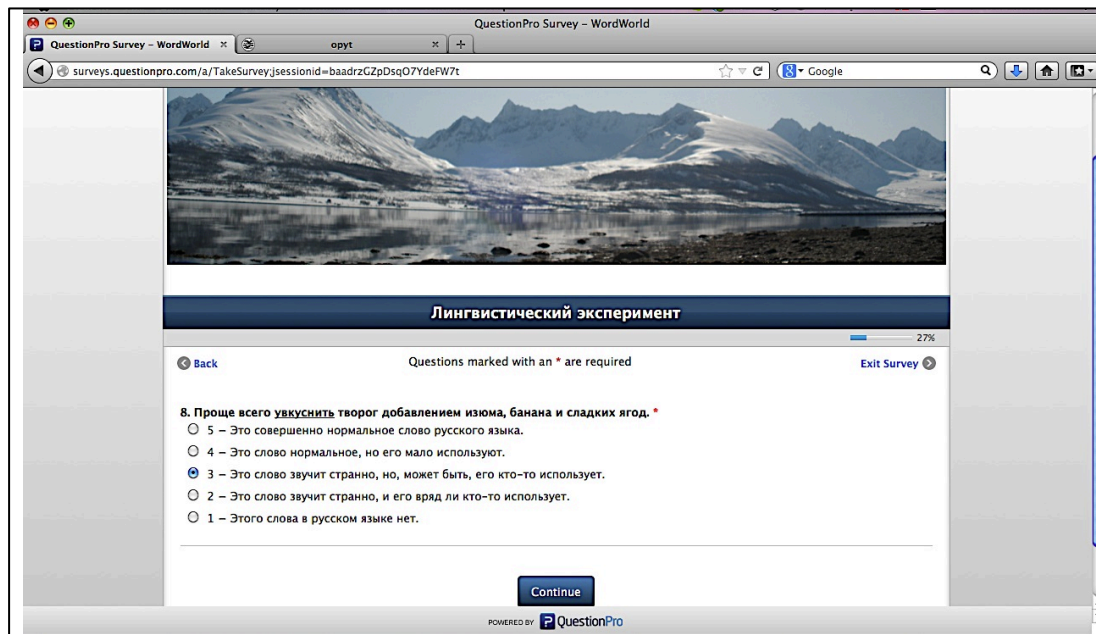
For children administration consisted in filling out a hard copy of a questionnaire form and was conducted in the school setting. Adults completed the survey over the internet, where they had to fill out a virtual questionnaire created in the software package <http://www.questionpro.com> (package “Professional”). The use of an online questionnaire form easily shared via internet is a common practice used in many recent

surveys of acceptability judgements (Keller & Asudeh 2001; Collins et al. 2009). The software made it possible to keep track of the data and make sure that people who participated online took the survey only once.

The stimuli sentences used in the experiment are presented in Appendix 6.

The introduction to the experiment collected sociolinguistic information about subjects' gender, age, level of education, area of expertise, and place of residence. This part was followed by instructions about the task, the list of scores and statements, and an illustrative example with a standard factitive verb. The next section told the subjects that they will be exposed to both existing and non-existing words, that they will have to evaluate sixty words, and that the tasks contain no typos. This part informed the subjects that they should not worry about incorrect responses, because the task is not about spelling competence but rather about speakers' linguistic intuition.

In the online questionnaire, the sentences were presented one at a time, one task per page. The scale of five scores accompanied with statements was given after each sentence, as shown below:



In the hard copy questionnaire, the sentences were presented five sentences per page. The scale of five scores was given after each sentence.

9.5.2.3 Subjects

The experiment was conducted over the course of two weeks in April, 2013. All subjects were native speakers of Russian who grew up, received their education, and currently live in Russia. We recruited the total of 121 participants including seventy children and fifty-one adults. This number excludes subjects who currently reside outside Russia, do not belong to the two age groups we are interested in, or failed to provide correct responses to control stimuli. In the three forthcoming subsections I will explain in detail the various sociolinguistic parameters that are commonly accounted for in psycholinguistic experiments of this kind.

9.5.2.3.1 Age groups and gender

We focused on two age groups of subjects: 1) middle and high school age children and 2) adults.

We tested seventy-five children of fourteen to seventeen years of age. At the time of participation in the experiment, the children attended the 7th and 9th grades. Only five children (6% of all children) failed to perform the task. It is hard to say whether they misunderstood the task or deliberately wanted to sabotage the assignment. They incorrectly evaluated about 50% of the items in the two control groups of stimuli (standard and nonce words). These children include one 14-year-old girl, two 15-year-old boys, and two 17-year-old boys. Their responses were excluded from the analyzed data.

Table 11 shows how the school-aged participants were distributed across age and gender. We account for acceptability judgments submitted by thirty-one boys and thirty-nine girls, yielding a total of seventy school-aged participants.

Age	Male	Female	Total
14	7	13	20
15	11	10	21
16	6	10	16
17	7	6	13
Total	31	39	70

Table 11: Distribution of school-aged participants across age and gender.

In order to test whether the acceptability of marginal verbs changes with age between middle school age and the period of adult professional activity, we left aside the responses of those subjects that lie between these two categories – namely those who are older than 17 but younger than 25 years. Thus, responses of fourteen subjects were excluded from our analysis.

We recruited fifty-one subjects who were 25-to-62 years old including sixteen men and thirty-five women. Most of these subjects (forty-five individuals, or 88 % of all adult subjects) were 25-31 years of age. Table 12 demonstrates how adult subjects were distributed across age and gender.

Age	Male	Female	Total				
25	1	7	8	31	0	2	2
26	3	12	15	33	1	0	1
27	2	8	10	39	1	1	2
28	4	2	6	40	0	1	1
29	2	0	2	48	0	1	1
30	1	1	2	62	1	0	1
				Total	16	35	51

Table 12: Distribution of adult participants across age and gender.

Overall, our results account for 121 subjects including seventy children and fifty-one adults. Among them, there are 47 males and 74 females. Figure 9 and Figure 10 show the percentage ratio of these two categories and demonstrate that in both groups the female subjects predominated.

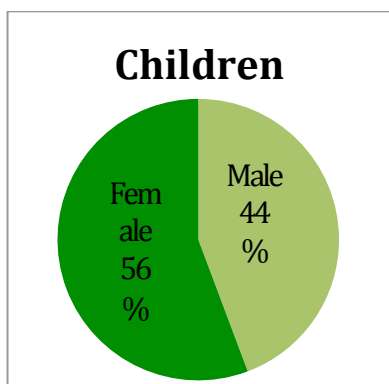


Figure 9: Distribution of school-aged subjects across gender.

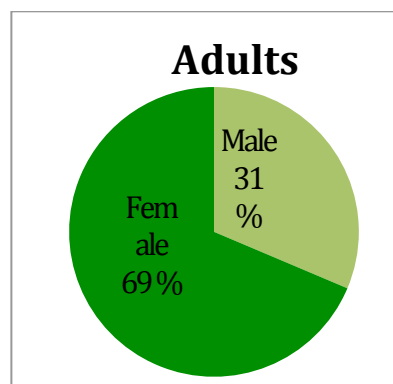


Figure 10: Distribution of adult subjects across gender.

9.5.2.3.2 Place of residence

All school-age participants live in the Russian city of Iževsk, where they attend school. Adult participants are residents of a large variety of Russian cities including Moscow, St. Petersburg, Iževsk, Samara, Nižnij Novgorod, Perm', Ufa, Tver', Smolensk, Kandalakša, Belgorod, and Blagoveščensk.

In addition, among adult subjects there were also twenty-one who participated in the survey from abroad because they currently live outside Russia. In particular, there were some Russian speakers from other Slavic countries like Belarus, Ukraine, Serbia, and Montenegro. Others participated while living in United States, Europe, or Scandinavia. We decided to exclude the responses of these participants from our analysis. This solution is motivated by the need to exclude any additional variables related to linguistic changes that speakers might undergo as a result of living abroad. The latter circumstance includes a number of sociocultural factors that are very difficult to control for and that can bias the results. These variables include the length and intensity of speakers' exposure to foreign languages while living abroad, the possibility of partial language attrition, as well as changes in lifestyle that might affect speakers' linguistic competence and preferences.

9.5.2.3.3 Profession

Among adult participants there were people of a wide variety of professions and areas of expertise. Adults recruited for this survey were employed in a variety of sectors such as telecommunication, mass media, computer maintenance, music, journalism, tourism, pharmacy, business, theater, cinema, ecology, and sport. Adult subjects included engineers, writers, doctors, lawyers, foreign language teachers, economists, psychologists, programmers, experts in nuclear security, locomotive engineers, and philologists. This variety of professions was welcome in this study in order to compile a representative sample of speakers. For the same reason, we did not exclude those subjects whose educational background was related to linguistics. Such subjects were very few and their responses were in line with the remainder of collected data.

9.5.3 Experimental results

The ultimate goal of the experiment was to test whether the acceptability scores assigned by speakers to verbal stimuli correlate with any of three factor variables: prefix, age, and word type. In particular, we focused on the following questions:

- Do scores of acceptability correlate with the prefix O- vs. U- (given that according to corpus data O- is more productive than U- in production of novel factitive verbs)?
- Do acceptability scores correlate with speakers' age (children 14-17 year old vs. adults 25-62 year old)?
- Do scores of acceptability correlate with the status of a verb and its token frequency in the RNC (standard verb with high token frequency vs. marginal verb with minimal token frequency vs. nonce verb with zero attestations)?

In order to find out what the statistically robust correlations are, we conducted several statistical analyses. The effect of the first two variables – prefix and age – turned out to be statistically significant at local domains (portions of data) but not in the data overall. The remaining variable – type of stimulus (standard vs. marginal vs. nonce) – was found highly significant. However, before I address each factor in turn, I will first discuss what type of data we deal with and what type of statistical analysis is appropriate for it.

9.5.3.1 A note on the level of measurement appropriate for collected data

The ambition of our methodology used in elicitation of acceptability judgements was to capture the intrinsic gradient nature of linguistic intuition. In our rating system implemented in the experiment we combined two types of scales: descriptive evaluative judgements provided a qualitative (categorical) scale, while a set of parallel scores from five to one formed a quantitative scale. In this subsection I address a standard objection typically used in discussion of this kind of experimental design especially when the data is subjected to a statistical analysis. This discussion is grounded in the key question of what kind of data we collected in the experiment and what kind of statistical tests treat this data appropriately.

The levels of data measurement are commonly classified in terms of four types of scales – nominal, ordinal, interval, and ratio. I will first outline the key properties of each type of scale and then turn to the problems with Likert-type rating scales and related experimental designs like ours.

The four types of scales named above are distinguished according to what kind of information they provide about data. A nominal scale contains categorical, or qualitative, data variables; it does not convey any quantitative information and does not imply any ordering of items (Cantos Gómez 2013: 37). In a nominal scale, all values have equal status and cannot be arranged in any particular order. Typical examples are yes and no alternatives, colors, genders, and types of nationality.

As opposed to nominal scales, ordinal ones not only categorize items but also reflect their ranking, or ordering. Ordinal scales establish a hierarchy of categories. They can either use descriptive statements or implement numeric variables in order to estimate a degree of a certain feature or characteristic. For example, a scale of words “liked slightly”–“liked moderately”–“liked a lot” or the numbers 1-2-3 can be used in movie ratings. Note, that in ordinal scales it is not certain whether the distances

between adjacent points of the scale are of equal magnitude. Returning to our example, the distance between “liked slightly” and “liked moderately” might be not the same as the distance between “liked moderately” and “liked a lot”.

By contrast, interval scales rank values that have equal intervals within each pair of adjacent units. Classical examples of interval scales are measures of length like centimeters or measures of temperature like degrees Celsius (Cantos Gómez 2013: 36). Importantly, the values that form an interval scale allow arithmetic calculations that compare their sizes, averages, and variation.

Lastly, a ratio scale has the key properties of interval scales but additionally has an interpretable and natural zero starting point, like in case of money or weight (Cantos Gómez 2013: 37).

Recall that a Likert-type scale contains a ranked set of points (usually seven), where the top and the bottom ends of the scale are descriptively categorized, for example, as “perfectly normal” and “unacceptable”.

There is a controversy in the literature as to whether Likert-derived data constitute an ordinal or an interval scale. This issue is important because the statistical techniques used for interval variables are not appropriate for ordinal variables. In particular, interval data can be subjected to parametric tests (like calculation of mean and variance), while ordinal data can be only explored via non-parametric tests like the chi-squared test (Cantos Gómez 2013: 236; Cohen et al. 2000: 317). The use of the wrong statistical test arguably leads to wrong conclusions about data.

Strictly speaking, the intervals between values on a Likert scale are not necessarily equal, but many researchers assume that they are. Cohen et al. (2000: 317) and Jamieson (2004) object against assuming an interval scale for Likert-type categories, and therefore they find it illegitimate to use parametric statistics for data obtained via Likert scales. Yet, Jamieson observes that in medical and social sciences it has become “a common practice to assume that Likert-type categories constitute interval-level measurements” (ibid: 1212). Similarly, Strobl et al. (2009: 323) mention the fact that “ordinally scaled variables, which are particularly common in psychological applications, are often treated as if they were measured on an interval or ratio scale”. Furthermore, *Dąbrowska (2010: 8) points out that a number of studies (Jaccard & Wan 1996; Labovits 1967, Kim 1975) have argued that “parametric tests are quite robust, so that violations of the intervalness assumption have relatively little impact on the results of the test”*. Moreover, *Dąbrowska 2010 states that “the use of parametric tests with data obtained using Likert scales has now become standard”* (cf. similar observations in *Blaikie 2003, Pell 2005*). *Dąbrowska (2010) herself uses a Likert-type scale in elicitation experiments and analyzes the responses with ANOVA and t-tests. Similarly, Bermel and Knittle (2012a,b) conduct an experiment using a Likert scale and explore their results with ANOVA statistics.*

The scale used in our experiment is different from the one used in *Dąbrowska 2010 and Bermel & Knittle 2012a* because we named not only the top and the bottom ends of the scale but also the three midpoints. Recall that in doing so we wanted to ensure a uniform interpretation of scale points across all subjects (cf. Section 9.5.2.1). Again, in the literature we find a controversy regarding such scales. Bermel & Knittle (2012a: 244) express a general concern that a scale with named midpoints would constitute a clear instance of non-interval data. On the other hand, a number of published studies in experimental syntax that make use of three-point scales for elicitation of grammaticality judgements find it appropriate to transfer such data into scores. Note that this practice takes place even when the options of grammaticality

judgements that subjects are presented with are not enumerated (Collins et al. 2009: 4). Sprouse (2007: 67) says that “unfortunately, there is no generally accepted non-parametric version of factorial ANOVA, and in fact, standard ANOVA is often reported for ordinal data in the psychological literature.” Crucially, Cowart (1997: 120) observes that “where a particular contrast is numerically large compared with variability around the relevant means, any statistical problems deriving from a failure to achieve interval level measurement are not likely to be consequential.”

Recall that in our experiment we used evaluative judgements aligned with the numeric scale of five points which is commonly used in Russian school grades. This scale implies an ordering of items and clearly contains quantitative information. This scale thus comprises ordinal values but not necessarily interval values. Yet, in our view, the scale of numeric points 5-4-3-2-1 supports the idea of approximately equal intervals between the midpoints. Moreover, we suggest that this scale contains a clear interpretable zero point, namely the score “1”: it corresponds to the descriptive statement “This word does not exist in the Russian language”. This means that if a stimulus receives the score of one point, it is evaluated by a subject as a non-word. Being a non-word, or being non-existent, is a clear interpretable zero point on the scale of acceptability scores. The fact that the scale employed in the experiment contains an interpretable zero starting point suggests the possibility that this scale might be at least “more than ordinal”, if not interval¹⁵⁷. In the context of the long-standing and continuing debate in the literature¹⁵⁸, we find it not entirely wrong to 1) translate the acceptability scores into numeric values and 2) apply parametric statistics like analysis of variance (ANOVA) to this data. However, because this strategy can be easily criticized, we additionally run other parametric and non-parametric statistical analyses that are specifically designed for handling ordinal data. We use three different models – Ordinal Logistic Regression, Regression Mixed-Effects Model for Ordinal Data, and a combined model of Classification Trees and Random Forests. Findings from both parametric and non-parametric statistical techniques contribute similar and complementary insights about our data.

9.5.3.2 Overview: Central tendencies in data distribution

Figures 11-14 plot the distribution of acceptability scores across four different factors – Gender (Figure 11), AgeGroup of subjects (Figure 12), Prefixes (Figure 13) and Word type categories (Figure 14). Gender is included only for the sake of comparison.

In each plot, the data is visualized in the shape of a rectangle, where the thick line indicates the MEDIAN score. A median is conventionally understood as the middle point, or the central score, of the distribution, “with half of the scores lying above and half of the scores falling below” (Cantos Gómez 2013: 3).

¹⁵⁷ Knapp (1990: 121) points out that the distinction between ordinal and interval scales of data measurement is often a challenge when one has to categorize a specific data set. Moreover, Knapp (ibid) suggests that a particular scale can be “ordinal, less than ordinal, or more than ordinal”, and that there are no agreed-upon rules for determining this.

¹⁵⁸ Ordinal-level variables are generally considered challenging for statistics. The ordinal/interval scale-and-statistics controversy is a long-standing and continuing debate in the literature. For the history of conflicting views see Gardner 1975; a brief summary is given in Knapp 1990.

These box-and-whiskers plots neatly visualize the central tendencies of data distribution. Comparing these plots, we observe the core differences in the overall impact of four factors.

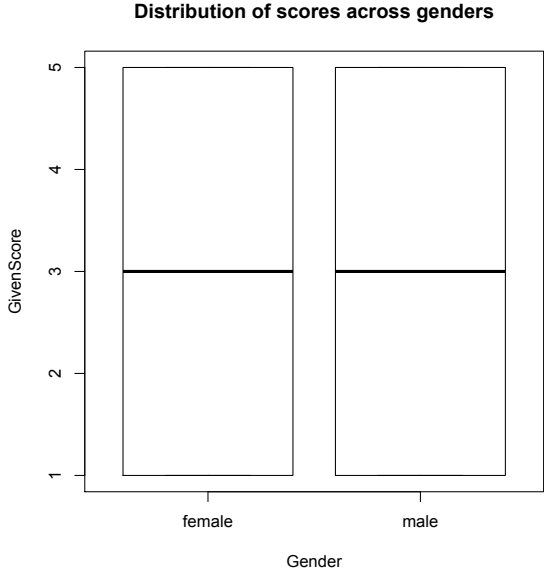


Figure 11: Impact of Gender.

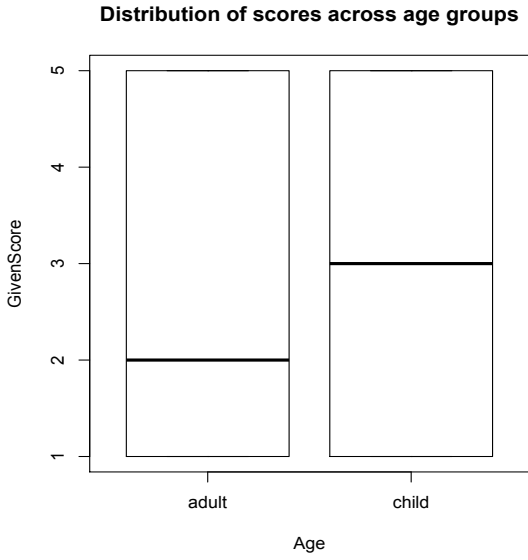


Figure 12: Impact of Age.

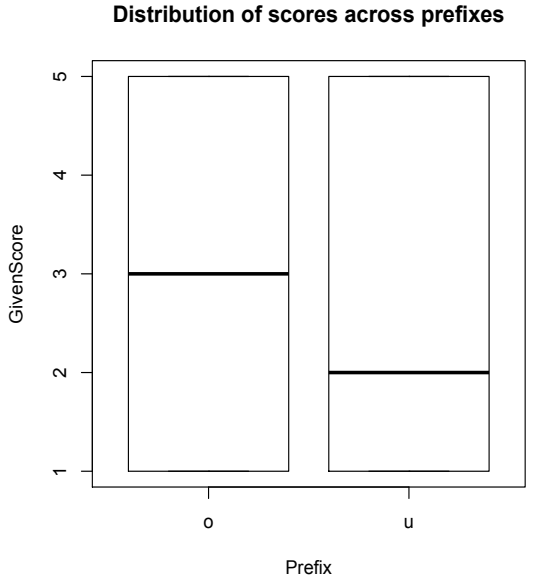


Figure 13: Impact of Prefix (O- vs. U-).

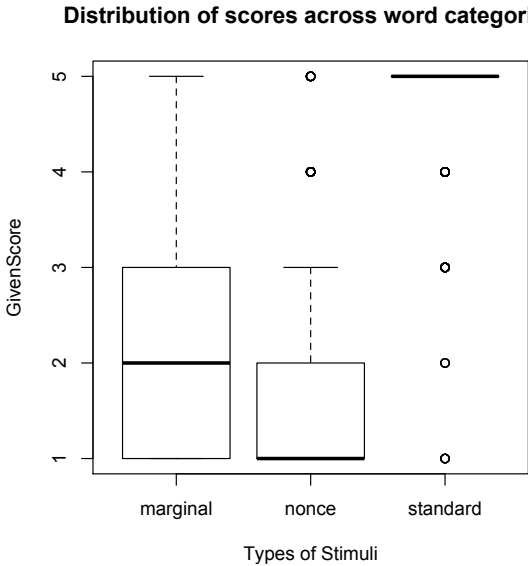


Figure 14: Impact of Word type.

We observe that Gender of a subject does not make any difference in terms of assigned scores (Figure 11), while Age of a subject makes a difference: children assign higher acceptability ratings than adults (Figure 12).

Likewise, verbs prefixed in O- overall tend to receive higher acceptability scores compared to U-verbs (Figure 13): half of O-verbs received scores higher than “3”, while half of U-verbs received scores higher than “2”.

Lastly, in Figure 14, we observe that types of stimuli (Word type) constitute three distinct patterns in terms of their ratings and MEDIANS: marginal verbs have the MEDIAN score “2” and in this sense they resemble nonce verbs that have the MEDIAN score “1”. Standard verbs, by contrast, receive the MEDIAN score “5” and form the most homogeneous group in terms of ratings, with only a few outliers represented in the plot

by small circles. Whiskers indicate that the ratings of marginal verbs can be as high as score “5” (= “This is an absolutely normal Russian word”), while whiskers of the nonce words reach only the score “3” (= “This word sounds strange, but someone might use it”). Overall, as shown in Figure 14, marginal verbs received from subjects surprisingly low acceptability scores: half of marginal stimuli received the lowest scores of 1 and 2.

9.5.3.3 Age

Table 13 aggregated the overall numbers of scores received by standard, marginal, and nonce factitive verbs within the two tested age groups – children and adults. The percentage ratio is given according to the maximal numbers of scores that twenty stimuli could receive within each group of subjects. The maximal number of scores for twenty verbs in the children’s age group is 7,000 (=5 points*20verbs*70 subjects). Likewise, the maximal number of scores for twenty verbal stimuli in the adult group is 5,100 (=5 points*20verbs*51 subjects).

Age group	Standard	Marginal	Nonce
Children: scores	6,896 = 98.5% of maximal number	3,400= 48.57% of maximal number	2,267= 32% of maximal number
Adults: scores	5,004 = 98.1% of maximal number	2,329= 45.66% of maximal number	1,402= 27.5% of maximal number

Table 13: Distribution of acceptability scores within two age groups of subjects.

Table 13 demonstrates that the proportion of scores with regard to the maximal number of scores within each stimulus-type condition is very similar: standard verbs were given 98.5% of the maximal score number by children, and 98.1% by adults; marginal verbs were assigned 48.57% of the maximal number by children, and 45.66% by adults; lastly, nonce stimuli were given 32% of the possible maximal number by children, and 27.5% by adults. The comparison of the two rows of percentages suggests that, overall, children give higher scores than adult speakers. In other words, children tend to rate stimuli as more acceptable and have a lower acceptability threshold than adults. This conforms to what we expected.

However, the statistical analysis suggests that the difference in the distribution of scores for the two age groups is not robust in this data. The chi-squared test shows that the difference in distributions of scores assigned by children and adults is statistically significant ($\chi^2 = 17.5$, $df = 2$, $p\text{-value} = 0.00015$), but the effect size is too small (Cramer’s $V = 0.028$) to consider it a reportable difference.

Another way to look at the difference between the two age groups of subjects is suggested in Table 14: here we compare children with adults in terms of the range of dispersion of scores received by individual stimuli within each stimulus type – Standard, Marginal, and Nonce.

If we sum the scores gained by each individual stimulus, we can see that individual stimuli form a continuum within each stimulus type. In this dispersed continuum, there is a word that received the largest sum of scores (MAX), and a word that received the smallest sum of scores (MIN). Among marginal words, the verb *opoxabit* ‘profane’ scored the largest overall number (i.e. sum of scores assigned by all subjects) in both age groups.

Subjects	Standard words	Marginal words	Nonce words
Adults (N=51)	MAX 255 (<i>ob''jasnit'</i>) - MIN 227 (<i>ožestočit'</i>) = DIS 28	MAX 198 (<i>opoxabit'</i>) - MIN 69 (<i>ukrasivit'</i>) = DIS 129	MAX 101 (<i>oblusit'</i>) - MIN 59 (<i>oduktit'</i>) = DIS 42
Children (N=70)	MAX 350 (<i>ob''jasnit'</i>) - MIN 322 (<i>ogolit'</i>) = DIS 22	MAX 281 (<i>opoxabit'</i>) - MIN 100 (<i>ukrasivit'</i>) = DIS 181	MAX 143 (<i>ogabit'</i>) - MIN 91 (<i>oduktit'</i>) = DIS 52

Table 14: Dispersal (DIS) of scores across individual stimuli within two age groups.

Adults gave this word a total of 198 points, while children gave it a total of 281 points. Similarly, in both age groups the smallest number of scores among marginal words was received by the verb *ukrasivit'* 'make prettier' – adults assigned it a total of 69 points, while children gave it a total of 100 points. The difference between MAX and MIN scores within each age group gives the range of dispersion (DIS) across individual stimuli. Table 14 demonstrates that in both age groups the dispersion range of marginal words is larger than that of the two other stimuli types. This data indicates that the variation in acceptability ratings across marginal words is greater in both age groups.

In order to compare the two strings of values for two age groups we calculated the mean value of dispersed scores: we divided the difference between the largest and the smallest scores (MAX-MIN) by the number of subjects in each group (51 adults and 70 children). Table 15 presents the resulting values which are very close to each other across the two age groups of subjects.

Subjects	Standard words	Marginal words	Nonce words
Adults (N=51)	28 / 51 = 0.549	129 / 51 = 2.529	42 / 51 = 0.823
Children (N=70)	22 / 70 = 0.314	181 / 70 = 2.585	52 / 70 = 0.742

Table 15: Mean of dispersed scores within two age groups.

A chi-squared test on the means in Table 14 supports the conclusion about the lack of difference between two age groups: the difference between two strings of values (22, 181, 52 for children; and 28, 129, 42 for adults) is not statistically significant – X-squared = 3.65, df = 2, p-value = 0.16. This provides us with additional evidence that the difference between the two age groups with regard to dispersion range of stimuli does not exist. In other words, children as a group and adults as a group assign equally diverse acceptability ratings to marginal words.

Note that Tables 14 and 15 contain information that we could not see in Table 13. Table 13 aggregates the overall number of scores assigned to all stimuli of three types and shows that marginal words lie closer to nonce words and farther from standard words. By contrast, Tables 14 and 15 introduce another measure of comparison – range of score dispersion within each stimulus type. Tables 14 and 15 demonstrate that in terms of dispersion and divergence within each group marginal words are different from both standard and nonce words. Marginal words constitute a category of their own which is characterized with a much greater degree of heterogeneity and non-uniformity than the two other word categories.

Lastly, we posed a third question: are children and adults different in terms of the distribution of response types? In other words, are adults more radical in their judgements than children: do they tend to assign minimal scores to unfamiliar words or do they express their doubts with higher acceptability ratings? In order to address this

question, we calculated the overall number of responses of each type – “5 point”, “4 points”, “3 points”, “2 points”, and “1 point” – given by the group of adults and the group of children. These raw numbers are presented in Table 16.

Age groups	5 points	4 points	3 points	2 points	1 point
Children (N=70)	1,440	280	546	658	1,276
Adults (N=51)	1,045	173	298	340	1,204
All subjects (N=121)	2,485	453	844	998	2,480

Table 16: Distribution of response types in two age groups.

This piece of data facilitates a number of observations. The response type “4 points” was the least popular for both age groups. The response types “3 points” and “2 points” were of similar popularity. The two extreme response types “5 points” and “1 point” account for two thirds of all responses, while a third part of responses is distributed among three intermediate types – “2”, “3”, and “4” points.

Figure 15 visualizes these proportions in terms of partial percentages (the vertical axis). The three bars represent children and adults separately (left-most bar and central bar) and together (right-most bar). The five colors indicate the five response types. Quantities of data that correspond to the same response type (the same score) are marked with identical color. The numbers inside each bar are the same as those given in Table 16.

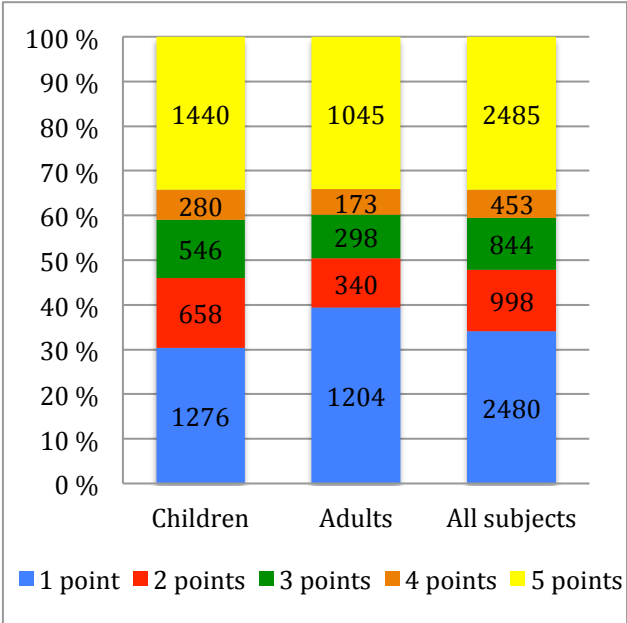


Figure 15: Distribution of response types in two age groups.

Again, we can see that the proportion of response types in the two age groups is very similar. However, there is one clear difference. As shown by the blue parts of the bars in Figure 15, the response type “1 point” is ten percent more frequent in the adult group of subjects than among children. This corresponds to those marginal and nonce stimuli that are rated as non-words by adults but receive higher ratings from children. This difference can be interpreted in terms of a higher threshold of word acceptability among adults as opposed to children. This means that when adult speakers encounter an unfamiliar word, in ten percent of cases they choose to rate it differently from children – as a non-word rather than as a possible but unfamiliar word with low frequency. In exactly these ten percent of cases, where the difference between the two age groups lies, teen-aged speakers accept marginal and even some nonce words as existing but rare or unknown to them. We can conclude that adults are ten percent more “radical” in their acceptability judgments than children when they face nonce and marginal words.

A chi-squared test on the distribution of response types given in Table 16 suggests that the difference between children and adults in this regard is indeed

statistically significant ($X^2 = 87.5$, $df = 4$, $p < 2.2e-16$). The effect size is small, but reportable (Cramer's $V = 0.1$).

Summing up, children and adult speakers show very similar patterns of responses triggered by three categories of stimuli. The difference between the two age groups is not significant in terms of how they evaluate the three categories of words and how dispersed their ratings are in each group of stimuli. However, we do find a difference with a small effect size in relative frequencies of response types: there is evidence in support of a higher threshold of word acceptability among adults, with about a ten percent difference between adults and teen-aged children.

9.5.3.4 Prefix

The next question is: does the prefix (O- vs. U-) have any impact on subjects' acceptability judgements? Are marginal verbs with the prefix O- better accepted than marginal verbs with the prefix U-? Do O-verbs receive higher acceptability ratings than U-verbs?

Because the difference between the two age groups turned out to be non-significant, in the following calculations we collapse the children's and adults' responses into one sample. Table 17 combines the data from two age groups and represents the distribution of overall scores across the two tested prefixes.

Prefix	Scores received by STANDARD words	Scores received by MARGINAL words	Scores received by NONCE words	Total score
Factitives in O-	5,871 = 97% of max	2,963 = 48.9% of max	1,912 = 31.6 % of max	10,746
Factitives in U-	6,029 = 99.6% of max	2,766 = 45.7% of max	1,757 = 29% of max	10,552

Table 17: Distribution of scores across the prefixes O(B)- and U-.

Recall that we expected that the higher token corpus frequency of the prefix O- (as opposed to U-) in novel factitive verbs would correlate with higher acceptability scores of marginal verbs with O- in the experiment. Table 17 shows that, indeed, this is the case: compare the total of 2,963 scores assigned to marginal factitives with O- as opposed to 2,766 scores assigned to marginal factitives with U-. When we compare these numbers with the overall maximal score possible (max), marginal factitive verbs prefixed in O- are perceived as more acceptable than marginal factitives prefixed in U-. This conforms to our prediction based on the higher productivity of the prefix O- in new coinages attested in the Russian National Corpus.

The maximal possible number of scores for a group of ten verbs evaluated by the total of 121 subjects is 6,050 (=5 points*10 verbs*121 subjects). We counted this maximal number as 100% and calculated the percentage ratio for each stimulus condition across the two prefixes. This gives us the following relative numbers: marginal factitives prefixed in O- received 48.9% of the possible maximal score, while marginal factitives in U- received 45.7%. Already these proportions suggest that the difference is not big. In order to evaluate the status of this difference with regard to the overall distribution of data, we subjected the raw numbers in Table 17 to a chi-squared test.

A chi-squared test showed that the difference between O- and U- in terms of received scores is statistically significant ($X^2 = 13.6$, $df = 2$, $p = 0.001$),

however the effect size is too small to count as a reportable difference (Cramer's $V=0.025$). This suggests that the contrast of O- and U- is not statistically robust.

Thus we should conclude that the contrast in type frequencies observed between the two prefixes in corpus data does not correlate with equally robust difference on the level of acceptability. However, the lack of a statistically significant difference in acceptability judgements for O- and U-verbs does not suggest that these prefixes lack a semantic contrast in standard and novel factitive verbs. Both O- and U-factitives were presented to subjects in contexts that fit their semantics. In this light, their similar acceptability might be indicative of the well-balanced experimental contexts chosen for the stimuli employed in our survey.

We further explored the difference between O- and U-stimuli from another perspective. In particular, we examined the dispersion of scores assigned to individual stimuli with the two prefixes. Tables 18 and 19 show how we calculated a range of dispersion (DIS), that is the difference between the maximal (MAX) and minimal (MIN) individual score within each group of verbs prefixed in O- and in U-.

Prefix	Range of dispersion		
	Standard words	Marginal words	Nonce words
Factitives in O- (N=30)	MAX 255 (<i>ob''jasnit'</i>) - MIN 227 (<i>ožestočit'</i>) = DIS 28	MAX 198 (<i>opoxabit'</i>) - MIN 98 (<i>omeždunarodit'</i>) = DIS 100	MAX 101 (<i>oblusit'</i>) - MIN 59 (<i>oduktit'</i>) = DIS 42
Factitives in U- (N=30)	MAX 255 (<i>ulučšit'</i>) - MIN 252 (<i>uxudšit'</i>) = DIS 3	MAX 149 (<i>udorožit'</i>) - MIN 69 (<i>ukrasivit'</i>) = DIS 80	MAX 77 (<i>ušadrit'</i>) - MIN 60 (<i>ukampit'</i>) = DIS 17

Table 18: ADULTS ONLY: Dispersion of scores across individual stimuli with two prefixes.

The strings of values in Table 18 refer to the data collected from adult subjects. A chi-squared test shows that in adults' responses the difference between the dispersion of scores in O-verbs is significantly different from the dispersion of scores in U-verbs ($X^2=16$, $df=2$, $p\text{-value}=0.00035$). The effect size is small to medium, close to medium (Cramer's $V=0.24$).

Similarly, we applied the same procedure to children's data presented in Table 19.

Prefix	Range of dispersion		
	Standard words	Marginal words	Nonce words
Factitives in O- (N=30)	MAX 350 (<i>ob''jasnit'</i>) - MIN 322 (<i>ogolit'</i>) = DIS 28	MAX 281 (<i>opoxabit'</i>) - MIN 115 (<i>ovnešnit'</i>) = DIS 166	MAX 143 (<i>ogabit'</i>) - MIN 91 (<i>oduktit'</i>) = DIS 52
Factitives in U- (N=30)	MAX 350 (<i>ulučšit'</i>) - MIN 344 (<i>uplotnit'</i>) = DIS 6	MAX 255 (<i>usovremenit'</i>) - MIN 100 (<i>ukrasivit'</i>) = DIS 155	MAX 121 (<i>utulit'</i>) - MIN 98 (<i>uloprit'</i>) = DIS 23

Table 19: CHILDREN ONLY: Dispersion of scores across individual stimuli with two prefixes.

The result of a chi-squared test on the strings of values in Table 19 indicated a statistically significant difference between O- and U-verbs ($X^2=17.2$, $df=2$, $p\text{-value}=0.00035$).

value = 0.00018). The effect size is small to medium, but slightly smaller compared to that for adults: Cramer's $V=0.2$.

Summing up the analysis of Tables 18 and 19, the difference between O- and U-verbs in range of dispersion of their ratings is statistically significant for both age groups of subjects, but has a slightly larger effect size for adults as opposed to children. In other words, the heterogeneity of scores given to O-verbs is larger than that of U-verbs. This suggests that verbs prefixed in O- as a group demonstrate more discrepancy and heterogeneity in subjects' judgements, as opposed to U-verbs. By contrast, verbs in U- as a group are perceived more uniformly.

9.5.3.5 Stimulus type

Before I turn to the statistical analysis of correlation between acceptability scores and stimulus types, I will compare two sets of raw numeric values obtained from the experiment and the corpus with regard to the sixty words we examined. Figures 16, 17, and 18 provide an overview of two numeric dimensions: the overall score assigned to verbs by adults and children in the experiment (Figure 16) and token frequencies of these verbs in the Modern Subcorpus of the Russian National Corpus – texts created in 1950-2013 (Figures 17, 18). For the ease of comparison, on each of these graphs verbs are ordered according to their token frequencies in the corpus. The vertical axis refers to the scores in Figure 16 and to corpus attestations in Figures 17 and 18. Comparing the acceptability ratings (top of the page) with corpus frequencies (bottom of the page) we can distinguish between three groups of words – standard, marginal, and nonce. However, there are some non-trivial differences which suggest that corpus frequencies do not entirely correspond to acceptability ratings assigned by speakers.

Let us first look at marginal verbs that are represented in the central part of the scope in Figure 16. Note that the two zigzag lines of acceptability scores in Figure 16 correspond to a decreasing line of token frequencies in Figures 17 and 18. Figures 16 and 18 differ the most in those parts of the graphs which represent marginal factitive verbs. This difference suggests that corpus frequencies cannot entirely predict the degree of acceptability for individual marginal verbs. For example, the verb *opoxabit'* 'profane' which has only a single corpus attestation in the experiment was evaluated overall as a more acceptable verb than some other marginal verbs with more corpus attestations like *udorožit'* 'make more expensive'. We can conclude that at the level of the small token frequencies (below ten attestations) that we deal with in the case of marginal verbs, the corpus data is not that informative and arguably cannot predict higher or lower acceptability rating of a marginal word.

The rightmost part of Figures 16 and 18 visualize nonce verbs. Because in the corpus such verbs are not attested, Figure 18 depicts an absolutely flat line of zero frequencies for the nonce verb group. In Figure 16 this corresponds to a wave-like red graph of acceptability scores assigned by children and a gradually decreasing blue line of acceptability scores assigned by adults. Whereas for marginal words the profiles of adults' and children's responses are rather similar, for nonce words they look different.

The leftmost part of Figure 16 and the entire Figure 17 visualize standard verbs. Here we observe the effect opposite to the nonce words profile: an almost flat line of acceptability scores depicted in Figure 16 corresponds to a declining line of various token frequencies of standard verbs in Figure 17. This result is natural because the experimental scale offered only one score for evaluation of highly frequent standard verbs – the score "5".

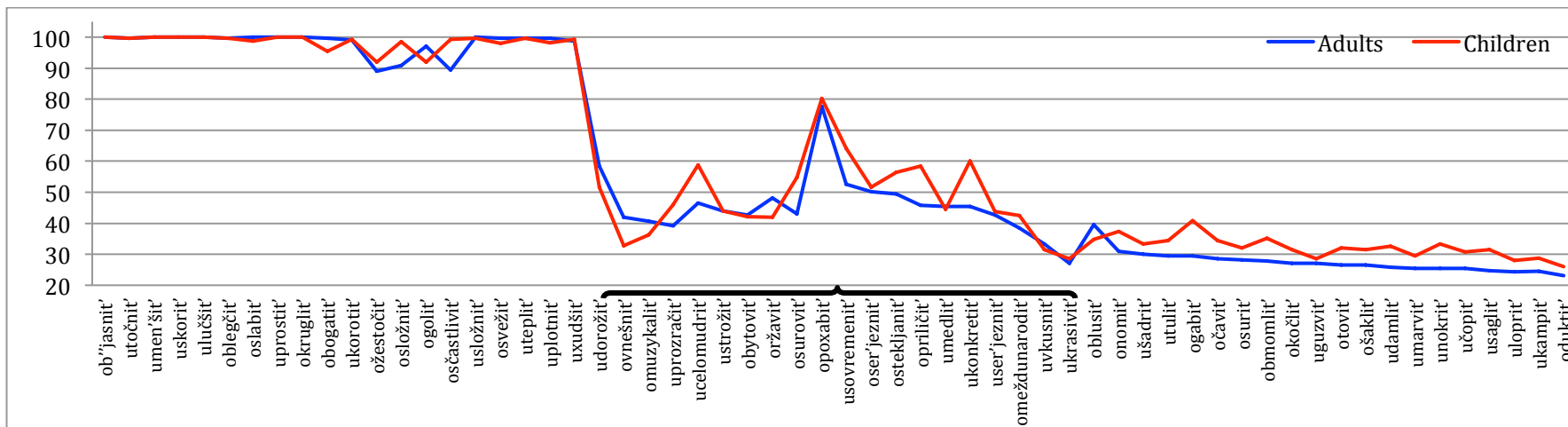


Figure 16: Percentages of total possible score assigned to individual stimuli by two age groups of subjects.

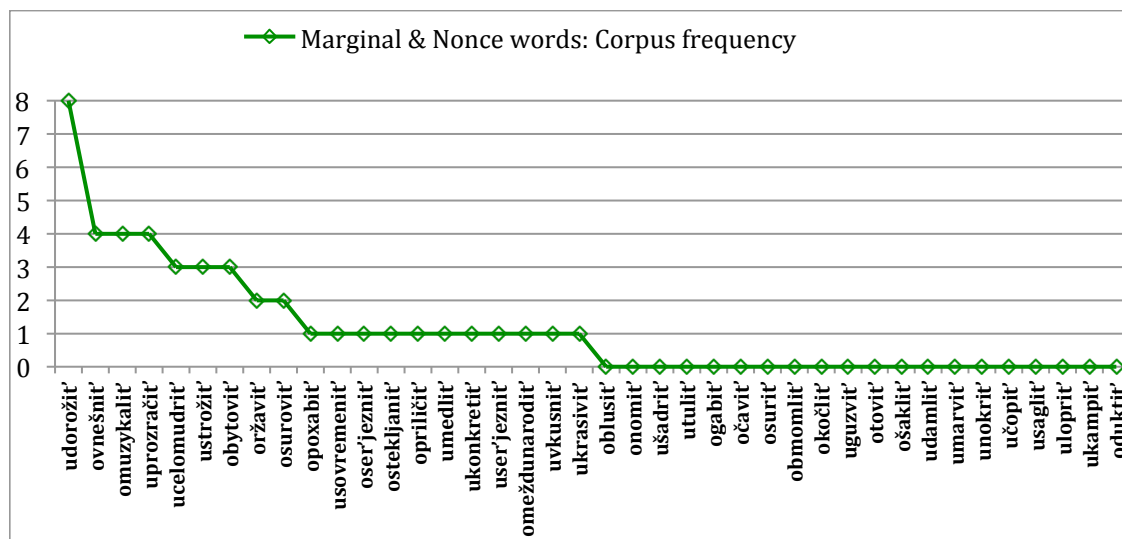
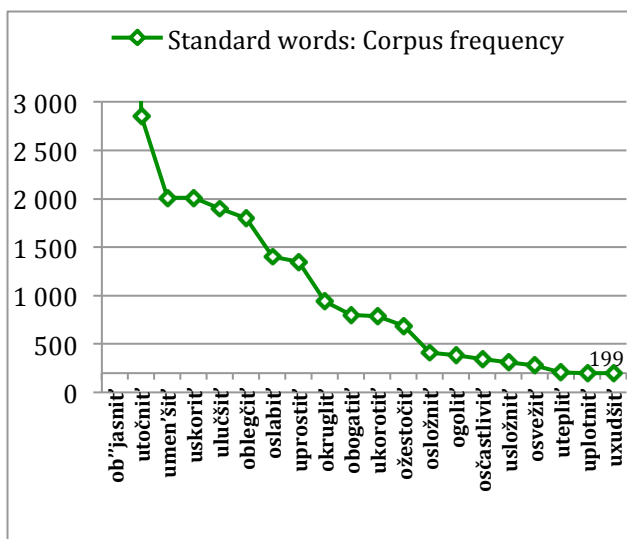


Figure 17, Figure 18: Number of attestations of experimental stimuli in the Modern Subcorpus of the Russian National Corpus.

Summing up this data overview, Figures 16, 17, and 18 visualize crucial correspondences between corpus frequencies and acceptability ratings and reveal their parallelism and differences. Corpus frequencies subdivide the sixty verbs used as stimuli into three clear groups that do not overlap: we turn the principal differences in their corpus frequencies into their categorical linguistic status – standard, marginal, and nonce words. By contrast, acceptability ratings suggest that the status of these three groups might be not categorical. In terms of overall total scores assigned to individual verbs, we observe not clear-cut boundaries but rather a continuum. In this continuum, some marginal verbs gain a smaller overall score than some nonce verbs. The difference between standard and marginal words might be smaller in terms of acceptability than it is in terms of frequency. In this continuum, we find overlaps of the zone of marginality and the zone of non-existence.

I will now examine the correlation between stimulus types and acceptability ratings. Given that the effect sizes for both age and prefix are ten times too small to be reported as significant, we combined the responses from the two age groups triggered by both prefixes and looked at the overall distribution of scores across the three types of stimuli. Table 20 presents this distribution and also calculates the percentage ratio of scores for each stimulus type with regard to the maximal number of scores – 12,100 (=5 points*20 verbs*121 subjects) taken here as 100%.

	Standard	Marginal	Nonce
Total score	11,900 (=98.3% of max)	5,729 (=47.3% of max)	3,669 (=30.3% of max)

Table 20: Combined data: children & adults, O- & U-).

In statistical analysis, we compared the three stimulus types (standard vs. marginal vs. nonce). Table 21 presents results of three calculation trials.

Opposition tested	Scores (combined data)	Chi-square test	Effect size
Standard vs. Marginal vs. Nonce	11,900 vs. 5,729 vs. 3,669	p-value < 2.2e-16	Cramer's V = 0.49 (large)
Standard vs. Marginal	11,900 vs. 5,729	p-value < 2.2e-16	Cramer's V = 0.35 (medium)
Marginal vs. Nonce	5,729 vs. 3,669	p-value < 2.2e-16	Cramer's V = 0.21 (small to medium)

Table 21: Results of chi-squared test for the three groups of stimuli.

The first chi-squared test was run for all three stimulus types. The smallest possible p-value and a large effect size (Cramer's V = 0.49) suggest that the difference between all three categories of words is significant and robust.

The second and the third chi-squared tests targeted the differences between only two classes – standard vs. marginal and marginal vs. nonce accordingly. While the difference in both cases turned out to be statistically highly significant (p-value < 2.2e-16 in both trials), the effect sizes measured by Cramer's V test were different. As shown in Table 21, the difference between standard and marginal stimuli in terms of scores is

bigger than the difference between marginal and nonce words: Cramer's $V = 0.35 >$ Cramer's $V = 0.21$. Summing up, in terms of acceptability, marginal words are much closer to nonce words than to standard words.

This result is also supported by ANOVA analysis run on combined data (both age groups and both prefixes all together)¹⁵⁹. Table 22 aggregates the key parameters that characterize each type of stimuli in terms of acceptability ratings.

Standard Verbs	Marginal Verbs	Nonce Verbs
MAX = 605	MAX = 479	MAX = 223
MEAN = 595	MEAN = 286.4	MEAN = 183.4
MIN = 549	MIN = 169	MIN = 150
stand dev = 15	stand dev = 67	stand dev = 19
variance = 235	variance = 4446	variance = 360

Table 22: Overall distribution of scores by Standard vs. Marginal vs. Nonce stimuli.

The boxplot in Figure 19 visualizes these parameters. Along the horizontal axis there are three types of stimuli – Standard, Marginal, and Nonce. The vertical axis reflects the distribution of scores.

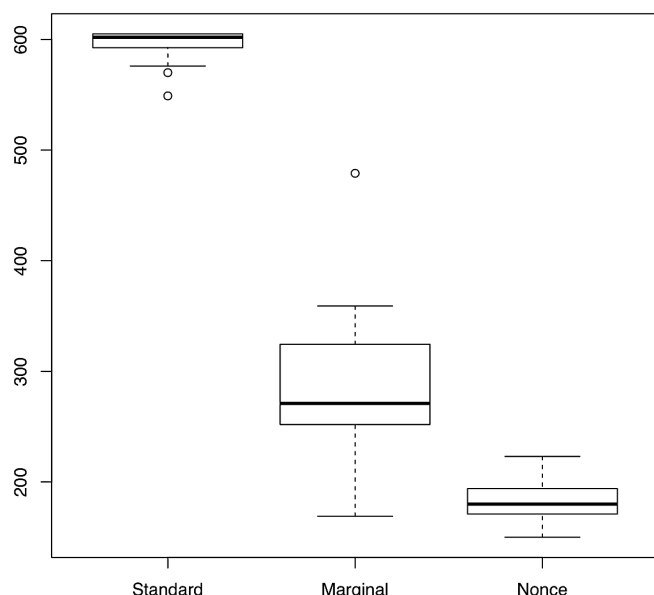


Figure 19: Three types of stimuli and distribution of acceptability ratings.

ANOVA analysis supports the idea that all three categories of words are perceived by speakers differently: the difference between distribution of acceptability scores across the three classes is significantly different ($F = 546$, $df = 2$, $p\text{-value} < 2.2e-16$).

Another crucial result is that marginal factitive verbs are evaluated by speakers more like nonce words rather than standard “normal” words. We can see this by comparing the MEAN values boldfaced in Table 22 and visualized as thick horizontal lines within the boxes in Figure 19. In terms of acceptability ratings, marginal words are much closer to nonce words than

to standard words. This conclusion suggests a compromise between Hypothesis 2 and Hypothesis 3.

Note that marginal words are semantically transparent, while nonce words are not. Thus, our finding that marginal words are rated more like nonce words than like standard words indicates that speakers are more sensitive to frequency than to semantic transparency. This suggests that frequency, which is related to performance, is a stronger factor than competence (ability to unpack morphological patterns). Therefore, memory may be a stronger factor than use of productive rules.

¹⁵⁹ I am grateful to Laura A. Janda who conducted the ANOVA analysis for this data. The R script is available at <http://ansatte.uit.no/laura.janda/PossWords/PossWords>.

On the other hand, marginal words exist on their own terms, different from both standard and nonce words in terms of much higher variation across stimuli. Table 22 demonstrates that marginal words are characterized by much higher standard deviation and variance (cf. two bottom lines of Table 22). The continuum that marginal factitives form is spread along a much longer scale than the continuum that we find among standard factitives and the continuum of nonce factitives.

Table 22 shows the highest (MAX) and the lowest (MIN) scores received by individual stimuli in each category. We can see that the variation across stimuli is much larger for marginal words (MAX-MIN=479-169=310 points) than for the two control groups (MAX-MIN=56 for standard words and MAX-MIN=73 for nonce words). This is reflected by the other two important measures of distribution – standard deviation and variance. These two measures demonstrate the dispersal of data with relation to the MEAN. Variance shows the mean of scores that deviate from the mean of distribution. Standard deviation is related to the value of Variance. Standard deviation equals the square root of the variance. Both measures show that the data is spread to a greater extent in the area of marginal words.

Moreover, high variation can be observed not only across individual marginal verbs but also across subjects: different subjects provide very different, sometimes contradictory judgements for the same marginal words. Examination of responses shows that most of the variation across subjects is related to evaluation of marginal words. Table 23 provides a few examples.

Marginal factitive	Gloss	Number of subjects who gave				
		5 scores (normal word)	4 scores	3 scores	2 scores	1 score (does not exist)
<i>usovremenit'</i>	'modernize'	22	26	27	18	28
<i>opriličit'</i>	'make decent'	9	25	33	22	31

Table 23: Variation across subjects regarding the same marginal stimuli.

Table 23 demonstrates that the verb *usovremenit'* 'modernize' was rated as a normal Russian word by twenty-two subjects and as a non-existing word by twenty-eight subjects. And there are equally many subjects that rated it in between these two extremes: twenty-six subjects decided that it is "a normal but rarely used word", twenty-seven subjects suggested that "this word sounds strange, but someone might use it", and eighteen subjects evaluated it as "a strange word unlikely to be used". Similarly, the novel marginal verb *opriličit'* 'make decent' also received very contradictory acceptability judgements. All scores received by each stimulus are presented in Appendix 7.

9.5.3.6 All factors in a single model: Advanced statistical modeling

In previous sections we looked separately at each of the possible factors that might determine the acceptability rating of a factitive verb. Another possibility is to include all possible factors into a single multifactorial model and let the model calculate the relative impact of each factor with regard to other factors, as well as their possible interactions.

In this section I present three different statistical models that I used in order to conduct a multifactorial analysis of experimental data. The three analyses offer comparable yet complementary insights into what determines the speaker's choice of an

acceptability score. Apart from the power to handle multifactorial analysis, there is another advantage of these models that is worth mentioning. Each of three models is appropriate for non-interval data. The first two models – Ordinal Logistic Regression and Ordinal Mixed-Effects Regression – are specifically designed for ordinally-scaled data. The third model – Classification and Regression Trees (CART) combined with Random Forests – is an alternative non-parametric statistical technique which does not impose interval status on non-intervally-scaled data either. Before I turn to the first model, I summarize the key features of our data.

The goal of this statistical analysis is to determine and evaluate the strength of the correlation between the dependent variable (also called an outcome response variable) and independent variables, or predictors. In our case, the dependent variable is the acceptability score elicited in the experiment, while the key predictor variables are stimulus type, prefix, and age of speaker.

It is crucial that the outcome response variable constitutes ordinal scale data, with scores of 1 to 5 points. This ordinal scale is an internally ordered set of categorical values, where it is not certain that the intervals between adjacent scores are of equal magnitude. Therefore, our goal is to find such a model that is suitable for ordinal data and does not impose interval status on this set of values.

It is also important that among possible predictors there are those that belong to so-called fixed-effects factors and random-effects factors (Baayen 2008: 241-250). All fixed-effects factors are categorical factorial predictors with the following values:

Fixed-Effects factors:

- WordType: standard, marginal, nonce
- AgeGroup: child, adult
- Prefix: O-, U-
- Gender: male, female

Note that this set of factors includes the sociolinguistic factor of Gender that we briefly addressed in 9.5.3.2 in Figure 11. At the same time, we exclude such sociolinguistic factors like City and Education because otherwise they would lead to a collinearity problem with the factor AgeGroup. The reason why collinearity would occur is that these three factors overlap with each other because all children come from the same city and have the same education level. Out of these overlapping factors we are interested in exploring the effect of AgeGroup, so we include it in our model. The factors City and Education are redundant and therefore excluded. The same collinearity problem would arise if a model would include both Corpus Frequency of stimuli and their WordType status. Since word categories (Standard, Marginal, and Nonce) used in the experimental design are established on the basis of corpus token frequencies of factitive verbs, we exclude Frequency from our model as a redundant overlapping variable.

Apart from fixed-effects factors we should also account for random-effects factors, so that our results would be applicable not only to the specific set of tested stimuli and the specific group of chosen subjects. The objective is to make conclusions that could go beyond this specific experiment and be generalized to other words with these linguistic parameters and to a broader population of speakers. Therefore, we distinguish the following random-effects factors that can affect the overall distribution:

Random-effects factors:

- Subject: 121 persons
- Stimulus: 60 verbs

Each of 121 subjects was exposed to each of sixty stimuli once. In total, we account for 7,260 datapoints (subjects' responses), where each response is characterized with twelve parameters – Stimulus, WordType of the stimulus, Prefix, individual anonymous Subject code, AgeGroup, Gender, as well as Context (sentence), Corpus frequency, subject's exact Age, Education level, Profession, and City. This yields 94,380 cells in an Excel spreadsheet. We are interested in statistical exploration of the first six parameters. Together with 7,260 datapoints they constitute 50,820 cells in Excel. The task of the statistical analysis is to infer from this data the key predictor factors that determine its overall distribution. The null hypothesis is that there are no statistically significant correlations among the variables. The alternative hypothesis is that such correlations do exist.

It is worth mentioning that three statistical models that we present account for the variability of data distribution as far as it is feasible for each of them. At the same time, the models focus on different aspects of data and reveal similar yet complementary insights about how the data is internally organized. I will present each model in turn, highlighting its advantages and the outcome, and then discuss the overall results. The complete R script of all models is available at <http://hdl.handle.net/10037.1/10078> in the document “R script O U EXPERIMENT ALL MODELS”.

9.5.3.6.1 Ordinal Logistic Regression

Logistic regression is a well established robust and powerful statistical technique that is widely used for multifactorial analysis (Strobl et al. 2009: 323; Baayen et al. 2013: 260). However, as Baayen (2008: 208) points out, a logistic regression analysis is appropriate for those dependent variables that are dichotomous, i.e. contain binomial values. In our case we are dealing with a multinomial dependent variable with five ordered values, where the score “5” is higher than the score “4”, the score “4” is higher than “3”, and so on. For such ordered dependent variables it is appropriate to use the kind of logistic regression which is specifically designed for ordinal data analysis – an Ordinal Logistic Regression (Baayen 2008: 208-214).¹⁶⁰

In this analysis we used the packages languageR, rms¹⁶¹, and MASS and the function `lrm()`¹⁶². The analysis was conducted using R version 2.15.0.

¹⁶⁰ “When a data frame is read into R, the levels of any factor are assumed to be unordered by default” (Baayen 2008: 209). Therefore, in order to make the outcome variable Score an ordered factor with levels 1<2<3<4<5 we used the function `ordered()`: `dat$Score=ordered(dat$Score, levels=c("E","D","C","B","A"))`.

¹⁶¹ Because the package ‘Design’ was removed from the CRAN repository, we used the package ‘rms’ instead.

¹⁶² What is crucial for the function `lrm()` of the Ordinal Logistic Regression model is that it “assumes that the effects of our predictors <...> are the same <...> across all levels of our ordered factor” (Baayen 2008: 212). Although this might somewhat simplify the outcome, we nevertheless obtain an important generalization about the statistically significant predictors of data distribution.

We explored the impact of four predicting factors – AgeGroup, Prefix, WordType, and Gender. The impact of Gender was found insignificant: Chi-Square= 0.33, df = 1, p-value = 0.56. The final and most optimal model included three factors that have significant effect on the choice of the Score¹⁶³ – WordType, AgeGroup, and Prefix. Note that here we account for main effects only. The ANOVA table suggests the following characteristics of these significant predictors:

Factor	Chi-Square	Degrees of freedom	p-value
AgeGroup	59.28	1	< .0001
Prefix	5.45	1	0.0195
WordType	3415.95	2	< .0001
TOTAL	3425.06	4	< .0001

Table 24: Outcome of the Ordinal Logistic Regression: Wald Statistics.

Note that the p-value for the factor Prefix is 0.02, which is smaller than 0.05¹⁶⁴. This means that the impact of Prefix should be considered significant, even though its significance is smaller than that of the factors WordType and AgeGroup which have p-value < .0001. If we compare the chi-square value of WordType (3415.95) with chi-square values of AgeGroup (59.28) and Prefix (5.45), we can conclude that the impact of WordType accounts for most of data, while the other two factors are very minor.

The summary of the Logistic Regression Analysis provides the measures of predictive strength of the model. All three important measures – C¹⁶⁵, Somer’s Dxy¹⁶⁶, and the R² index (Harrel 2001: 248; Baayen 2008: 204) – are high and indicate the high predictive strength of the model.

		Model Likelihood		Discrimination		Rank Discrim.	
		Ratio Test		Indexes		Indexes	
Obs	7260	LR chi2	7618.29	R2	0.689	C	0.855
max deriv	7e-12	d.f.	4	g	3.136	Dxy	0.710
		Pr(> chi2)	<0.0001	gr	23.016	gamma	0.754
				gp	0.380	tau-a	0.518
				Brier	0.119		

Table 25: Outcome of the Ordinal Logistic Regression.

Summing up, in the Ordinal Logistic Regression analysis we approached the dependent variable Score as ordinal data. This analysis shows that three factors are statistically

¹⁶³ The formula that we used: `dat.lrm2=lrn(Score ~ AgeGroup + Prefix + WordType, data=dat, x=T, y=T)` and the command used was `anova(dat.lrm2)`.

¹⁶⁴ “In language and linguistic research it is customary to take an *alpha* decision level of 5 per cent ($p < 0.05$). This means that there is less than 5 per cent probability that rejecting the null hypothesis will be an error”. (Cantos Gómez 2013: 49; cf. also Baayen 2008: 188).

¹⁶⁵C is the index of concordance between the predicted probability and the observed response. According to Baayen (2008: 204), “[w]hen C takes the value 0.5, the predictions are random, when it is 1, prediction is perfect. A value above 0.8 indicates that the model may have some real predictive capacity”. In our case, C is higher than 0.8, which suggests that the model has a high predictivity.

¹⁶⁶ Somer’s Dxy is an index of a rank correlation between predicted probabilities and observed responses. According to Baayen (2008: 204), “this measure <...> ranges between 0 (randomness) and 1 (perfect prediction).”

significant predictors of acceptability scores – WordType and AgeGroup (with p-values < 0.0001, or ***) and Prefix (with p-value=0.0195, or *)¹⁶⁷.

9.5.3.6.2 Regression Mixed-Effects Model for Ordinal Data

The Ordinal Logistic Regression model presented in the previous section accounts for the fixed-effects factors, namely WordType, AgeGroup, and Prefix. However, apart from these factors, the experimental data can also be affected by random-effects factors like the bias of individual subjects and individual stimuli.

Both subjects and stimuli are sampled randomly from the overall population of speakers and words, but we want to obtain a generalization about the data that would go beyond these specific subjects and specific stimuli. In other words, we need a model that can generalize over the bias of individual subjects and stimuli and determine a tendency which predominates over the effect of random effects.

A common model for experiment-based data where multiple subjects respond to multiple items is a Mixed-Effects Model (Baayen 2008: 242-302). Mixed-effects models are primarily used to explore data with nominal binomial dependent variables (0/1, A/B) (e.g. Tagliamonte & Baayen 2012) or continuous numerical dependent variables, for example reaction time (e.g. Baayen 2008: 242-302).

In order to account for a multinomial ordinal dependent variable by means of a mixed-effects model, I used the package Ordinal in its latest version 2013.9-13¹⁶⁸ available in R version 3.0.2. I used the function `clmm()` which can handle the crossed random-effects structure of two factors – Subject and Stimulus.¹⁶⁹ It is worth mentioning that technically the Regression Mixed-Effects Model is a parametric model, but it does not assume a normal distribution for the response. In this sense, it does not make dubious parametric assumptions about the data.

In this analysis, the factors Gender and Prefix were found insignificant in terms of predicting the dependent variable Score. After elimination of these factors, the most optimal fitted model¹⁷⁰ indicated the significant effects of two factors – WordType and AgeGroup. Again, we take into account only main effects. Tables 26 and 27 report on model’s output concerning the random- and fixed-effects factors:

Groups	Name	Variance	Standard Deviation
SubjectCode	(Intercept)	1.091	1.045
Stimulus	(Intercept)	1.043	1.021

Table 26: Random-effects factors.

In the right-most column of Table 27 the following common codes of significance are used: the two-star symbol ** indicates that the p-value of the contrast between two age groups is at the level of 0.001, and the three-star symbol *** suggests that the contrast between word types is larger and lies at the level of even smaller p-value.

¹⁶⁷ According to the common set of codes that indicate significance, the number of stars corresponds to the level of significance: 0 = ‘***’, 0.001 = ‘**’, 0.01 = ‘*’, 0.05 = ‘.’, 0.1 = ‘.’.

¹⁶⁸ See the description at <http://cran.r-project.org/web/packages/ordinal/index.html>

¹⁶⁹ I am indebted to Rune Haubo Bojesen Christensen for pointing out this possibility to me.

¹⁷⁰ The formula used is: `fm2 <- clmm(Score ~ AgeGroup + WordType + (1|Stimulus) + (1|SubjectCode), data=dat, Hess=TRUE)`.

	Estimate	Std. error	z value	Pr(> z)	
AgeGroup-child	0.5803	0.2013	2.883	0.00394	**
WordType-nonce	-1.7791	0.3292	-5.405	6.48e-08	***
WordType-standard	7.4203	0.3712	19.991	< 2e-16	***

Table 27: Fixed-effects factors: Coefficients.

Summing up the outcome of Regression Mixed-Effects Model, the impact of only two fixed-effects factors was found statistically significant – WordType and AgeGroup. The effect of WordType is more significant than that of AgeGroup. Note that the impact of Prefix which had the largest p-value (*) in the Ordinal Logistic Regression, is found insignificant in the Mixed-Effects model.

9.5.3.6.3 Classification and Regression Trees (CART) and Random Forests

Classification and Regression Trees (also abbreviated as CART) is a new method that is quickly gaining popularity in genetics, medicine (Strobl et al. 2009: 324), social sciences, and linguistics (cf. most recent applications in Tagliamonte & Baayen 2012 and Baayen et al. 2013). Classification and Regression Trees is a non-parametric statistical technique which is appropriate for non-interval data. In particular, CART analysis provides a powerful tool to explore an ordinally scaled dependent variable (Faraway 2006: 253-268; Baayen 2008: 148-164). The Trees method has many advantages and has proven to give robust results, comparable with more traditional models like Logistic Regression, and even to give more accurate predictions, especially regarding complex multifactorial interaction effects which cannot be identified by parametric techniques (Baayen 2008: 154; Baayen et al. 2013). In a linear model like Logistic Regression the predictors are analyzed in a linear way in order to model their impact on the response (dependent) variable. By contrast, nonparametric regression models like Trees do not employ linearity and are often more flexible in modeling combinations of predictors (Faraway 2006: v). Because Trees do not hold any assumptions about the normal distribution of the response variable (as opposed to the logistic regression model), the Trees method can cope with any data structure and type and is highly recommended for unbalanced datasets. Robust results of this method are achieved by the use of recursive partitioning, bootstrapping, bagging, and cross-validation (cf. Strobl et al. 2009 for details). Apart from the high processing capacity to handle a large number of predictors non-linearly, the CART analysis also offers measures of variable importance, or predictive strength of tested variables. Variable importance ranking is available via the extension of the CART method to the so-called Random Forest approach. A Random Forest is an ensemble of Classification or Regression Trees which produces a scale of variable importance. The scale makes it possible to compare all tested predictors with each other in terms of their strength.

How exactly do Classification and Regression Trees work? CART is an algorithm-based method (Faraway 2006: 253). The outcome of the CART analysis is a graphically plotted “tree” created via a recursive partitioning of data. The Tree represents an algorithm of data partitioning which consists of recursive binary splits, each based on one variable. The Tree outlines a decision procedure of predicting the values of the dependent variable. As a result, recursive splits subdivide the entire data set into several non-overlapping subsets of data. Each split reduces the error and increases the “purity” of a subset of data points (the “principle of impurity reduction”, cf. Strobl et al. 2009: 326). The Tree is optimal at each split. However, each local split is not necessarily

globally optimal, meaning that a factor that might have a significant effect locally in the Tree, might be insignificant with regard to the entire data set.

Both Classification Tree (henceforth Ctree) and Regression Tree (henceforth Rtree) employ recursive partitioning but differ in terms of the types of response data the Tree is used for. A Ctree applies to factorial dependent variables and treats the values of a dependent variable as a categorical scale, while an Rtree applies to numerical dependent variables (Baayen 2008: 148). Because Ctree and Rtree handle different kinds of data, they differ in mechanisms of data partitioning. Ctree makes splits according to the principle of increasing purity of a node: after each split the subgroups of data observations should become purer, or more of the same kind. An Rtree employs the residual sum of squares as a criterion for splitting the nodes (Faraway 2006: 261). In addition, Rtree also computes the mean within each partition (ibid: 261).

I used both Ctree and Rtree for the experimental data. Comparing the outcomes of different treatments of the same data set was both interesting from the methodological perspective and reasonable in the light of uncertainty about the status of acceptability scores in terms of measurement scale.

The two resulting Trees¹⁷¹ turned out to be very similar but at the same time not entirely identical. The Ctree of acceptability ratings is presented in Figure 20. The Ctree treats the values of the dependent variable Score as categorical data: A = score “5”; B = score “4”; C = score “3”, D = score “2”, and E = score “1”. By contrast, the Rtree of acceptability ratings presented in Figure 21 approaches scores as numerical ordinally scaled data: from 5 points to 1 point.

Although the Ctree expands to the left, while Rtree stretches to the right, they make almost identical splits, just in different order. Crucially, both Trees demonstrate that WordType determines the major split of data at the root node (node 1), followed by Prefix as the second level, and AgeGroup at the third level.

The root node is the same in both Trees – WordType. Note that the decision rule of the root node partitions data into two large subsets, grouping together marginal and nonce verbs and setting them apart from standard verbs. We can interpret this partitioning in terms of a close connection between marginal and nonce verbs in terms of their similar acceptability ratings and a larger distance of marginal verbs from standard verbs. Recall that this generalization is also supported by the ANOVA analysis: marginal verbs as a group pattern more similarly to nonce verbs than to standard verbs.

In both trees standard verbs are further split according to Prefix. Terminal (leaf) nodes 12 and 13 of the Ctree (Figure 20) and nodes 3 and 4 of the Rtree (Figure 21) demonstrate that standard verbs prefixed in U- as a group receive slightly higher acceptability ratings (i.e. are better excepted) than standard verbs prefixed in O-. In particular, the plots of terminal nodes show that among verbs prefixed in O- there are more outliers that receive scores lower than “5” than in the group of U-verbs. This is supported by a total of 2,420 datapoints (see the numbers that appear on the node square boxes).

In both trees, in the branch opposite to standard verbs, WordType further determines the split into marginal and nonce stimuli. In the Ctree, marginal verbs are further subdivided according to Prefix and AgeGroup. These subsequent splits suggest that marginal verbs prefixed in U- (node 4 in Figure 20) receive slightly more rejections (score “1”) than O-verbs (nodes 6 and 7).

¹⁷¹ Both analyses were carried out in R version 2.15.0.

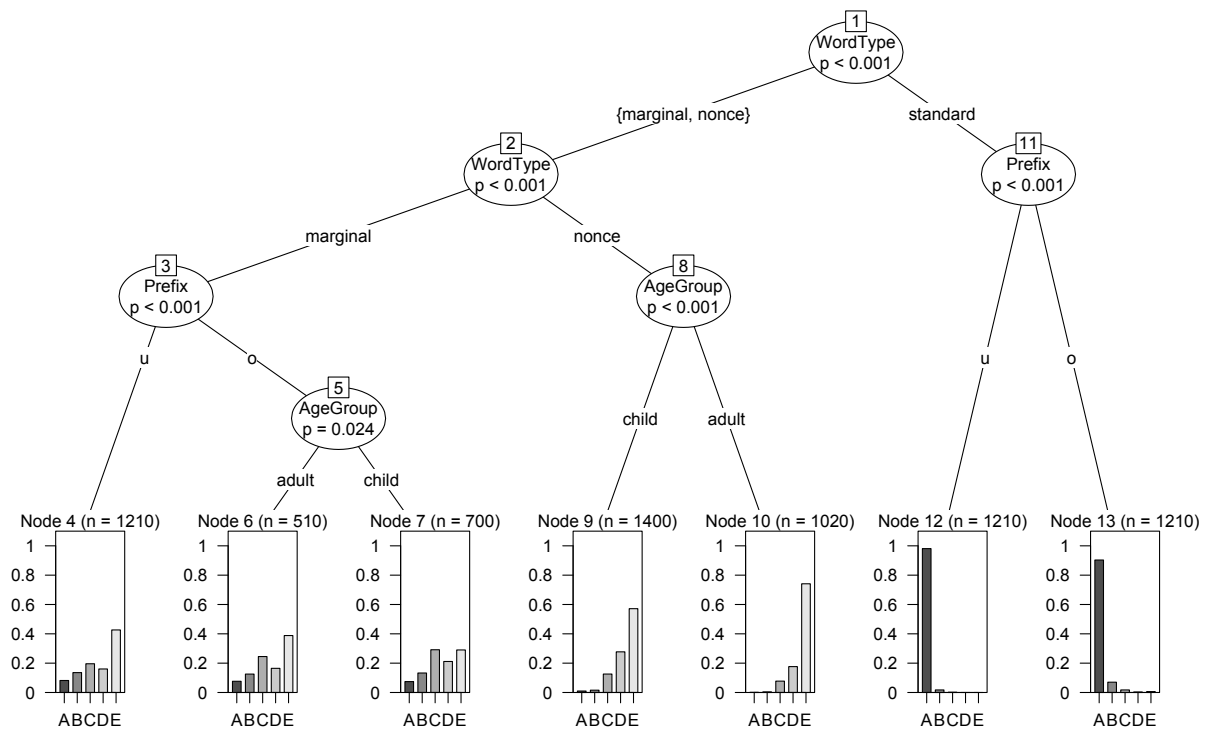


Figure 20: Classification tree of acceptability ratings: scores are treated as categorical data: A-score "5"; B-score "4"; C-score "3"; D-score "2"; E-score "1".

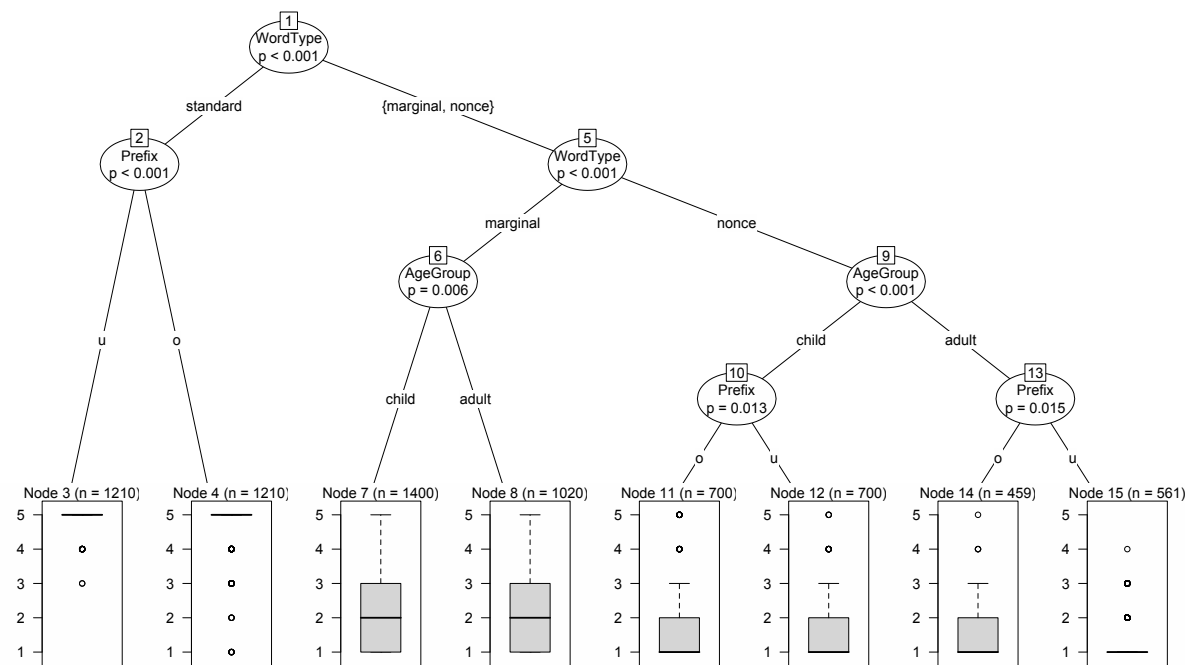


Figure 21: Regression tree for of acceptability ratings: scores are treated as numerical ordinal data – from 5 points to 1 point.

Meanwhile, for O-verbs we can observe an interaction effect of Prefix and AgeGroup: adult speakers tend to reject marginal verbs prefixed in O- more often than children (compare the bars representing “E” score in nodes 6 and 7). The Ctree suggests an AgeGroup effect for nonce verbs as well (see node 8 in Figure 20). Again, adults tend to completely reject nonce verbs regardless of their prefix more often than children do (compare “E” bars in terminal nodes 9 and 10).

Interestingly, the Rtree has the same predictors in the marginal & nonce branch, slightly rearranging their order of application. Marginal verbs are partitioned according to AgeGroup, however the difference between adults and children in this domain must be very small because nodes 7 and 8 (Figure 21) look identical. The group of nonce verbs, by contrast, is affected by the interaction of AgeGroup and Prefix: for adults both O- and U-verbs pattern pretty much the same (compare the nodes 11 and 12), while for children nonce verbs prefixed in O- (as opposed to U-verbs) tend to be more acceptable or at least more diverse in terms of their ratings and include more outliers with scores higher than “1”.

Summing up, both trees show high-level interactions of WordType, Age group, and Prefix. Both Ctree and Rtree visualize what is going on in the data distribution with respect to three factors and offer interesting insights. The structure of both trees is surprisingly similar: in both trees WordType is the most important factor, while Prefix and AgeGroup play their roles locally, making rather slight differences. The effects of AgeGroup and Prefix are statistically significant and optimal only within the scope of each local split. The role of these factors in the overall data distribution is different (much smaller), as clearly shown in the Random Forest analysis.

In order to compare the two Random Forest analyses consider Figures 22 and 23. Both figures present barplots of variable importance scores for factorial predictors of acceptability ratings. Figure 22 presents the outcome of the Random Forest analysis of acceptability scores taken as categorical data (A, B, C, D, E), while Figure 23 is the result of the Random Forest analysis of acceptability scores taken as ordinal data (A>B>C>D>E).

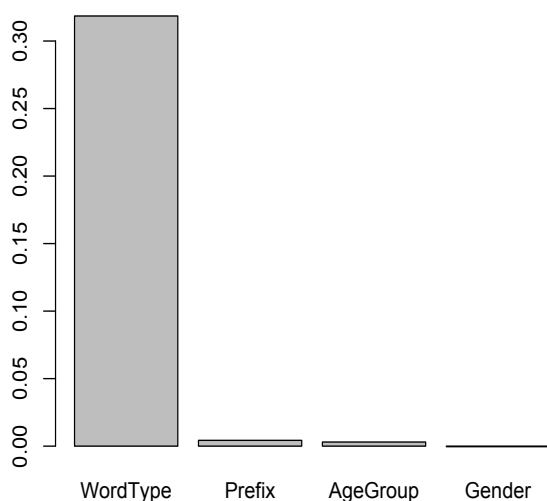


Figure 22: Variable importance scale for categorical data (A, B, C, D, E).

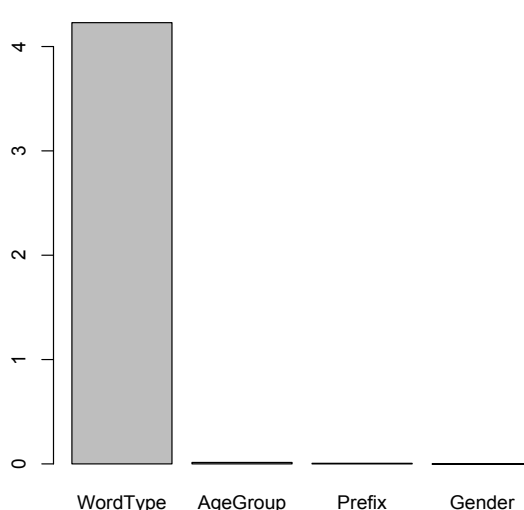


Figure 23: Variable importance scale for ordinal data (A>B>C>D>E).

Both barplots visualize a scale of relative importance, where the predictors of the dependent variable Score are ranked according to their relative strength. Each bar represents one predictor. Both plots depict the same four factors and arrange them almost identically. First of all, both plots show that WordType is absolutely the strongest predictor, while the impact of other factors is close to zero. Both plots show that Gender is the weakest predictor of all (recall that it appeared in neither of the Trees). Prefix and Age Group are ranked differently: Forest analysis of categorical data (Figure 22) suggests that Prefix is slightly stronger than AgeGroup, while Forest analysis of ordinal data (Figure 23) supports the reverse ranking, with a stronger impact of AgeGroup followed by Prefix. However, the difference between the importance scores of these two factors is very small in both plots.

9.5.3.7 Discussion of experimental results

The goal of this section is to summarize the outcomes of various statistical techniques and highlight what was consistent throughout all analyses.

We approached the data from different perspectives, applying both parametric and non-parametric statistics and also models specifically designed for handling ordinal data. The outcomes of different models are similar yet complementary. In the Ordinal Logistic Regression we found a significant effect for all three factors. In the Ordinal Regression Mixed-Effects Model only WordType and AgeGroup showed a significant effect. In the Trees and Forests analysis only WordType was the major predictor while AgeGroup and Prefix gain significance within local subsets of data. I suggest that the latter model is the most insightful and fruitful regarding this data. As a non-parametric test, Classification and Regression Tree method demonstrates that the importance of a factor can belong to different “levels”: what is crucial at the level of a local split (AgeGroup and Prefix) might have very small overall predicting power considering the entire dataset, while other factors (like WordType) can determine the major trend of data distribution, as we saw in the major split of the Trees and the highest bar in the Random Forest plots. The outcome of Random Forest analyses indicates that AgeGroup and Prefix do have some importance but their effect is very small. Indeed, this effect is revealed in high level interactions of the factors depicted in the Classification and Regression trees.

The major role of WordType is supported by Trees, Random Forests, ANOVA test, Ordinal Logistic Regression Model and Ordinal Mixed-Effects Regression Model.

The relatively small importance of Prefix revealed by the Random Forest analysis is comparable with the outcome of Ordinal Logistic Regression, where Prefix is the least significant of three factors; and is also parallel to the result of Ordinal Mixed-Effects Regression, where Prefix is not found to be significant at all.

The low predictive strength of AgeGroup revealed by Random Forest corresponds to what was found by chi-squared tests and ANOVA test. At the same time, this contradicts with the result of the Ordinal Logistic Regression and the

Mixed-Effects Regression analyses, where the effect of AgeGroup was found to be statistically significant.

Lastly, the effect of Gender is insignificant according to all models where it was tested.

9.6 Conclusions

This chapter provides a number of insights.

First, corpus data shows that the phenomenon of factitive verbs is broader than assumed in the literature. There are seventeen aspectual prefixes that can be employed in the formation of factitive verbs. In order to account for the choice of the prefix, I propose *the Spatial Motivation Hypothesis*, claiming that the spatial meaning of the prefix plays a major role.

Next, I explore whether this hypothesis applies also to O- and U-, which are the most frequent and productive prefixes in factitive verbs. Although their semantics in factitive verbs is bleached, I proposed that O- and U- encode different types of changes of states and describe this in terms of *the Scalarity Hypothesis*. Moreover, I suggest that the semantic contrast between these prefixes can be traced back to their spatial meanings. This conclusion also lends support to *the Spatial Motivation Hypothesis*.

Because the choice of the prefix is motivated by prefix spatial semantics, I conclude that the prefixes employed in factitive verbs represent distinct morphemes rather than allomorphs of a single morpheme-verbalizer.

My experimental study targets those factitive verbs that have marginal status in Modern Russian. I tested whether the prefix (O- vs. U-) correlates with higher or lower acceptability of novel coinages in perception of native speakers. Statistical models applied to elicited acceptability scores indicate that the importance of the prefix is relatively small.

In terms of acceptability, marginal words pattern closer to nonce words than to standard words. This finding might be explained by the linguistic culture specific for Russia, which implies strong linguistic norms and in particular strong concern for the purity of proper literary language.

In future research, one could explore the processing of marginal words with regard to standard and nonce words even further – by means of such technologies as eyetracking and reaction times. In particular, we could test how speakers behave when exposed to marginal words: do they slow down or back up, and whether their behavior is more parallel to their reaction on standard or nonce stimuli.

Chapter 10

Conclusions

In this chapter I bring together my findings and summarize the contribution of this dissertation.

What do we learn about allomorphy from this work? I suggest that the phenomenon of allomorphy is broader than is traditionally assumed. The traditional understanding of allomorphy is based on absolute criteria of identical meaning and complementary distribution of morpheme variants. I have tried to show that this model of allomorphy is a theoretical construct, an idealization, “abstracted away” from the empirical phenomenon that we can observe in linguistic data. The data shows that submorphemic semantic differences and distributional overlap are not uncommon properties of morpheme variants and are attested even for those allomorphs that linguists agree on. In this dissertation I have tried to “undo” this idealization of the model of allomorphy by confronting the theory with data. As a result, I propose a more realistic and accurate theoretical model of this phenomenon that is flexible enough to capture both clear standard cases and deviations.

What exactly does the new model of allomorphy suggest? First of all, it suggests that this phenomenon has a gradient nature and therefore can be best described in terms of a radial category with a central prototype, standard exemplars and non-standard deviations.

The Prototypical Allomorphic relationship is characterized by the closest and most automatic association of formants. Such a relationship is typically manifested in the case of phonologically conditioned allomorphs that are perfectly identical in terms of semantics, even if they have more than one meaning. An example of such Prototypical Allomorphy was given in 3.2.2 in the analysis of the prefixes RAZ- ~ RAS- ‘apart’ conditioned by the active, productive and exceptionless phonological process of regressive voicing assimilation in Modern Russian.

I further suggested that allomorphy conditioned by factors other than (or in addition to) active phonology are in principle less prototypical because these factors make the alternation of allomorphs less straightforward and automatic. These factors can belong to the levels of morphophonology, morphology, semantics, register, or history. Moreover, the distribution can be governed by a mixture of several competing factors. However, as long as the criteria of semantic “sameness” and distributional complementation are perfectly satisfied, we are dealing with Standard Allomorphy. An example of such a Standard (but not Prototypical) Allomorphic relationship was given in 3.3.1 in the account of the vocalization of consonant-final Russian prefixes (RAZ- ~ RAZO- ‘apart’) conditioned by interacting phonological and morphophonological factors.

Next, contrary to the common understanding of allomorphy, I extend this term to non-standard cases which represent deviations from the two criteria. I define Non-Standard allomorphy as a relationship of forms that fail to satisfy either the criterion of identical meaning or the criterion of complementary distribution or both criteria. These violations can take place due to partial distributional overlap of formants or partial semantic dissimilarity. Despite these deviations, there should be strong evidence suggesting that these formants are connected in a single perceptible morpheme. In particular, they should exhibit strong semantic similarity which exceeds their semantic

divergence. Moreover, these forms should have a robust pattern of distribution. Both of these characteristics should be ideally justified in terms of statistically robust generalizations over extensive data sets. A powerful technique that can incorporate large amounts of data and identify significant factors which govern the distribution is statistical modeling. Statistical modeling offers a different dimension of data analysis which can incorporate all data available and detect the core trends that organize it. In a way, a statistical approach to allomorphy mediates between the idealized criteria and the complex realities of empirical data. In other words, statistical models make it possible to measure deviations from the idealized criteria, revealing relationships that are nevertheless important, meaningful, and would otherwise go undetected. Thus, we can capture the key properties of data distribution and reveal a relationship among forms, if it is there.

Non-Standard Allomorphy is largely an unexplored terrain. In this dissertation I have looked at pairs of prefixes that might be candidates for Non-Standard Allomorphic relationship: S- vs. SO- 'together' (Ch. 4), O- vs. OB- 'around' (Ch. 5), PERE- vs. PRE- 'across' (Ch. 6), VZ- vs. VOZ- 'up' (Ch. 7), and VY- vs. IZ- (Ch. 8). In addition, I presented a peculiar case of a strong and close relationship of rival prefixes that, as I argue, are non-allomorphic and should be rather characterized in terms of distinct morphemes – O- vs. U- (Ch. 9).

Table 1 summarizes all the case studies discussed in this dissertation. Column (i) lists all analyzed pairs of prefixes, column (ii) specifies the chapter where they are addressed, (column iii) shows how many verbs with these prefixes were collected. Columns (iv) – (vii) address the criteria that are relevant in making subtle distinctions between these cases. Finally, column (viii) identifies the status of each pair of prefixes in terms of subtypes of allomorphy.

I propose that all examined pairs of prefixes can be located on a scale presented in column (viii) of Table 1. The scale is a range of different degrees of closeness of mutual relations that prefixes can have – from “regular” allomorphic status within a single morpheme (the top extreme) to distinct morphemes that lack any systematic association with one another (the bottom extreme). The top extreme of the scale is Prototypical Allomorphy represented by RAZ-/RAS- which are purely phonologically conditioned. The next level on the scale is Standard Allomorphy that we find in RAZ-/RAZO-. It is followed by Non-Standard Allomorphy observed in cases of S-/SO-, O-/OB-, PERE-/PRE-, VZ-/VOZ-, and VY-/IZ-. Each pair of prefixes exhibits certain semantic differences and partial distributional overlap, i.e. verbs that can attach both prefixes. In the four former cases, I examined data that resulted from the historical process of semantic and distributional divergence of former phonological variants of a single morpheme. In each case I observed that the distribution has motivated some semantic differences among the alternants. If a morpheme’s variants are conditioned not only phonologically but also morphologically, this leads to strong distributional and semantic specification of the variants, which can result in their divergence and even a complete morphological split as the end-point of this process. I suggest that the pairs of prefixes VZ-/VOZ- and VY-/IZ- represent borderline cases, having a hybrid status between Non-Standard Allomorphy and Non-Allomorphy. VZ- and VOZ- share the spatial prototype ‘Up’ but differ in terms of the scale of the trajectory: the former suggests a short path, whereas the latter implies a long path of the same motion.

i	ii	iii	iv	v	vi			vii		viii
Prefixes	Ch.	Number of verbs	Formal similarity	Etymological relationship	Semantics			Distribution		Status
					# of shared submeanings	Shared prototype	Distinct profiles	Size of overlap	Conditioning factors	
RAZ-RAS-	3	200	similar	related	share all 7 submeanings	share 'APART'	No	no overlap	phonology	Prototypical allomorphy
RAZ-RAZO-	3	210	similar	related	share all 7 submeanings	share 'APART'	No	no overlap	phonology & morphophonology	Standard allomorphy
S-SO-	4	1,156	similar	related	share all 6 submeanings	share both 'DOWNWARD' & 'CENTRIPETAL'	Yes: in CONCOMIT. ACTION	15 minimal pairs	phonology, morphophonology, register, semantics	Non-Standard allomorphy
O-OB-OBO-	5	1,037	similar	related	share all 15 submeanings	share 'AROUND'	Yes: spatial vs. change-of-state	23 minimal pairs	phonology, semantics (type of base), prosody	Non-Standard allomorphy
PERE-PRE-	6	945	similar	related	share 8 out of 14 submeanings	share 'TRANSFER OVER/ACROSS'	Yes: spatial vs. intensity	22 minimal pairs	grammatical classes: verbs vs. non-verbs	Non-Standard allomorphy
VZ-VOZ-	7	384	similar	related	share all 9 submeanings	share 'UPWARD' but differ in <i>height</i>	Yes: spatial, metaphor., aspectual	21 minimal pairs	semantics, register, aktionsart	Non-Standard allomorphy (borderline case)
VY-IZ-	8	998	not similar	different sources	share 10 out of 12 submeanings	share 'OUT OF', but don't share 'ZIGZAG'	Yes: 'OUT OF' vs. 'EXHAUST'	112 minimal pairs	semantics, register (prosody)	Non-Standard allomorphy (borderline case)
O-U-	9	155	not similar	unrelated	share the submeaning 'make X be Y'	different prototypes 'AROUND' and 'AWAY'	Not applicable	17 minimal pairs	qualitative vs. relational adjectival base	Non-Allomorphy; Closely associated rival morphemes
PRE-PRI-PRED-	6	10	similar	unrelated	no shared submeanings	different prototypes	Not applicable	some overlap	different semantics	Non-Allomorphy; Different morphemes with no association

Table 1: The overview of Russian aspectual prefixes discussed in this dissertation.

VY- shares the prototypical meaning 'OUT OF CONTAINER' with the loan IZ-, and their relationship can arguably be viewed as Non-Standard Allomorphy. At the same time, VY- does not share the spatial image schema 'ZIGZAG' of the "other" IZ-, native to Russian, so they should rather be analyzed as distinct morphemes. What is crucial for both borderline cases is the complex situation with regard to the prototype that is shared by the prefixes only to some degree.

Non-Standard Allomorphies are followed on the scale by Non-Allomorphy. Here I have located the case of O-/U- which is a pair of distinct morphemes with different spatial prototypes. At the same time, there is a strong association between the rival prefixes within this pair. In this sense, they are different from prefixes that have no association with one another. The bottom extreme of the scale is represented by the prefixes PRE-, PRI- and PRED- that were briefly discussed in Chapter 6. Despite their formal phonological similarity, there is nothing that ties them together in their modern use.

What are the parameters that have weight in assessing the status of candidates for allomorphy? I have addressed these parameters in Table 1 (Columns (iv) – (vii)) and explicate them below in the text.

Formal similarity

- Are the formants similar in their formal phonological shape?
[This property is expected but not obligatory. Formants that are similar in their phonological shape are more likely to be associated with each other.]

Etymological origin

- Are the formants historically related?
[This property is expected but not obligatory. A shared diachronic source often facilitates or motivates close semantic connection of the candidates as well as their phonological similarity.]

Semantics

- Are the formants polysemous or do they have only one meaning?
- Do these formants share all of their submeanings, the majority of submeanings or a minor part of them?
- What predominates – semantic similarity or semantic difference?
- What is the status of shared and non-shared senses?
[Semantic difference in several individual submeanings might be sub-morphemic difference; Semantic difference in the central prototypical sense rather suggests distinct morphemes.]
- Do these formants share the central sense – the semantic prototype?
[Sharing the spatial prototypical sense strongly points towards semantic similarity and allomorphy; Different prototypes or different construals of the same prototype suggest distinct morphemes.]
- Do these formants have distinct profiles in their semantics? If the difference in distributions across different senses is statistically significant – how large is the effect size?

Distribution

- Which factors determine the distribution of the prefixes: phonology, morphology, history, stylistics, semantics?

- Do the formants have complementary distribution or do they tolerate any overlap?
- How large is the area of the distributional overlap?
[The smaller the area of overlap, the more allomorphic is the relationship.]
- In contexts where the formants overlap, do they show any contrast in terms of semantics or register or are they completely interchangeable with one another?

Mixed conditioning

- Do semantic factors compete with other factors in conditioning the choice of the formant?
- Which factor has more power?
[Predominance of phonological or grammatical factors points towards allomorphy; Predominance of semantic factors points towards distinct morphemes.]

What does the recognition of Non-Standard Allomorphy buy us? This is a crucial amendment to the traditional all-or-nothing model of allomorphy. It means that the two definitional criteria can be violated but the allomorphic relationship is still there. In this sense, the canonical convention of a purely identical meaning and a perfect complementary distribution is no longer decisive in assessing allomorphic status. Instead, we accept that semantic “sameness” and complementarity are scalar variables that can vary and can be measured. In this light, there is no more need to limit ourselves to clear unproblematic cases that satisfy both semantic and distributional criteria. We are free to look at clear cases from a broader perspective and use them as a standard to evaluate data that is less clear and more complex. Such complications involve multiple factors that might conflict or interact with one another while affecting the distribution of forms. We can account for polysemous formants, which might share some of their submeanings and diverge in other ones, or even become specialized for specific uses.

The case studies that I have examined in this dissertation present a four-fold contribution. First, I provide comprehensive datasets that contain abundant information and can be used in future research. Second, I present detailed accounts of specific data and provide non-trivial conclusions about persistent issues of Russian derivational morphology. Nearly all of my analyses present the first corpus-based account of the issue at hand. Third, these case studies have a methodological value. They are examples that demonstrate applications of concrete techniques of data analysis including semantic modeling of polysemy and statistical modeling of multifactorial dependencies. Last but not least, my findings can be used in teaching Russian as both a native and a foreign language.

This dissertation opens up new directions for research. First, I recommend exploring Non-Standard Allomorphic relations in other languages. It is still an open question as to how common such properties of allomorphs like submorphemic differences and distributional overlap are. Also, it is fruitful to compare my findings on allomorphy of Russian prefixes with corresponding morphemes in other Slavic languages. Thirdly, one should look at whether the Non-Standard Allomorphy in aspectual prefixes has had an impact on the historical development of the Russian aspect. Fourth, revising the model of allomorphy in terms of a radial category suggests that other linguistic notions might need rethinking and reevaluation. It is crucial to bring our theory of language in general and our models of concrete linguistic phenomena in

particular back to the data, make them more psychologically realistic, accurate, and elaborate. In other words, to “undo” the idealization.

Appendix 1

Overview of data explored in this dissertation

Chapter	Prefixes	Type of data	Number of verbs
Chapter 3	RAZ-, RAS-, RAZO-	Modern Russian verbs	210
Chapter 4	S-, SO-	Modern Russian verbs	998
Chapter 5	O-, OB-, OBO-	Modern Russian verbs	1,037
Chapter 6	PERE-, PRE-	Modern Russian verbs	945
Chapter 7	VZ-, VOZ-	Modern Russian verbs	384
Chapter 8	VY-, IZ-	Modern Russian verbs	989
Chapter 9	O-, U- in factitive verbs	Modern Russian verbs	155
Total:	15 prefixes		4,718

Chapter	Experiments	# datapoints (responses of subjects)
Chapter 5	O-, OB-	3,878
Chapter 9	O-, U- in factitive verbs	7,260
Total:		11,138

Databases and R scripts for statistical analyses are available at the Tromsø Repository of Language and Linguistics (TROLLing) (<http://opendata.uit.no/>).

The direct link is <http://hdl.handle.net/10037.1/10078> (all files are located at *Data & Analysis*).

The full citation is

Endresen, Anna, 2014, "Non-Standard Allomorphy in Russian Prefixes: Corpus, Experimental, and Statistical Exploration", <http://hdl.handle.net/10037.1/10078> UiT Open Research Data [Distributor] V1 [Version]

Appendix 2

Expansion of the prefix S- in Modern Russian

The table below presents data showing that the prefix S- can replace ten other prefixes with no considerable change in the meaning of the resulting verb.

The table is organized as follows. The leftmost column lists the prefix that is replaced with the prefix S-. Neighboring columns provide the standard prefixed verb, its S-prefixed equivalent, and the English gloss for both verbs. The rightmost column illustrates the use of S-derivatives with representative examples culled from the RNC.

Pref.	Standard Russian verb	Formation with S-	English gloss	Example from the RNC
za-	<i>zaintrigovat'</i>	<i>syntrigovat'</i>	'intrigue'	<i>Vy ne spjaneete bystro? – Ja ulybnulsja.</i> [V. Makanin. <i>Andegraund, ili geroj našego vremeni</i> (1996-1997)] 'Won't you get drunk quickly? – I smiled.'
	<i>zamerznut'</i>	<i>smerznut'</i>	'freeze'	
	<i>zaxotet'</i>	<i>sxotet'</i>	'start wanting'	
	<i>zap'janet'</i>	<i>sp'janet'</i>	'become intoxicated'	
pod-	<i>podkaraulit'</i>	<i>skaraulit'</i>	'guard someone, find the right moment'	<i>Nelli Sergeevna <...> skaraulila moment, kogda svekrov' byla vo dvore...</i> [V. Astaf'ev. <i>Proletnyj gus'</i> (2000)] "Nelli Sergeevna found the right moment , when her mother-in-law was in the yard..."
ot-	<i>otporot'</i> (<i>pugovicu</i>)	<i>sporot'</i> (<i>pugovicu</i>)	'remove (e.g. a button) from fabric'	<i>Komandoval imi vysokij belokuryj krasavec, odetyj v šinel' so sporotymi našivkami.</i> [V. Grossman. <i>Žizn' i sud'ba</i> (1960)] "Their commander was a blond handsome man dressed in a coat with stripes that had been removed. " <i>Podskazite, pliz, gde možno sprazdnovat' pervyj den' roždenija?</i> [Detskie razvlečenija (forum) (2006)] "Could you give me some tips, please, where one can celebrate a birthday for a one-year-old?"
	<i>otprazdnovat'</i>	<i>sprazdnovat'</i>	'celebrate'	
	<i>otrepetirovat'</i>	<i>srepetirovat'</i>	'rehearse'	
vz-	<i>vspotet'</i>	<i>spotet'</i>	'sweat'	<i>Ves' xutor na pristup šel: – Ty čto narod sbalamutil!</i> [B. Ekimov. Čikomasov (2001)] "The whole village was coming into attack. – How come you troubled the people!" <i>Eželi k obedu ne spašeš' polosu, pridu – uši vse do odnogo oborvu.</i>
	<i>vskipjatit'(sja)</i>	<i>skipjatit'(sja)</i>	'boil (trans.)'	
	<i>vskipet'</i>	<i>skipet'</i> (<i>about čajnik, samovar</i>)	'boil (intrans.)'	
	<i>vzbalamutit'</i>	<i>sbalamutit'</i>	'trouble, stir up, make	

			nervous'	[V. Belov. Plotnickie rasskazy (1968)] 'If you don't plow up the row by lunchtime, I will come and tear off all your ears.'
	<i>vzbesit'sja</i>	<i>sbesit'sja</i>	'become crazy'	
	<i>vskopat'</i>	<i>skopat'</i> (<i>ogorod</i>)	'dig up, dug'	
	<i>vspaxat'</i>	<i>spaxat'</i>	'plow, furrow'	
	<i>vspomnit'</i>	<i>spomnit'</i>	'remember'	
	<i>vspuxnut'</i>	<i>spuxnut'</i>	'swell'	
pri-	<i>prigodit'sja</i>	<i>sgodit'sja</i>	'be used for smth'	[N]a balkone <...> <i>stojal ogromnyj škaf, v ktorom <...> xranilis' šurupy, gvozdi i vsjakaja metalličeskaja meloč', sposobnaja kogda-nubud' sgodit'sja: v svobodnoe vremja papa masteril mebel'</i> [S. Spivakova. Ne vsě (2002)] 'On the balcony, there was a huge cupboard where we kept screws, nails, and metal trifles that might some time be utilized : in his spare time, dad was making furniture.'
	<i>prigotovit'(sja)</i>	<i>sgotovit'(sja)</i>	'cook, prepare oneself'	– <i>Xočeš', ja priedu sgotovlju čto-nibud'?</i> – predložila Anna. [V. Tokareva. Svoja pravda // «Novyj Mir», 2002] 'If you want, I can come and cook something', Anna suggested.'
po-	<i>pogibnut'</i>	<i>sgibnut'</i>	'perish'	Byl u menja drug <...> <i>Žal', sgib gde-to, govorjat.</i> [A. Ivanov. Serdce Parmy (2000)] 'I had a friend. Unfortunately, he died somewhere, they say.' <i>On stal osvivaivat'sja, sžaril jaičnicu, vyslušal po radio treskurečuju reč'...</i> [A. Azol'skij. Lopušok // «Novyj Mir», 1998] 'He started making himself at home, fried eggs, listened to someone's crackling speech on the radio...' <i>Prišla i ego pora u Poliny sčast'ja spytat'.</i> [Vinogradnaja loza (2001)] 'The time was right to try his luck with Polina.'
	<i>poprobovat'</i>	<i>sprobovat'</i>	'try'	
	<i>povelet'</i>	<i>svelet'</i>	'command'	
	<i>poborot'</i>	<i>sborot'</i>	'conquer down'	
	<i>pokvitat'sja</i>	<i>skvitat'sja</i>	'settle accounts, pay the debt'	
	<i>ponadobit'sja</i>	<i>snadobit'sja</i>	'become needed once'	
	<i>poznakomit'sja</i>	<i>soznakomit'sja</i>	'become acquainted'	
	<i>požarit'(sja)</i>	<i>sžarit'(sja)</i>	'fry'	
	<i>popytat'(sčastja)</i>	<i>spytat'</i>	'try'	
iz-	<i>ispugat'sja</i>	<i>spugat'sja</i>	'become afraid'	<i>On vybralsja iz kabiny i kriknul: «Ilja Petrovič', <...> ničego ne sdelaješ', sžarilsja ves'.</i> [O. Pavlov. Stepnaja kniga (1990-1998)] 'He climbed out of the car and shouted: Ilja Petrovič', there is nothing one can do, I am all overheated .'
	<i>izžarit'sja</i>	<i>sžarit'sja</i>	'become overheated'	
	<i>istratit'</i>	<i>stratit'</i>	'spend'	
	<i>ispeč'(sja)</i>	<i>speč'(sja)</i>	'cook, bake'	
	<i>iskoverkat'</i>	<i>skoverkat'</i>	'distort, deform,	

			mispronounce a word'	<i>Xočeš' ja tebe pirožkov speku.</i> [A. Mišarin. Belyj, belyj den' // «Oktjabr'», 2003] 'If you want, I will bake some stuffed buns.'
	<i>ispačkat'</i>	<i>spačkat'</i>	'make dirty'	
	<i>izurodovat'</i>	<i>surodovat'</i>	'disfigure, uglify'	
	<i>ispolnit'</i>	<i>spolnit'</i>	'perform a task, fulfil'	<i>Tol'ko karakter ego sportilsja, ne slušaet menja!</i> [V. Oseeva. Dinka proščaetsja s detstvom (1969)] 'But his character is ruined , he does not listen to me!'
	<i>isportit'(sja)</i>	<i>sportit'(sja)</i>	'spoil'	
	<i>izvajat'</i>	<i>svajat'</i>	'form, sculpture, cast'	
	<i>iskupat'(sja)</i>	<i>skupnut'sja</i>	'take a quick swim'	<i>[K]ul'tovaja figura byla svajana rukoj skul'ptora-professionala.</i> [N. Voronel'. Bez prikras. Vospominanija (1975-2003)] 'The iconic figure was sculpted by a professional.'
	<i>istopit'(sja)</i>	<i>stopit'(sja)</i>	'heat'	
	<i>istomit'sja</i>	<i>stomit'sja</i>	'languish until being worn out'	
	<i>iskrošit'</i>	<i>skrošit'</i>	'crumble'	– <i>Skaži svoemu, čtob banju stopil!</i> [I. Adamackij. Utešitel' // «Zvezda», 2001] 'Tell your guy to heat up the sauna!' <i>Naposledok, polnost'ju skrošiv malen'kij ostatok melka, obrabotal xolodil'nik.</i> [M. Butov. Svoboda // «Novyj Mir», 1999] 'The last thing I did was to crumble the short remainder of the chalk and treated the refrigerator [with it].'
u-	<i>uvidet'sja</i>	<i>svidet'sja</i>	'see each other once'	– <i>Menja, konečno, za èto nado s raboty svol'njat', – prežnim golosom skazal Aniskin.</i> [V. Lipatov. Derevenskij detektiv (1967-1968)] 'No doubt, one should fire me for this, Aniskin said in the same voice.'
	<i>uvolit'(uvol'njat')</i>	<i>svol'njat'</i>	'fire, dismiss from work'	
	<i>ukrast'</i>	<i>skrast'</i>	'steal'	– <i>Potom proguljaemsja, skradem v noči kakuju-nibud' virtual'nuju relikviju, razdelim dobyču, i ja pojdu tvorit'.</i> [D. Karalis. Roman s geroinej // «Zvezda», 2001] 'Afterwards, we would go for a walk, steal some virtual relic at night time, split the booty, and I would go to do some work.'
vy-	<i>vyprjavit'(sja)</i>	<i>sprjavit'(sja)</i>	'straighten'	<i>Posredi ètogo počti prijatel'skogo razgovorca ruka ego rezko sprjamilas' v lokte...</i> [O. Pavlov. Delo Matjušina (1996)] 'In the middle of this almost friendly conversation, his arm abruptly straightened at the elbow.'
	<i>vykurit'(pačku sigaret)</i>	<i>skurit'</i>	'smoke all the supply; degradate under constant smoking'	

				<p><i>K koncu rabočego dnja, vypiv nemerennoe količestvo čašek kofe i skuriv počti pačku sigaret, Vasilij Semenovič nakonec privel v porjadok pas'jans...</i> [P. Galickij. Cena Šagala (2000)] 'Toward the end of day, having drunk numerous cups of coffee and smoked near a pack of cigarettes, Vasilij Semenovič finally put the solitaire in order.'</p>
pro-	<i>promoknut'</i>	<i>smoknut'</i>	'become wet'	<p>– <i>Musečka, u baby vse plat'je na spine v puzyrjax. Sutjužit' nado.</i> [T. Mospan. Podium (2000)] 'Sweetie, this woman's dress is all bunched up on her back. One should do the ironing.'</p> <p><i>[Š]erst' vsja smokla v zol'noj židkosti.</i> [Kratkaja ènciklopedija skornjaka. 1999] 'The fur has become wet in the lime liquor.'</p>
	<i>proutjužit'</i>	<i>sutjužit'</i>	'iron flat'	

Appendix 3

Experiment on O- and OB-: Stimuli

Narratives from the Questionnaires A and B (verbal stimuli; stress marks according type A)

1. **Копáть** – рыть, заниматься земляной работой.

Необходимо было провести работы по изоляции фундамента, однако после случившегося наводнения все здание кругом по периметру было невозможно.

2. **Черту́ть** – проводить линию, черту.

Чтобы сорвать папоротник, нужно в ночь Ивана Купала разостлать около растения священную скатерть, вокруг себя на земле ножом круг, окропить папоротник святой водой и молиться. Сорвав цветок, нужно спрятать его за пазуху и бежать без оглядки.

3. **Гу́звить** – (о птице) прихрамывать, притворяясь, что одно крыло сломано.

Птица заметила лису и стала манить ее прочь от гнезда. Приподняв одно крыло, как будто оно было сломано, птица отбежала чуть подальше, вокруг камня, и, подождав, когда лиса последует за ней, взмыла вверх.

4. **Тка́бить** – перемещаться на вертолете.

Чтобы взглянуть на критическую ситуацию собственными глазами, глава правительства несколько раз вокруг метеорологической станции и рассудил, что пора принять экстренные меры.

5. **Рва́ть** – выдергивать с силою, резким движением.

Когда все нужные детали отшлифованы, необходимо защитную пленку шириной 4-5 см по всему *периметру изделия*.

6. **Ду́ктить** – кое-как с непривычки перемещаться на высоких каблуках.

Цокая высокими каблуками, дама прошла вдоль скамейки, кое-как вокруг клумбы, вдруг зацепилась за что-то и чуть не упала.

7. **Ра́жнить** – перемещаться на индийском слоне.

Согласно древней индийской традиции, жених должен подъехать к дому своей невесты верхом на слоне, торжественно вокруг дома и поставить слона на одно колено.

8. **Ю́пить** – передвигаться на лыжах без лыжных палок, энергично работая руками.

Сегодня юпить на лыжах было трудновато: ветер дул в лицо, да еще гололед. Так что вокруг леса я только один раз, зато сто раз пожалел, что палки с собой не взял.

9. **Ша́клить** – перемещаться верхом на верблюде.

Согласно древней легенде, если вокруг египетской пирамиды Хеопса в Гизе ровно девять раз, то загаданное в этот день желание обязательно исполнится.

10. **Мо́млить** – передвигаться, весело пританцовывая.

В тот вечер Андрей Иванович был в лучшем расположении духа. Сияя лучезарной улыбкой, он вокруг рояля, на котором играла Лизавета, затем подсел к ней поближе и стал переворачивать ноты.

11. **Га́бить** – передвигаться очень медленно, приставляя пятку одной ступни к носку другой.

По дороге ему встретился гигантский камень больше человеческого роста. Он вокруг камня, потрогал подошвой его мшистый бок и продолжил путь.

12. **Е́хать** – передвигаться на колесах.

Те, кто сдавал на права, знают, что одно из первых заданий на экзамене – аккуратно вокруг столба.

13. **Кóчлить** – перемещаться вперед на одной ноге.

Мишка часто соревновался со Славиком, кто сможет быстрее вокруг песочницы в одну сторону и обратно.

14. **Жра́пить** – с трудом перемещать на веревке ящик с боеприпасами.

Силы Павла были на исходе, однако он собрал все мужество и вокруг землянки.

15. **Но́крить** – передвигаться на коленках.

Маша боролась с сорняками около часа. Она не один раз по периметру вокруг всей морковной грядки, пока добилась, чтобы на ней не осталось ни одного сорняка.

16. **Ползту́** – передвигаться на животе.

Змея страшно зашипела. Она медленно вокруг веранды и, остановившись, долго смотрела на мангуста.

17. **Ца́вить** – перемещаться на пиратской шхуне.

Наш корабль вокруг Австралии всего за два месяца.

18. **Ху́шнить** – передвигаться, пиная перед собой футбольный мяч.

Ведя мяч к воротам, Кирсанов вокруг противника, пробежал еще несколько шагов и неожиданно для всех вдруг забил гол.

19. **Ви́ть** – скручивать, сплетать.

Я думал, как усажу ее рядом с собой на скамейку, руку вокруг ее талии и поцелую.

20. **Гно́рить** – передвигаться на руках вверх ногами, улыбаясь от счастья.

Когда я вдруг узнал, что меня приняли на первый курс, я готов был вокруг всего здания консерватории, крича на всю улицу о своем успехе.

21. **Па́тлить** – перемещаться на лодке.

Увидев прямо по курсу дом, стоящий на воде, мы подгрести поближе и решили вокруг него, чтобы посмотреть, есть ли под ним хоть маленький кусочек суши или нет.

22. **Чту́сить** – передвигаться, сильно шатаясь.

Данила шел на хутор долго и с трудом. Зато было, о чем потом рассказать: и о том, как он вылетел из кабака, и как приятно дул ветер в спину на раздолье, и как он, в конце концов, вокруг мельницы и наткнулся там на Емельяна.

23. **Зóприть** – перемещаться на коне.

Дворец был так велик, что даже на отличном английском скакуне нельзя было вокруг него за один день.

24. **Круж́ить** – передвигаться по кругу.

Стремление США Китай военными базами, поддержка независимости Тайваня, а также военное сотрудничество с Индией подталкивали Пекин к ответным мерам.

25. **На́дить** – передвигаться, держа в руках большой букет цветов.

Говорят, чтобы окончательно вскружить голову женщине, нужно вокруг нее более 200 раз.

26. **Спúлить** – перемещаться на водных лыжах.

Честно говоря, я думал что вокруг такого небольшого озера – пара пустяков. Однако, встав на лыжи, я убедился, что не все так просто.

27. **Тóвить** – перемещаться на одноколесном велосипеде.

Вдруг из-за кулис появился хохочущий клоун. Балансируя на одном колесе, он вокруг фокусника, выхватил у него цилиндр и достал из него еще двух кроликов.

28. **Ча́вить** – передвигаться пешком, играя на гитаре.

И нечего Вам всё вокруг дома моего ходить! Один раз Вы вокруг дома – ну, думаю, ладно, а Вы – и другой, и третий! Нужно же и честь знать!

29. **Плестú** – перевивать, соединяя в одно целое.

Когда-то, в дополимерную, допенопластовую эпоху, почти единственным способом защитить бутылку от ударов было ее камышом или соломой.

30. **Зúпить** – перемещаться в санях на собачьей упряжке.

Ему снилось, что он уже стремительно зупит на упряжке из десяти лохматых хаски в направлении долгожданного Северного полюса, и вот ему остается преодолеть последнюю сотню метров, победно вокруг заветной точки притяжения и закрепить гордый флаг завоевателя так, чтобы его не унесло ветром.

31. **Лóприть** – передвигаться, громко топая ногами.

В бане живет банник. Это особый банный дух, лохматый и вредный. Он любит шпарить кипятком и кидаться камешками из печки. Поэтому перед тем как идти париться, говорят, надо банника припугнуть, а для этого нужно вокруг бани, покрикивая и грозя веником.

32. **Ва́ждить** – перемещать коляску с ребенком.

Фонтан был так велик, что когда Марина вокруг него, малыш уже мирно спал.

33. **Бежа́ть** – двигаться, быстро отталкиваясь от земли ногами.

Чтобы пройти нужную дистанцию, лыжники должны были вокруг снежного поля 111 раз.

34. **Ша́дрить** – перемещаться на метле.

Когда Баба-Яга увидела, что золотое яичко разбито, а курочка Ряба сбежала с Колобком, она вскочила на метлу, разок-другой вокруг своей избушки и отправилась в погоню.

35. **Лúсить** – тихонько перемещаться в свое удовольствие.
Что может быть лучше, чем поутру вылить себе на голову в ванной ведро ледяной воды, разок-другой вокруг стадиона, позавтракать овсяной кашей, а затем спешить на работу, улыбаясь яркому солнечному дню, который так приятно начался.
36. **Жúсклить** – передвигаться с закрытыми глазами.
Играли в жмурки. Миша вокруг комнаты, шаря перед собой руками, но никого не поймал. Все со смехом разбежались.
37. **Катúть** – везти какой-нибудь предмет, поставленный на колеса, или ехать самому.
Этот велосипед – вещь историческая. На нем я дважды вокруг всего света!
38. **Вúрлить** – двигаться, легко лавируя между препятствиями.
При помощи нового автоматического управления на этом космическом корабле можно вокруг земного шара, ловко уклоняясь от встречных метеоритов.
39. **Жáхлить** – перемещаться на двухметровых ходулях.
Этой весной во время разлива Нила вода поднялась так высоко, что местные жители вынуждены были удлинить свои обычные ходули на полметра. Они с трудом добрались до храма Хапи, повелителя наводнений, вокруг храма, однако вход был затоплен водой и войти в него было просто невозможно.
40. **Скóлить** – перемещаться в лодке на веслах.
В поисках удобной бухты нам пришлось вокруг всего острова. Причалили мы, в конце концов, у северного мыса, поскольку окрестность выглядела вполне приветливо и рифов у берега видно не было.
41. **Гнúть** – изгибать, отклонять.
Они добежали до парка, по кругу танцплощадку и спрятались в зарослях за эстрадой.
42. **Рóглить** – перемещаться на роликовых коньках.
Мишке вчера купили новые ролики. Теперь для полного счастья нужно вокруг школы, чтобы все видели, скататься до магазина, несколько раз упасть на мягкие налокотники и заполучить парочку царапин, чтобы потом было чем похвастаться.
43. **Пúрить** – передвигаться, хватаясь лапами и хвостом за ветки деревьев.
Маугли поднял голову и увидел на деревьях с полдюжины обезьян. Самая крупная обезьяна вокруг него и быстро спустилась по лиане на землю.
44. **Ялить** – перемещаться по воздуху (о воздушном шаре).
Воздушный шар поднялся в небо. Он плавно проплыл по воздуху до самой горы, вокруг ее вершины и полетел дальше, чуть задевая встречные облака.
45. **Чúпить** – перемещаться со скоростью 70 километров в час.
Из пункта А выехал грузовик. Проехав расстояние 105 км., он по периметру вокруг леса, площадь которого имела форму квадрата и составляла 180 км².
46. **Вéять** – воздействовать легкой струей воздуха.
Расстегнув ремешок, он выпустил рубашку наружу и попытался влажное, разгоряченное тяжелой работой тело.
47. **Хóпить** – стремительно перемещаться, прыгая при этом на скакалке.

И вот командные соревнования начались. Первым заданием было добежать до середины дорожки, взять лежащую на земле скакалку, затем хопить что есть силы до корзины с кеглями, вокруг корзины, а после – спешить обратно к своей команде.

48. Мúрлить – передвигаться на цыпочках, изображая кошку.

У маленькой Любы просто талант изображать разных животных. Вчера она мастерски играла мышку, а сегодня, смотрю, – она уже мурлит в сторону папы, вокруг кресла, где он сидел, а потом вдруг прыг – к нему и как замыкает!

49. Жгáвить – перемещаться в мешке.

Ребята придумали новое развлечение: кто быстрее сможет вокруг газонокосилки тети Шуры.

50. Гнáть – быстро ехать, вести автомобиль на большой скорости.

Петров завел мотор, за пару секунд автомобиль вокруг здания районного центра, и они во всю мочь помчались по сухой проселочной дороге.

51. Знúпить – перемещаться в санях, запряженных северными оленями.

Дед Мороз спешил с подарками. Всё собирать по списку пришлось в последний момент, поэтому он прыгнул в сани, разок вокруг волшебного леса, удостоверился, что ни один заветный мешок он не оставил, и помчался на санях развозить подарки.

52. Кáмпить – передвигаться, вглядываясь в лица встречных людей.

Он доехал до станции метро, вышел на улицу и стал ждать Тоню. Сначала она опаздывала на 5 минут, потом на 10. Вдруг ему пришло в голову, что, может быть, Тоня уже приехала и ждет его у другого выхода. Он вокруг всего здания метро, но Тоню не встретил.

53. Тúлить – медленно перемещаться, от скуки пиная найденный на дороге камень.

Блок вышел из дома и стал тихонько тулить в сторону табачной лавки. Там ему предстояло ждать еще битых полчаса. Он уже успел вокруг лавки 4 раза, когда, наконец, из-за угла появился Анненский.

54. Плыть – передвигаться по поверхности или в глубине воды.

Остров был небольшим, поэтому мы решили его по *кругу* и присмотреть бухту получше.

55. Сúрить – перемещаться босиком.

Земля разъезжалась, трещины уходили вглубь, дуб уносило водой. Захар метнулся к дому, выскочил с верёвкой, накинул петлю на сучок отъезжающего дуба, вокруг дерева и стал изо всех сил тянуть, соединять землю.

56. Бóстить – перемещаться, считая шаги и измеряя длину участка.

Знаете, у землемеров есть своя технология для вычитывания площади участка. Это только кажется, что достаточно по периметру вокруг поля – и все дела. Это, извините меня, уже вчерашний день.

57. Дáмлить – воздушно вальсировать.

Юнкер подхватил Юлию и закружил ее в вихре вальса. Горели свечи, гремела музыка. Они несколько раз вокруг зала, потом голова у Юлии закружилась, и она без сил упала на диванные подушки.

58. Валúть – беспорядочно бросать, класть в большом количестве куда-либо.

Готовясь к зиме, крестьянину нужно избу кругом землей, утыкать мхом и защитить от стужи соломой.

59. **Бúклить** – передвигаться на задних лапах.

Когда приходят гости, Пампи любит показывать фокусы. Если покрутить перед ее носом кусочком печенья, она последует за ним и может даже вокруг журнального столика.

60. **Нестú** – перемещать, возводить.

Потом *город* завоевали римляне, они дали ему новое имя, *стеной* и разбили традиционную римскую планировку.

61. **Са́глить** – передвигаться задом наперед, изредка оглядываясь, чтобы не упасть.

Говорят, если саглить регулярно, то можно выработать приличное чувство равновесия. Я начала с того, что вчера два раза вокруг телевизора, но в результате чуть не уронила фарфоровую вазу, которая на нем стояла.

62. **Щúлить** – передвигаться, щурясь от солнечного света.

Яркое солнце слепило глаза. Эдик несколько раз вокруг машины, но так и не нашел предательски отвалившийся болтик.

Narratives from the Questionnaire C (adjectival stimuli)

1. **Легкий** – исполняемый, преодолеваемый без большого труда, усилий.

В прошлом году библиотека закупила дополнительную партию книг для студентов, что разительно учебный процесс.

2. **Глухой** – не способный слышать.

Тут Игнат совсем рассвирепел: «Ты что, глухой? Не слышишь, что ли, что я тебе говорю?» А Федор, усмехаясь, проговорил в бороду: «Да ты меня, брат, своим криком совсем».

3. **Надый** – требовательный к чистоте и порядку.

Когда дедушка узнал о приезде Паши, он стал таким надым, что все просто диву дались. А сколько разговоров потом было, что это известие его так сильно

4. **Тулый** – неуверенный в себе.

Катерина Николаевна всю ситуацию представила совсем в ином свете, так что чиновник оказался во всей этой истории робким, нерешительным и тулым человеком. Да-да, именно тулым. Хотя заметьте, что это Катерина Николаевна сделала его тулым, а иначе сказать – его.

5. **Нокрый** – умеющий хорошо готовить.

Кирилл всегда мечтал стать отличным поваром, однако то, что он готовил, есть никто не мог. После школы кулинаров его словно подменили! За пару месяцев его научили и варить, и жарить, и выпекать, сделали нокрым и подкованным в разных технологиях, то есть его.

6. **Общий** – содержащий только самое существенное, без подробностей.

Перед Кириллом Викторовичем стояла непростая задача: все результаты полугодовой работы в лаборатории профессора Ильинского теперь нужно было суммировать и представить в общем виде, то есть максимально

7. **Чупый** – немного выпивший.

Этот напиток очень освежает, утоляет голод и жажду, вызывает небольшой пот и делает сонливым после утомительной езды. Однако, если он пролежит в погребе года два-три, то может даже, особенно такого непривычного, как Вы.

8. **Важдый** – испытывающий сильное чувство патриотизма.

Издавека ветер доносил звуки раздольной русской песни. Николай вышел в поле, вдохнул вечерней прохлады и остро ощутил, как родной мотив и простые слова сделали его глубоко важным, тронули его до слез, его.

9. **Гнорый** – употребляющий в пищу только сладкое.

Машуня у нас теперь не ест ни кашу, ни суп, ни второе! Только сладкое и ест! Это бабушкины сахарные крендели и рогалики с джемовой начинкой ее так Что теперь будешь с ней делать?

10. **Хопый** – испытывающий непреодолимую тягу к покупкам.

Каждая получка бросает Марианну в омут неутомимого шопинга. И в этот раз, как всегда, очередная зарплата изменила экономную Марианну до неузнаваемости: сделала ее кокетливой, озорной и хопой, иными словами, ее и закрутила в вихре новых магазинов и покупок.

11. **Злой** – полный злобы.

Вчерашний разговор до того меня, что я сегодня без волнения не мог думать о случившемся.

12. Гузвый – красочный.

Пришла осень и принесла с собой свежесть, ветер и новые краски. Осень раскрасила листья в золото и пурпур, застелила тропы мягким ковром, превратила лес в богато украшенный, гузвый терем. Осень-мастерица постаралась на славу, преобразила лес, его.

13. Товый – тоскующий по дому.

После короткого разговора с сестрой по телефону Артем вдруг весь ушел в себя, стал молчаливым и грустным. Всем было понятно, что это телефонный разговор его так

14. Жахлый – проводящий все время за чтением книг.

Если вы вдруг решите сделаться ходячей энциклопедией, то первым делом нужно обзавестись богатой библиотекой, которая могла бы вас

15. Круглый – имеющий форму круга.

Чтобы произнести немецкий звук ÿ в слове Мюнхен, нужно губы, как бублик.

16. Саглый – имеющий пристрастие к компьютерным играм.

Новая компьютерная игра «Юпитер», которая только что вышла в продажу, может любого, даже взрослого. Она так захватывает, что просто невозможно оторваться.

17. Щулый – легко обижающийся.

Постоянная критика со стороны учителей развила в Любе сильный комплекс неполноценности и сделала ее невероятно шулой девочкой, иными словами, ее.

18. Жгавый – имеющий навязчивую идею постоянно мыть руки.

Еще год назад Павел где-то прочитал о многочисленных микробах, населяющих поверхности различных предметов – овощей, фруктов, рук, мебели и пр. С тех пор он моет руки по сто раз в сутки, все кипятит и дезинфицирует. Ума не приложу, что за книга произвела такое сильное впечатление на его воображение, так его

19. Нагой – не имеющий на себе покрывала.

Елена Николаевна резко встала с кресел и направилась к двери. Накидка упала к ее ногам и белые плечи.

20. Чавый – с хорошими манерами поведения.

Когда мы нашли Мусю во дворе, она была совсем дикой и сначала только царапалась и кусалась. Но домашняя обстановка, молочко и манная кашка Мусю вскоре – она стала ручной и чавой.

21. Момлый – томно закатывающий глаза и вздыхающий.

Бесконечные дамские романы, которыми Роза зачитывалась в последнее время, ее

22. Американский – имеющий отношение к Америке.

Хеппи-энд – обязательный элемент голливудских фильмов. Однако такой счастливый поворот сюжета на американский манер уже давно перешагнул границы США. Чтобы

увеличить кинопрокат и потенциальную зрительскую аудиторию, даже Индия стремится свои современные фильмы.

23. Сурый – погруженный в уныние.

На выходных ребята хотели поехать кататься на роликах, однако погода была никудышная: с пятницы зарядил дождь. Рома и Славик повесили носы, а вот Павлика это совсем не, у него в запасе всегда было много затей.

24. Дамлый – испытывающий чувство ностальгии.

Серая дождливая погода, старые пластинки, в одиночестве проведенный вечер – все это заставило Станислава Николаевича мысленно перенестись в былые дни, сделало его дамлым и чувствительным – его.

25. Зопрый – имеющий выдающиеся музыкальные способности.

Регулярные занятия музыкой сильно развили Катин голос и слух – ее. Более того, теперь, когда она читала ноты, вокруг уже звучала музыка, наполняя душу звуками и переливами.

26. Дуктый – невесомый, воздушный.

Надю было просто не узнать. Она стала изумительно дуктой. Кажется, это новое платье и прическа ее так

27. Юпый – суеверный.

Недавно Галина Петровна вдруг сделалась донельзя юпой. А причиной тому был черный кот, который перешел ей дорогу. После того кота и посыпались на нее все несчастья: и каблук сломался по дороге в магазин, и сумку утащили, и автобус опоздал. Этот случай Галину Петровну, сделал ее грозой всех котов, особенно черных.

28. Мрачный – угрюмый.

Вчера Миша пришел домой из школы мрачный, как туча. «Миша, что тебя так?» – спросила его мама.

29. Патлый – забывчивый.

Раньше баба Шура все помнила, а теперь говорит: «Голова – решето, положишь в него, а из дырки и выпадет. Что поделаешь! Это старость меня!»

30. Жусклый – предающийся лени.

Постоянное сидение у телевизора не привело ни к чему, только Толика, то есть сделало его жусклым и неповоротливым, а это, в свою очередь, стало раздражать и маму, и бабушку, и Наташу.

31. Вшивый – имеющий много вшей.

Очередное летнее нашествие насекомых вместе с недостатком гигиены в походных условиях так солдат, что каждый новый переход заставлял их невыносимо страдать.

32. Спулый – не способный работать правой рукой.

Данила вылетел из кабака очень неудачно, повредил правую руку так сильно, что его Оставалось либо на время сделаться левшой, либо звать на покос соседей.

33. Ражный – чрезмерно расточительный.

Неожиданный выигрыш в воскресной лотерее Захара, и он решил отправиться на ярмарку и купить там подарки всем своим домашним.

34. **Немецкий** – свойственный немцам, характерный для них.
Четыре года, проведенные в Германии, заметно сержанта. Особенно это касалось его внешнего вида.
35. **Шадрый** – не способный видеть объемные предметы.
В автокатастрофе никто из пассажиров тяжело не пострадал, однако у Владислава случилось сотрясение мозга, которое его, сделав медлительным и рассеянным.
36. **Светлый** – не темный.
Оказывается, можно значительно волосы не только разными химикатами вроде перекиси водорода, а естественными народными средствами, например раствором крапивы и ромашки.
37. **Зупый** – никому не доверяющий.
Михаила столько раз обманывали на рынке, что он стал зупым и подозрительным. Да такое и любого бы, не только его.
38. **Габый** – восхищенный красотой природы.
Поездка в Монголию сделала Гришу истинным любителем походов и палаток, габым и неутомимым путешественником, его.
39. **Русский** – свойственный русским людям по языку, обычаям.
Жители Смоленской области изначально были белорусами, и только полстолетия назад их окончательно
40. **Кампый** – полный надежд и новых планов.
Возможность получить двойной отпуск не просто Никитина, а превратила его в жизнерадостного и энергичного человека.
41. **Знупый** – идущий полным ходом, интенсивный.
Англия стремилась добиться того, чтобы ее внешняя торговля стала процветающей и знупой. В этом стремлении внешнюю торговлю Англия учреждала в чужеземных странах колониальные владения.
42. **Бостый** – умеющий изготавливать красивую глиняную посуду.
Два года, проведенные в мастерской, не только помогли Емельяну развить нужную сноровку, но просто напросто его, сделали настоящим мастером своего дела.
43. **Голый** – без убранных.
Все картины, тарелочки и фотографии со стен сняли и упаковали в коробки. Мебель перевозили постепенно. Когда совсем стены, в комнате поселилось эхо.
44. **Шаклый** – разочарованный.
Костя что есть силы спешил на почту, поэтому то, что ее закрыли на полчаса раньше, его и повергло в глубокое уныние.
45. **Лусый** – не способный есть рыбу.
В детстве Вику так много кормили рыбой, что в результате ее, так что теперь на рыбу она смотреть не может.
46. **Жрапый** – берущий взятки.

Бюрократический дух и подобострастная атмосфера, царящие в организации, и начальника, сделав его корыстным и безнравственным.

47. **Мелкий** – незначительный по величине, размеру, стоимости.

Закрывать государственный исторический музей – значит умалить значение государственной истории, его подлинно национальное содержание и достоинство.

48. **Ткабый** – стеснительный.

Выйдя на сцену лицом к огромному залу зрителей, Поливанов почувствовал, что его , так что от стеснительности он был не в силах произнести ни слова.

49. **Лопрый** – стойкий, несгибаемый.

Многочисленные препятствия и жизненные трудности закалили его характер, укрепили волю, его, сделав стойким и решительным.

50. **Чтусый** – сосредоточенный на деталях.

Скрупулезная профессия инженера наложила отпечаток и на его характер: Муравьева, сделав еще к тому же дотошным, педантичным и пунктуальным.

51. **Буклый** – гордый своим успехом.

В погоне за карьерными достижениями Влад добился повышения по службе, и это придало ему уверенности в себе, самоуважения, его и сделало еще более амбициозным.

52. **Грубый** – жестокий, неучтивый, неделикатный в обращении.

В облике разных по характеру и возрасту фронтовиков А.Т. Твардовский показал, что война не их души.

53. **Ялый** – не способный переносить визг.

В детском садике дети всегда громко визжали. Поработав там несколько месяцев, Вероника уже не могла переносить детский визг, у нее начинала болеть голова, работа ее окончательно

54. **Мурлый** – плавно текущий, медленный.

В тот вечер Елена Николаевна баловала всех занимательными историями. Помню, однако, что в рассказе о катании на санках ее бесконечные подробности и детали значительно повествование, а мне не терпелось узнать, чем же дело кончилось.

55. **Пурый** – не переносящий транспорт.

Постоянные утомительные поездки на метро и автобусах, электричках и маршрутках туда и обратно, сделали Решетова хронически усталым, нервным и пурым, его, так что по выходным он предпочитал никуда не ездить.

56. **Кочлый** – полный спокойствия.

Андрей так переживал, что просто места себе не находил. Даша дала ему выпить настоя трав, и это сняло волнение, расслабило и Андрея, сделав его вдруг спокойным и невозмутимым.

57. **Вурлый** – полный инициативы.

Участие в новом инженерном проекте не только сделало из апатичного и рассеянного Виталия собранного и деловитого разработчика, но еще и его.

58. **Цавый** – неискренний, лицемерный.

Необходимость постоянно выслуживаться и угождать начальству
Молчалина, сделав его подлым и хитрым.

59. **Роглый** – фашистский.

Нищета и националистские настроения в Германии 30-х годов сделали многих людей сторонниками фашизма, их.

60. **Хушный** – привередливый.

Любу дома так избаловали, что в садике она теперь почти ничего не ест. Воспитательница говорит, что это дома Любу , вот она и стала такой привередой.

61. **Сколый** – знаменитый.

Изобретение микроскопа было скачком в развитии оптики. Оно сделало Антони ван Левенгука по-настоящему знаменитым, иными словами его.

62. **Живой** – полный жизни, энергии.

Приезд офицеров местное общество, кроме того, в городке появился генерал. Званые обеды, которые он часто устраивал, собирали много народа.

Appendix 4

Experiment on O- and OB-: Nonce words

#	Adjectival stimuli	Definition of the nonce adjective	Verbal stimuli	Definition of the nonce verb
1	<i>bostyj</i>	<i>umejuščij izgotovljat' krasivuju glinjanuju posudu</i> 'able to make beautiful dishes out of clay'	<i>bostit'</i>	<i>peremeščat'sja, sčitaja šagi i izmerjaja dlinu učastka</i> 'move counting steps and measuring the distance'
2	<i>buklyj</i>	<i>gordyj svoim uspešom</i> 'proud of one's success'	<i>buklit'</i>	<i>peredvigat'sja na zadnix lapax</i> 'move on back paws'
3	<i>vurlyj</i>	<i>polnyj iniciativy</i> 'full of initiative'	<i>vurlit'</i>	<i>dvigat'sja, laviruju meždu prepjatstvijami</i> 'move, maneuvering between obstacles'
4	<i>važdyj</i>	<i>ispytyvajuščij sil'noe čuvstvo patriotizma</i> 'experiencing a strong feeling of patriotism'	<i>važdit'</i>	<i>peremeščat' koljasku s rebenkom</i> 'move a carriage with a child'
5	<i>guzvyj</i>	<i>krasočnyj</i> 'colorful'	<i>guzvit'</i>	<i>(o ptice) prixramyvat', pritvorjajas', čto odno krylo slomano</i> '(of a bird) limp pretending that a wing is broken'
6	<i>gabyj</i>	<i>vosxiščennyj krasotoj prirody</i> 'fascinated with the beauty of nature'	<i>gabit'</i>	<i>peredvigat'sja očen' medlenno, pristavljaja pjatku odnoj stupni k nosku drugoj</i> 'move very slowly placing one foot to the front of the other heel to toe'
7	<i>duktyj</i>	<i>nevesomyj, vozdušnyj</i> 'weightless, lightweight'	<i>duktit'</i>	<i>koe-kak s neprivyčki peremeščat'sja na vysokix kablukax</i> 'move with difficulty on high heels'
8	<i>damlyj</i>	<i>ispytyvajuščij čuvstvo nostalgii</i> 'experiencing a feeling of nostalgia'	<i>damlit'</i>	<i>vozdušno val'sirovat'</i> 'waltz lightly'
9	<i>žaxlyj</i>	<i>provodjaščij vse vremja za čteniem knig</i> 'spending all time reading books'	<i>žaxlit'</i>	<i>peremeščat'sja na dvuxmetrovyx xoduljax</i> 'move on two-meter-long stilts'
10	<i>žusklyj</i>	<i>predajuščijsja leni</i> 'being lazy'	<i>žusklit'</i>	<i>peredvigat'sja s zakrytymi glazami</i>

				'move with eyes closed'
11	<i>zopryj</i>	<i>imejuščij vydajuščiesja muzykal'nye sposobnosti</i> 'having outstanding musical abilities'	<i>zoprit'</i>	<i>peremeščat'sja na kone</i> 'move on a horse'
12	<i>zupyj</i>	<i>nikomu ne doverjajuščij</i> 'trusting nobody'	<i>zupit'</i>	<i>peremeščat'sja v sanjax na sobač'ej uprjažke</i> 'move dog-sledding'
13	<i>jupyj</i>	<i>suevernyj</i> 'superstitious'	<i>jupit'</i>	<i>peredvigat'sja na lyžax bez lyžnyx palok, ènergično rabotaja rukami</i> 'move skiing without poles and intensively working the arms'
14	<i>jalyj</i>	<i>ne sposobnyj perenosit' vizg</i> 'not able to tolerate screaming'	<i>jalit'</i>	<i>peremeščat'sja po vozduxu (o vozdušnom šare)</i> 'move by air (about a hot-air balloon)'
15	<i>kočlyj</i>	<i>polnyj spokojstvija</i> 'full of calmness'	<i>kočlit'</i>	<i>peremeščat'sja vpered na odnoj noge</i> 'move forward on one leg'
16	<i>kampyj</i>	<i>polnyj nadežd i novyx planov</i> 'full of hopes and new plans'	<i>kampit'</i>	<i>peredvigat'sja, vgljadyvajas' v lica vstrečnyx ljudej</i> 'move staring at people's faces'
17	<i>lusyj</i>	<i>ne sposobnyj est' rybu</i> 'not able to eat fish'	<i>lunit'</i>	<i>tixon'ko peremeščat'sja v svoe udovol'stvie</i> 'move along at a comfortable pace'
18	<i>lopyj</i>	<i>stojkij, nesgibaemyj</i> 'firm, indestructible'	<i>loprit'</i>	<i>peredvigat'sja, gromko topaja nogami</i> 'move while loudly stamping one's feet'
19	<i>murlyj</i>	<i>plavno tekuščij, medlennyj</i> 'floating smoothly, slow'	<i>murlit'</i>	<i>peredvigat'sja na cypočkax, izobražaja košku</i> 'move on tiptoe pretending to be a cat'
20	<i>momlyj</i>	<i>tomno zakatyvajuščij glaza i vzdyxajuščij</i> 'languorously rolling eyes and sighing'	<i>momlit'</i>	<i>peredvigat'sja, veselo pritancovyvaja</i> 'move dancing joyfully'
21	<i>nadyj</i>	<i>trebovatel'nyj k čistote i porjadku</i> 'demanding everything to be tidy and in order'	<i>nadit'</i>	<i>peredvigat'sja derža v rukax bol'šoj buket cvetov</i> 'move carrying a big bouquet of flowers'
22	<i>nokryj</i>	<i>umejuščij xorošo gotovit'</i> 'able to cook well'	<i>nokrit'</i>	<i>peredvigat'sja na kolenkax</i> 'move on one's knees'
23	<i>puryj</i>	<i>ne perenosjaščij transport</i>	<i>purit'</i>	<i>peredvigat'sja, xvatajas' lapami i xvostom za vetki derev'ev</i>

		'not able to tolerate travel by vehicle'		'move grabbing branches of the tress by paws and a tail'
24	<i>patlyj</i>	<i>zabyvčivyj</i> 'forgetful'	<i>patlit'</i>	<i>peremeščat'sja na lodke</i> 'move by boat'
25	<i>roglyj</i>	<i>fašistskij</i> 'fascist'	<i>roglit'</i>	<i>peremeščat'sja na rolikovyx kon'kax</i> 'move on roller skates'
26	<i>ražnyj</i>	<i>črezmerno rastočitel'nyj</i> 'extremely wasteful'	<i>ražnit'</i>	<i>peremeščat'sja na indijskom slone</i> 'move riding an Indian elephant'
27	<i>saglyj</i>	<i>imejuščij pristrastie k komp'juternym igram</i> 'obsessed with playing computer games'	<i>saglit'</i>	<i>peredvigat'sja zadom napered, izredka ogljadyvajas', čtoby ne upast'</i> 'walk backwards, looking over one's shoulder occasionally so as not to fall'
28	<i>suryj</i>	<i>pogružennyj v unynie</i> 'dejected'	<i>surit'</i>	<i>peremeščat'sja bosikom</i> 'move barefoot'
29	<i>tulyj</i>	<i>neuverennyj v sebe</i> 'lacking self-confidence'	<i>tulit'</i>	<i>medlenno peremeščat'sja, ot skuki pinaja najdennyj na doroge kamen'</i> 'move slowly kicking a stone found on a road from boredom'
30	<i>tovyj</i>	<i>toskujuščij po domu</i> 'being homesick'	<i>tovit'</i>	<i>peremeščat'sja na odnokolesnom velosipede</i> 'move by unicycle'
31	<i>xopyj</i>	<i>ispytyvajuščij nepreodolimuju tjagu k pokupkam</i> 'shopoholic'	<i>xopit'</i>	<i>stremitel'no peredvigat'sja, prygaja pri ètom na skakalke</i> 'move fast jumping over a jump rope'
32	<i>xušnyj</i>	<i>priveredlivyj</i> 'picky'	<i>xušnit'</i>	<i>peredvigat'sja, pinaja pered soboj futbol'nyj mjač</i> 'move kicking a football ball in front'
33	<i>cavyj</i>	<i>neiskrennij, licemernyj</i> 'insincere, hypocritical'	<i>cavit'</i>	<i>peremeščat'sja na piratskoj šxune</i> 'move sailing a pirate ship'
34	<i>čupyj</i>	<i>nemnogo vypivšij</i> 'a little drunk'	<i>čupit'</i>	<i>peremeščat'sja so skorost'ju 70 kilometrov v čas</i> 'move at the speed of 70 km per hour'
35	<i>čavyj</i>	<i>s xorošimi manerami povedenija</i> 'with good manners'	<i>čavit'</i>	<i>peredvigat'sja peškom, igraja na gitare</i> 'move on foot while playing guitar'
36	<i>šadryj</i>	<i>ne sposobnyj videt' ob'emnye predmety</i>	<i>šadrit'</i>	<i>peremeščat'sja na metle</i> 'move on a broom'

		'not able to see volumetric objects'		
37	<i>šaklyj</i>	<i>razočarovannyj</i> 'disappointed'	<i>šaklit'</i>	<i>peremeščat'sja verxom na verbljude</i> 'move sitting on a camel's back'
38	<i>ščulyj</i>	<i>legko obizajuščijsja</i> 'touchy, being easily offended'	<i>ščulit'</i>	<i>peredvigat'sja, ščurjas' ot solnečnogo sveta</i> 'move while squinting at the sunshine'
39	<i>gnoryj</i>	<i>upotrebljajuščij v pišču tol'ko sladkoe</i> 'eating only sweet food'	<i>gnorit'</i>	<i>peredvigat'sja na rukax vverx nogami, ulybajas' ot sčast'ja</i> 'move walking on palms upside down smiling happily'
40	<i>žrapyj</i>	<i>beruščij vzjatki</i> 'taking bribes'	<i>žrapit'</i>	<i>s trudom peremeščat' na verevke jaščik s boepripasami</i> 'move a box of ammunition on a rope with difficulty'
41	<i>znupyj</i>	<i>iduščij polnym xodom, intensivnyj</i> 'going at full speed, intensive'	<i>znupit'</i>	<i>peremeščat'sja v sanjax, zaprjažennyx severnymi olenjami</i> 'move riding a sleigh drawn by reindeers'
42	<i>čtusyj</i>	<i>sosredotočennyj na detaljax</i> 'concentrated on details'	<i>čtusit'</i>	<i>peredvigat'sja, sil'no šatajas'</i> 'move while staggering'
43	<i>žgavyj</i>	<i>imejuščij navjazčivuju ideju postojanno myt' ruki</i> 'having an obsessive idea to wash one's hands all the time'	<i>žgavit'</i>	<i>peremeščat'sja v meške</i> 'move (jumping) in a bag'
44	<i>spulyj</i>	<i>ne sposobnyj rabotat' pravoj rukoj</i> 'not able to work with the right hand'	<i>spulit'</i>	<i>peremeščat'sja na vodnyx lyžax</i> 'move on waterskis'
45	<i>skolyj</i>	<i>znamenityj</i> 'famous'	<i>skolit'</i>	<i>peremeščat'sja v lodke na veslax</i> 'move by paddling a boat'
46	<i>tkabyj</i>	<i>stesnitel'nyj</i> 'shy'	<i>tkabit'</i>	<i>peremeščat'sja na vertolete</i> 'move by helicopter'

Appendix 5

Minimal pairs of verbs in PERE- and PRE-

There are 14 pairs of standard verbs:

- | | | |
|--|---|--|
| (1) <i>pere-byt'</i> 'overcome a period of time' | – | <i>pre-byt'</i> 'last, remain constant' |
| (2) <i>pere-byvat'</i> 'visit many places' | – | <i>pre-byvat'</i> 'be, reside, be in a state of |
| (3) <i>pere-dat'(sja)</i> 'give to, send to' | – | <i>pre-dat'(sja)</i> 'commit to; betray' |
| (4) <i>pere-gorodit'</i> 'partition off, barrier' | – | <i>pre-gradit'</i> 'obstruct, block' |
| (5) <i>pere-lomit'(sja)</i> 'break in two, fracture' | – | <i>pre-lomit'(sja)</i> 'refract (ray), interpret' |
| (6) <i>pere-poručit'</i> 're-assign to smb' | – | <i>pre-poručit'</i> 'entrust, commit' |
| (7) <i>pere-rvat'(sja)</i> 'tear apart' | – | <i>pre-rvat'(sja)</i> 'break off, interrupt' |
| (8) <i>pere-seč'(sja)</i> 'cross (a road)' | – | <i>pre-seč'(sja)</i> 'cut short, stop an activity' |
| (9) <i>pere-stat'</i> 'stop' | – | <i>pre-stat'</i> 'stop doing smth' |
| (10) <i>pere-stavit'(sja)</i> 'move' | – | <i>pre-stavit'sja</i> 'die' |
| (11) <i>pere-stupit'</i> 'step over' | – | <i>pre-stupit'</i> 'violate (a law)' |
| (12) <i>pere-terpet'</i> 'overcome suffering' | – | <i>pre-terpet'</i> 'undergo, endure' |
| (13) <i>pere-vorotit'(sja)</i> 'turn over' | – | <i>pre-vratit'(sja)</i> 'convert into, transform' |
| (14) <i>pere-sytit'</i> 'make too staffed' (this verb is only attested in participial forms) | | |
| – <i>pre-sytit'</i> 'sate with' | | |

There are 8 additional pairs where one of the members is marginal and very infrequent in the corpus:

- (15) *pere-klonit'(sja)* (marginal) 'bend over; attract to' – *pre-klonit'(sja)* 'incline, bend, bow'
- (16) *pere-ložit'(sja)* 'lay to another place; put too much' – *pre-ložit'(sja)* (marginal) 'shift, transfer'
- (17) *pere-menit'(sja)* 'change' – *pre-menit'(sja)* (marginal) 'change'
- (18) *pere-pirat'sja* (marginal) 'argue with' – *pre-pirat'sja* 'argue with'
- (19) *pere-selit'sja* 'move, migrate' – *pre-selit'sja* (marginal) 'move to another place'
- (20) *pere-slušat'(sja)* 'listen to many (songs)' – *pre-slušat'(sja)* (marginal) 'disobey'
- (21) *pere-tvorit'* (marginal) 're-create' – *pre-tvorit'* 'turn into, change into'
- (22) *pere-vodit'(sja)* 'transfer, take across' – *pre-vodit'(sja)* (marginal) 'transfer to'

One false minimal pair which exists due to homonymous base verbs:

- pere-zret'* 'overripen' (< *zret'* 'ripen') – *pre-zret'* 'disdain' (< *zret'* 'see')

Appendix 6

Experiment on O- and U- in factitive verbs: Stimuli

Standard verbal stimuli prefixed in O-

- (1) *Nado ne prosto **ob**'jasnit', a čtoby ešče stalo ponjatno!*
'One should not just **explain**, one should explain so that the point would become clear.'
- (2) *Ljuboe straxovanie prizvano **oblegčit'** bremja vozmožnoj imuščestvennoj otvetstvennosti.*
'Any insurance serves to **lighten** the burden of possible property liability.'
- (3) *Čtoby **oslabit'** ili soveršenno uničtožit' bol', nužno tščatel'no sobljudat' rekomendacii vrača.*
'In order to **reduce** or completely stop the pain, one should thoroughly follow their doctor's guidelines.'
- (4) *Rešat' zadači nado bez pomošči kal'kuljatora, pri neobxodimosti sleduet otvety **okruglit'** do celyx čisel.¹*
'In working out an assignment one should not use a calculator; if needed, one should **express** the outcomes **in round numbers**.'
- (5) *Tak vy možete bystro **obogatit'** svoj professional'nyj opyt.*
'In such a way you can quickly **enrich** your professional experience.'
- (6) *Te obstojatel'stva dolžny byli sdelat' nas vragami i **ožestočit'** naši duši.*
'Those circumstances had to make us enemies and **harden** our hearts.'
- (7) *Stremitel'naja èkspansija evro možet **osložnit'** naši vyplaty po vnešnemu dolgu.*
'The fast expansion of the euro can **complicate** our payments of external debt.'
- (8) *Ona ulybnulas' Miše, kak staromu znakomomu, i velela **ogolit'** život.*
'She smiled at Miša like as if he was a long-time acquaintance and asked him to **bare** his stomach.'

¹ In the pilot version of the experiment we used another context for the verb *okruglit'* 'round': *V mašine ja protjanul den'gi. Katerine ostavalos' tol'ko **okruglit'** glaza.* 'In the car I offered her money. Caterina could only **round** her eyes (in surprise).' Although this context presents a more spatial and more concrete meaning 'make something round (in shape)' of the verb *okruglit'* compared to its secondary metaphoric submeaning 'approximate, express in round numbers', the former use was banned by subjects as atypical. Being exposed to the two expressions *okruglit' glaza* 'round one's eyes' and *okruglit' čislo* 'round a number', all subjects preferred the latter in terms of a greater naturalness and higher frequency. As a result, we changed the experimental context for the one presented in (4).

- (9) *Skazano èto bylo tak, slovno svojim pojavleniem on dolžen byl srazu, siju minutu vsëx osčastlivit'.*
 'It was said in such a way, as if his arrival was supposed to **make** everybody **happy** at once, at this very moment.'
- (10) *S pomošč'ju ètoj očiščajuščej maski možno legko uvlažnit' kožu i osvežit' cvet lica.*
 'By means of this cleansing mask one can easily moisturize the skin and **freshen** the complexion.'

Standard verbal stimuli prefixed in U-

- (11) *Vlasti Južnoj Korei pytajuťsja utočnit' vse obstojatel'stva proiššedšego.*
 'The authorities of South Korea are trying to **get more specific information** on the facts of the matter.'
- (12) *Primenenie ètoj mikrošxemy dast vozmožnost' umenšit' gabarity i ves bytovoj texniki.*
 'The implementation of this microchip will make it possible to **reduce** the dimensions and weight of household appliances.'
- (13) *Novaja sistema pozvoljaet uskorit' dostavku gruzov i povysit' bezopasnost' personala.*
 'The new system makes it possible to **speed up** transportation and to increase staff safety.'
- (14) *Èto upražnenie dolžno ulučšit' osanku i pomoč' pri naprjaženii v spine.*
 'This exercise is supposed to **improve** your body posture and to help with muscle tension in your back.'
- (15) *Ja vezde zamečal stremlenie kolleg podognat' sobytija pod štampy, uprostit', sdelat' ponjatnymi dlja čitatelja.*
 'I would notice everywhere that my colleagues tried to squeeze events into standard cliches, **simplify**, and make them comprehensible for the reader.'
- (16) *Mladšij brat rešil ukorotit' svoju familiju. Teper' on podpisывalsja — Danč.*
 'The younger brother decided to **shorten** his surname. From now on he spelled his name as Danč.'
- (17) *Reforma možet sil'no usložnit' strukturu mestnogo samoupravlenija.*
 'The reform can **make** the system of local self-government far **more complex**.'
- (18) *Ja pytajus' uznat', čem možno uteplit' pol v starom kamennom dome na cokol'nom ètaže.*
 'I am trying to find out what materials one can use in order to **provide heat insulation** for a basement floor in an old stone house.'
- (19) *Čtoby izbežat' problem, nužno kanavu zasypat' gruntom i xorošen'ko ego uplotnit'.*

'In order to avoid problems, one should cover the ditch with soil and **seal** it properly.'

- (20) *Spirt možet tol'ko **uxudšit'** rabotu central'noj nervnoj sistemy.*
'Pure alcohol can only **worsen** (impair) the functioning of the central nervous system.'

Marginal verbal stimuli prefixed in O-

- (21) *Prikinut'sja mertvym – èto ne prosto nepodvižno ležat'! Nado ešče **ostekljanit'** glaza i praktičeski ne dyšat'!*
'Playing dead is not just lying motionless! One should also **make** one's eyes **glassy** and almost stop breathing.'
- (22) *Čukovskij preziral avtorov, staravšixsja kak možno skoree **oser'ěznit'** rebenka.*
'Čukovskij disdained the authors who tried to **make** children **serious** as early as possible.'
- (23) *Davno pora kak-to **opriličit'** naše obščenie bolee mjagkimi vyraženijami.*
'It's high time we **made** our interaction **respectable** by using gentler expressions.'
- (24) *Gollivud uxitrilsja **opoxabit'** počti vse šedevry literatury.*
'Hollywood has managed to **profane** almost all masterpieces of fiction.'
- (25) *U nas klub vrode meždunarodnyj, predlagaju i èmblemu **omeždunarodit'**, možno flagami, libo nazvanijami stran.*
'Our club is kind of international; I suggest that we should **internationalize** the logo as well, may be with flags or names of different contries.'
- (26) *Nam nužno nad vorami i vzjatočnikami **osurovit'** zakon!*
'We should **make** the legislation on thieves and bribe-takers **severe**.'
- (27) *Dumaju, v svobodnom dostupe net nikakogo ximičeskogo rastvora, kotoryj mog by za dve nedeli prilično **oržavit'** metall.*
'I think there is no freely available chemical solution that could properly **rust** the metal in two weeks.'
- (28) *Pered teatral'nym xudožnikom stoit trudnaja zadača – ne **obytovit'** proisxodjaščee na scene, ostat'sja v ploskosti poëtičeskogo jazyka.*
'A stage designer has a challenging task – not to **vulgarize** the onstage action and to keep to the level of poetic language.'
- (29) *Ljudi, naxodjas' v sil'nyx duševnyx terzanijax, inogda nanosjat sebe rany, kusajut lokti. Èto instinktivnyj pozyv **ovnešnit'** stradanie.*
'When people experience strong emotional suffering, they sometimes injure themselves, kick themselves. This is an instinctive impulse to **exteriorize** the suffering.'

- (30) *Zdes' pojavilos' umenie kompozitora **omuzykalit'** russkuju reč', sozdat' svobodnyj, živoy i napevnyj rečitativ.*
'Here the composer demonstrated his talent to **muzicalize** the sound of the Russian speech, to create a free-floating, vivid and melodious recitative.'

Marginal verbal stimuli prefixed in U-

- (31) *Usovremenit' arxitekturu v gorodax Rossii možno bylo by putem snosa vetxix domov v centre mnogix gorodov.*
'It could be possible to **modernize** the architecture in Russian cities by demolishing shabby houses in many city centers.'
- (32) *My xoteli pesnju **user'ëznit'** i pridat' peniem nemnogo fol'klornosti.*
'We wanted to **make** the song **sound more serious** and add a flavour of folk music as well.'
- (33) *Devčonki, podskazite, kak možno ostanovit' ili **umedlit'** rost neželaemyx volos.*
'Girls, give me some tips on how one can stop or **slow down** the growth of unwanted hair.'
- (34) *S rekami tože sleduet razobrat'sja. Nužno **ukrasivit'** ix nazvanija, podnjat' prestiž rodneyx prostorov.*
'One should look into the issue with rivers as well. It is necessary to **make** the names of rivers **prettier**, to enhance the prestige of our home grounds.'
- (35) *Ljubaja krupnaja firma Moskvy možet ètot spisok prodlit', **ukonkretit'** i razvit'.*
'Any large firm in Moscow can extend, **concretize** (make more specific), and develop this list.'
- (36) *Prošče vsego **uvkusnit'** tvorog dobavleniem izjuma, banana i sladkix jagod.*
'The easiest way to **make** cottage-cheese **taste better** is to add raisins, banana, and sweet berries.'
- (37) *Telesnoe uveč'e možet **ucelomudrit'** dušu.*
'A bodily injury can **purify** the soul.'
- (38) *Za dva goda nam udalos' ubrat' korrupcionnye sostavljajuščie, navesti porjadok, **uprozračit'** procedury.*
'In two years we managed to get rid of corrupted institutions, put things right, **make** the procedures **more transparent**.'
- (39) *Evrokomissija namerena **udorožit'** avtomobili v srednem na 6000 dollarov.*
'The European Commission intends to **make** cars on the average 6,000 dollars **more expensive**.'
- (40) *Načal'stvu ničego ne ostavalos', kak podvintit' gajki: **ustrožit'** režim bylo lučšim sposobom pokončit' s putanicej.*

'The authorities could do nothing but "tighten the bolts": to **toughen** the regime was the best way to stop the confusion.'

Nonce factitive verbs with the prefix O-

Since nonce verbs do not have specific meanings, I have translated them using "default" factitive English equivalents *change* and *affect*.

- (41) *Esli budete kormit' rebenka ryboj každyj den', to možete ego tak **oblusit'**, čto on u vas potom na rybu smotret' ne budet.*
'If you feed a child with fish every day, you can **change** him **in some way** so that he would not be able to even look at fish.'
- (42) *Kostja čto est' sily spešil na počtu. To, čto ee zakryli na polčasa ran'še, ne moglo ego ne **ošaklit'**.*
'Kostja was hurrying to get to the post office. The fact that it closed a half an hour earlier than it was supposed to, could not but **affect** him **in some way**.'
- (43) *Novye komp'juternye igry mogut **otovit'** ljubogo: na èto rabotaet i grafika, i cvetovaja gamma, i sjužet.*
'New computer games can **affect** anyone: for this purpose they employ a certain graphic design, color range, and plot.'
- (44) *S pjatnicy zarjadil dožd', no Slavika èto ne moglo **osurit'**. U nas v zapase byla ečše odna zateja.*
'Since Friday a heavy rain set in, but it could not **affect** Slavik². We had one more plan in our supplies.'
- (45) *Po moim ožidanim, novoe plat'e i pričeska dolžny Dašu tak **oduktit'**, čto ee budet prosto ne uznat'.*
'According to my expectations, the new dress and hairstyle must **change** Daša in such a way that she will become simply beyond all recognition.'
- (46) *Danila vyletel iz kabaka neudačno, povredil pravuju ruku tak sil'no, čto ego moglo sil'no **okočlit'**.*
'Danila dashed out of the pub ill-fated; he damaged his right arm so badly that it could strongly **affect** him.'
- (47) *Kogda my našli Musju vo dvore, ona byla sovsem dikoj, no na moločke udalos' ee **očavit'**, i ona vskore stala ručnoj.*
'When we found Musja³ in the yard, she was absolute wild, but we managed to **change / affect** her with milk, and she soon became tame.'
- (48) *Skrupuleznaja professija ne mogla ne **ogabit'** Muravjeva, sdela ego dotošnym i pedantičnym.*

² A diminutive form of the Slavic male name Slava, Svjatoslav.

³ A common name for a female-cat.

'A punctilious profession could not but **affect** Muravjev in such a way that he became meticulous and pedantic.'

- (49) *Solenuju morskiju vodu **obnomit'** očen' prosto, nado tol'ko znat' tehnologiju.*
'It is very easy to **change / alter** salty sea water, one needs only to know the technology.'
- (50) *Ostaetsja **obmmlit'** devušek poslednim raundom šopinga, i zavtra oni budut milymi i laskovymi.*
'What remains to do is to **affect** the young girls with the last round of shopping, and tomorrow they will be nice and sweet.'

Nonce factitive verbs with the prefix U-

- (51) *Blagodarja reguljarnym zanjatijam muzykoj Alle udalos' značitel'no **ušadrit'** svoj golos i slux.*
'Thanks to the regular music lessons Alla managed to significantly **change** her singing and hearing.'
- (52) *Maša i tak nikogo ne slušaetsja, a babuškinj gostincy mogut ee tol'ko ešče bol'se **uloprit'**.*
'Maša disobeys everyone as she is, but grandmother's gifts can do nothing else than **affect** her even more (towards even greater disobedience).'
- (53) *Dlinu trapecii sledovalo **utulit'** napolovinu.*
'The length of trapezia had to be **changed** by half.'
- (54) *Čtoby mexanizm nakonec zarabotal, priššlos' **uguzvit'** napor vody v dva raza.*
'In order to get the machine to function finally, one had to **alter (increase or reduce)** water pressure twofold.'
- (55) *V rezul'tate sajť udalos' prilično **usaglit'**, privnesti nedostajuščego losku v ego dizajn.*
'As a result, we managed to decently **change** the website, and to eliminate the lack of gloss in its design.'
- (56) *Ja izo vsej sil staralsja **unokrit'** etiketku, pri ètom soxraniv strogost' oformlenija.*
'I did my best to **change** the label and to preserve the sparse style at the same time.'
- (57) *Nužno **učopit'** uslovija truda, podnjat' zarplatu, obščij uroven' žizni, a potom uže srezat' privilegii.*
'One should first **change** working conditions, raise salaries and the overall living standard, and only afterwards cut off the benefits.'
- (58) *Esli vy voz'mete s soboj sobaku, pridetsja **ukampit'** stoimost' èkskursii na 40 rublej.*
'If you take the dog along, we will have to **change** the price of the tour by forty rubles.'

- (59) *Vskore biznes pošel v goru, pojavilas' vozmožnost' ego **udostit'** i ešče bol'se razvit'.*
'Soon the business became successful, and it became possible to **alter** it and to develop it even more.'
- (60) *Skazali, čto goditsja, nado tol'ko čut'-čut' **umarvit'** tekst, čtoby soderžanie bylo ponjatnee.*
'They said that it is good enough, one should only **change** the text slightly so that the message would be easier to understand.'

Appendix 7

Experiment on O- and U- in factitive verbs: Results

Acceptability scores assigned to factitive verbs in O- and U-

Factitive verb	WordType	Freq in RNC texts 1950-2012	Adults-sum of all scores	Children sum of all scores	Adults % of max possible score	Children % of max possible score	Adults' responses					Children's responses				
							Score "5"	Score "4"	Score "3"	Score "2"	Score "1"	Score "5"	Score "4"	Score "3"	Score "2"	Score "1"
ob''jasnit'	S	18,149	255	350	100	100	51	0	0	0	0	70	0	0	0	0
utočnit'	S	2,860	254	349	99.6	99.7	50	1	0	0	0	69	1	0	0	0
umen'šit'	S	2,010	255	350	100	100	51	0	0	0	0	70	0	0	0	0
uskorit'	S	2,008	255	350	100	100	51	0	0	0	0	70	0	0	0	0
ulučšit'	S	1,899	255	350	100	100	51	0	0	0	0	70	0	0	0	0
oblegčit'	S	1,802	254	349	99.6	99.7	50	1	0	0	0	69	1	0	0	0
oslabit'	S	1,401	255	346	100	98.8	51	0	0	0	0	67	2	1	0	0
uprostit'	S	1,350	255	350	100	100	51	0	0	0	0	70	0	0	0	0
okruglit'	S	939	255	350	100	100	51	0	0	0	0	70	0	0	0	0
obogatit'	S	800	254	334	99.6	95.4	50	1	0	0	0	60	6	3	0	1
ukorotit'	S	787	253	348	99.2	99.4	49	2	0	0	0	65	4	1	0	0
ožestočit'	S	686	227	322	89	92	31	15	2	1	2	48	16	6	0	0
osložnit'	S	410	232	345	90.9	98.5	43	2	1	1	4	67	1	2	0	0
ogolit'	S	387	248	322	97.2	92	44	7	0	0	0	50	14	4	2	0
osčastlivit'	S	343	228	348	89.4	99.4	42	9	0	0	0	65	4	1	0	0
usložnit'	S	311	255	349	100	99.7	51	0	0	0	0	69	1	0	0	0
osvežit'	S	280	254	343	99.6	98	50	1	0	0	0	64	5	1	0	0
uteplit'	S	205	254	349	99.6	99.7	50	1	0	0	0	69	1	0	0	0
uplotnit'	S	201	254	344	99.6	98.2	50	1	0	0	0	64	6	0	0	0
uxudšit'	S	199	252	348	98.8	99.4	49	1	1	0	0	68	2	0	0	0
udorožit'	M	8	149	181	58.4	51.7	9	13	8	7	14	11	7	18	10	24
ovnešnit'	M	4	107	115	41.9	32.8	0	0	7	14	30	0	1	10	22	37
omuzykalit'	M	4	104	127	40.7	36.2	3	2	14	7	25	0	3	17	14	36
uprozračit'	M	4	100	161	39.2	46	2	5	8	10	26	1	12	20	11	26
ucelomudrit'	M	3	119	206	46.6	58.8	4	10	9	4	24	11	19	11	13	16
ustrožit'	M	3	112	154	43.9	44	2	8	9	11	21	1	9	18	17	25
obytovit'	M	3	109	148	42.7	42.2	1	6	11	14	19	2	3	19	23	23
oržavit'	M	2	123	147	48.2	42	2	7	15	13	14	0	5	27	8	30
osurovit'	M	2	110	192	43.1	54.8	3	6	12	5	25	5	10	29	14	12
opoxabit'	M	1	198	281	77.6	80.2	20	16	9	1	5	25	25	16	4	0

usovremenit'	M	1	134	255	52.5	64.2	8	7	11	8	17	14	19	16	10	11
oser'jeznit'	M	1	128	181	50.1	51.7	5	8	14	5	19	1	14	24	17	14
ostekljanit'	M	1	126	198	49.4	56.5	4	8	15	5	19	8	13	20	17	12
opriličit'	M	1	117	205	45.8	58.5	1	10	13	6	21	8	15	21	16	10
umedlit'	M	1	116	156	45.4	44.5	6	6	10	3	26	8	5	15	9	33
ukonkretit'	M	1	116	210	45.4	60	5	7	9	6	24	11	18	16	10	15
user'jeznit'	M	1	109	153	42.7	43.7	3	7	11	3	27	3	9	15	14	29
omeždunarodit'	M	1	98	149	38.4	42.5	0	1	15	14	21	3	4	21	13	29
uvkusnit'	M	1	85	111	33.3	31.4	0	0	11	12	28	0	2	12	11	45
ukrasivit'	M	1	69	100	27	28.5	0	0	4	10	37	0	1	6	15	48
oblusit'	N	0	101	122	39.6	34.8	0	0	4	14	33	0	1	11	27	31
onomit'	N	0	79	131	30.9	37.4	0	1	6	13	31	0	2	15	25	28
ušadrit'	N	0	77	117	30.1	33.4	0	0	9	8	34	0	3	9	20	38
utulit'	N	0	75	121	29.4	34.5	0	0	7	10	34	0	2	7	31	30
ogabit'	N	0	75	143	29.4	40.8	1	0	7	6	37	8	1	13	12	36
očavit'	N	0	73	121	28.6	34.5	0	0	5	12	34	1	1	8	28	32
osurit'	N	0	72	112	28.2	32	0	1	4	10	36	0	1	11	17	41
obmomlit'	N	0	71	123	27.8	35.1	0	1	2	13	35	1	4	11	15	39
okočlit'	N	0	69	111	27	31.4	0	0	4	10	37	0	2	5	25	38
uguzvit'	N	0	69	100	27	28.5	0	0	4	10	37	0	2	5	14	49
otovit'	N	0	68	112	26.6	32	0	0	5	7	39	0	0	13	16	41
ošaklit'	N	0	68	111	26.6	31.4	0	0	4	9	38	1	0	12	13	44
udamlit'	N	0	66	114	25.8	32.5	0	0	4	7	40	0	1	8	25	36
umarvit'	N	0	65	103	25.4	29.4	0	0	3	8	40	0	0	7	19	44
unokrit'	N	0	65	117	25.4	33.4	0	0	3	8	40	0	1	8	18	43
učopit'	N	0	65	108	25.4	30.8	0	0	3	8	40	2	1	9	14	44
usaglit'	N	0	63	111	24.7	31.4	0	0	1	10	40	1	0	12	13	44
uloprit'	N	0	62	98	24.3	28	0	0	2	7	42	0	0	4	20	46
ukampit'	N	0	60	101	24.5	28.8	0	1	1	4	45	0	0	6	19	45
oduktit'	N	0	59	91	23.1	26	0	0	1	6	44	0	0	2	17	51

Word type: S – standard; M – marginal; N – Nonce

Purple color shading marks those marginal verbs that receive very diverse acceptability judgements.

References

- Adamovič, Ja. M. et al. 1980. *Sučasnaja belaruskaja literaturnaja mova: Marfalogija*. Vyšejšaja škola.
- Alekseeva, A. P. 1978. *Iz istorii pristavočnogo glagolnogo slovoobrazovanja (na primere obrazovanij s OB i O)*. Avtoreferat dissertaciji. Leningrad.
- Anderson, Stephen R. 1992. *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Anderson, Stephen R. To appear. The Morpheme: Its Nature and Use. In Baerman, Matthew (ed.) *The Oxford Handbook of Inflection*. Available at http://cowgill.ling.yale.edu/sra/morpheme_final.pdf Accessed 16.08.2014.
- Anderson, Stephen R. & Wolfgang U. Dressler. 2003. Morphology. In Frawley, William J. (ed.) *International Encyclopedia of Linguistics*. Vol. 3. Oxford: Oxford University Press. 81-85.
- Andrews, Edna. 1984. A semantic Analysis of the Russian Prepositions/Preverbs O(-) and OB(-). In *The Slavic and East European Journal*. 28.4. 477-492.
- Apresjan, Jurij D. 1995. *Izbrannye trudy. Vol. 1. Leksičeskaja semantika. Sinonimičeskie sredstva jazyka*. 2 ed. Moskva: Jazyki russkoj kul'tury.
- Apresjan, Jurij D. 2005. O Moskovskoj semantičeskoj škole. In *Voprosy jazykoznanija*. 1. 3-31.
- Apresjan, Jurij D. 2009. *Issledovanija po semantike i leksikografii. Vol. 1: Paradigmatika*. Moskva: Jazyki slavjanskix kul'tur.
- Aronoff, Mark. 1983. Potential Words, Actual Words, Productivity and Frequency. In *Proceedings of the 13th international congress of linguists*, August 29 - September 4, 1982. Tokyo. 163-171.
- Aronoff, Mark. 1994. *Morphology by itself: Stems and inflectional classes*. (Linguistic Inquiry Monograph 22). Cambridge, MA: MIT Press.
- Aver'janova A. P. 1974. Semantičeskie različija glagol'nyx obrazovanij s prstavkami o- i ob- (obo-). In *Voprosy semantiki*. 1. Leningrad: Izdatel'stvo Leningradskogo universiteta. 28-36.
- Aver'janova, Galina N. 2008. *Russkie glagol'nye prstavki*. Moskva: Russkij jazyk.
- Axmanova, Olga S. 1957. *Očerki po obsčej i ruskoj leksikologii*. Moskva. Izdatel'stvo ministerstva prosvješčenija RSFSR.
- Baayen, R. Harald. 2008. *Analysing linguistic data. A practical introduction to Statistics using R*. Cambridge University Press.
- Baayen et al. 2013 – Baayen, R. Harald, Laura A. Janda, Tore Nessel, Anna Endresen, Anastasia Makarova. 2013. Making choices in Slavic: Pros and cons of statistical methods for rival forms. In *Russian Linguistics 37* (Special issue "Space and Time in Russian Temporal Expressions"). 253-291.
- Baranovskaja, V. G. 1974. Glagoly s prstavkoj vъz- v drevnerusskom jazyke 11-14 vv. In *Voprosy slovoobrazovanja i leksikologii drevnerusskogo jazyka*. Moskva: Nauka. 122-137.
- Bard et al. 1996 – Bard, Ellen Gurman, Dan Robertson and Antonella Sorace. 1996. Magnitude Estimation of Linguistic Acceptability. In *Language*. 72.1. 32-68.
- Barykina et al. 1989 – Barykina, Alefina N., Valerija V. Dobrovol'skaja, Stella N. Merzon. 1989. *Izučeniye glagol'nyx prstavok*. Moskva: Russkij jazyk.
- Basova et al. 2003 – Basova, Galina, Aleksandr Kačura, Aleksandr Kixno, Valentina Musienko, Nina Ozerova, Galina Olijnyk, Neonila Romanova, Elena Snitko, Ol'ga

- Tvoronovič. 2003. *Sopostavitel'naja grammatika russkogo i ukrainskogo jazykov*. Kiev: Naukova dumka.
- Baudouin de Courtenay, Jan A. 1881a. Podrobnaja programma lekcij v 1876-1877 uč. godu (=Detailed Program of Lectures (1877-78)). In *Izvestija Kazanskogo universiteta*, 1879, issue 2: 353-382; issue 6: 445-492; 1880 issue 6: 370-480; 1881: 481- 608. Kazan', Warsaw. Reprinted in Baudouin de Courtenay 1963: 108-117. Translated into English in Stankiewicz 1972.
- Baudouin de Courtenay, Jan A. 1881b. Nekotorye otdely sravnitel'noj grammatiki slavjanskix jazykov. In *Russkij filologičeskij vestnik*. Vol. 5. 265-344. Reprinted in Baudouin de Courtenay 1963: 118-126. Translated into English in Stankiewicz 1972.
- Baudouin de Courtenay, Jan A. 1895. *Opyt teorii fonetičeskix alternacij. Glava iz psixofonetiki*. (=An Attempt at a Theory of Phonetic Alternations. A chapter from psychophonetics). (Translation of Versuch einer Theorie der phonetischen Alternationen: ein Kapitel aus Psychophonetik. Strasburg. 1895). 265-347. Translated into English in Stankiewicz 1972.
- Baudouin de Courtenay, Jan A. [Boduén de Kurtené, Ivan A.] 1963. *Izbrannye trudy po obščemu jazykoznaniju*. Vol. 1, 2. Moskva: Izdatel'stvo akademii nauk SSSR.
- Bauer, Laurie. 1988. *Introducing linguistic morphology*. Bristol: Edinburgh University Press.
- Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.
- Bauer, Laurie. 2004. *A Glossary of Morphology*. Washington, D.C.: Georgetown University Press.
- Bauer, Laurie. 2012. *Grammaticality, acceptability and the notion of possible word*. A paper presented at the Conference "Data-Rich Approaches to English Morphology" in Victoria University Wellington, New Zealand.
- Baydimirova (Endresen), Anna & Svetlana Sokolova. 2010. Pristavki VZ- i VOZ- v ruskom jazyke: "Sed'maja voda na kisele" ili "dva v odnom"? Paper presented at November Seminar on Russian at the University of Tromsø, 26.11.2010.
- Baydimirova (Endresen), Anna. 2010. *Russian aspectual prefixes O, OB, and OBO: A Case Study of Allomorphy*. Master's thesis. University of Tromsø. Available at <http://www.ub.uit.no/munin/handle/10037/2767> Accessed 15.08.2014.
- Beard, Robert. 1982. The plural as a lexical derivation. In *Glossa*. 16. 133-48.
- Beard, Robert. 1987. Morpheme order in a Lexeme/Morpheme based morphology. In *Lingua*. 72. 1-44.
- Beard, Robert. 1995. *Lexeme-morpheme base morphology: A general theory of inflection and word-formation*. SUNY Series in Linguistics. State University of NY Press.
- Będkowska-Kopczyk, Agnieszka. 2012. Semantic and grammatical features of *o-/ob-* in verbs of emotion in Slovene. In *Jezikoslovje*. 13.1. 19-38.
- Będkowska-Kopczyk, Agnieszka. 2012. The prefix *vz-* and its allomorphs in verbs of emotional experience in Slovene. Grammatical evidence for upward motion metaphors of EMOTIONS. In *Filologie, areál a praxe: inovativnost v současných oborech*. Sv.II. Praha: Verbum. 82-105.
- Będkowska-Kopczyk, Agnieszka & Wojciech Lewandowski. 2012. *O-/ob- in Slovene and o(be)- in Polish. A comparative cognitive analysis*. Paper presented at Slavic Cognitive Linguistics Conference in Zagreb.
- Beekes, Robert. 2010. *Etymological Dictionary of Greek*. Vol. 1. Leiden, Boston.

- Belozercev, G. I. 1966. *Sootnošenie knižno-slavjanskix i narodno-razgovornyx èlementov v drevnerusskom literaturnom jazyke XI-XII vv. (Na materiale obrazovanij s prstavkami IZ- i VY- prostranstvennogo značenija)*. Avtoreferat dissertacii. Moskva.
- Bermel, Neil & Luděk Knittl. 2012a. Corpus frequency and acceptability judgements: A study of morphosyntactic variants in Czech. In *Corpus Linguistics and Linguistic Theory*. 8.2. 241-275.
- Bermel, Neil & Luděk Knittl. 2012b. Morphosyntactic variation and syntactic constructions in Czech nominal declension: corpus frequency and native-speaker judgements. In *Russian Linguistics*. 36.1. 91-119.
- Berneker, Erich D. 1924. *Slavisches etymologisches wörterbuch*. Heidelberg: C. Winter.
- Blaikie, Norman. 2003. *Analyzing qualitative data*. London: Sage Publications.
- Bloch, Bernard. 1947. English verb inflection. In *Language*. 23. 399-418.
- Bloomfield, Leonard. 1923. Review of de Saussure (1922). In *Modern Language Journal*. 8. 317-319. Reprinted in Hockett, Charles (ed.) 1970. *A Leonard Bloomfield Anthology*. Bloomington: Indiana University Press. 106-108.
- Bloomfield, Leonard. 1933 [1961]. *Language*. Reprint. New York: Henry Holt.
- Boguslavskij, Andrej. 1963 [2001]. Glagol'naja prefiksacija v sovremennom ruskom jazyke. In Gindin S. I. (ed.) *Moskovskij lingvističeskij žurnal (Glagol'nye prefiksy i prefiksial'nye glagoly)*. 5.1. 2001. Moskva: Izdatel'skij centr RGGU. 7-35. Original edition: Boguslawski, A. 1963. Prefiksacja szasownikowa we współczesnym języku rosyjskim. Wrocław, Warszawa, Kraków.
- Bondarko, Aleksandr V. et al. 1996. *Teorija funkcional'noj grammatiki. Kačestvennost', količestvennost'*. St. Petersburg: Nauka.
- Bonet et al. 2007 – Bonet, Eulàlia, Maria-Rosa Lloret and Joan Mascaró. 2007. Allomorph selection and lexical preferences: Two case studies. In *Lingua*. 117. 903-927.
- Booij, Geert. 1995. *The Phonology of Dutch*. Oxford: Clarendon Press.
- Booij, Geert. 1997. Allomorphy and the autonomy of morphology. In *Folia Linguistica*. 31. 25-56.
- Booij, Geert. 2005. *The grammar of words: an introduction to linguistic morphology*. Oxford: Oxford University Press.
- Booij, Geert. 2010. *Construction morphology*. Oxford: Oxford University Press.
- Botvinik, Maria. 2009. *Central Meaning of the Verbal Prefix VY- in Russian*. Paper presented at Slavic Cognitive Linguistics Conference, Prague, October 15-17, 2009.
- Britain, David. 2007. Grammatical variation in England. In Britain, David (ed.) *Language in the British Isles*. Cambridge University Press. 75-104.
- Britain, David & Seu Fox. 2009. The regularisation of the Hiatus Resolution System in British English. A Contact-Induced 'Vernacular Universal'?" In Filppula, Markku, Juhani Klemola and Heli Paulasto (eds.) *Vernacular Universals and Language Contacts. Evidence from Varieties of English and Beyond*. London, New York: Routledge. 177-205.
- Brugman, Claudia & George Lakoff. 1988 [2006]. Cognitive topology and lexical networks. In Small, Steven L., Garrison W. Cottrell, and Michael K. Tanenhaus (eds.) *Lexical Ambiguity Resolution: Perspectives from Psycholinguistics, Neuropsychology, and Artificial Intelligence*. 477-508. San Mateo, CA: Morgan Kaufmann. Cited from the reprint in Geeraerts, Dirk (ed.) 2006. *Cognitive Linguistics: Basic Readings*. Berlin, New York: Mouton de Gruyter. 110-139.

- Burzio, Luigi. 2006. Lexicon versus grammar in English morpho-phonology: Modularity revisited. In *Korean Journal of English Language and Linguistics*. 6. 437-464.
- Bussmann, Hadumod. 1996. *Routledge Dictionary of Language and Linguistics*. London & New York: Routledge.
- Bybee, Joan L. 2010. Diachronic Linguistics. In Geeraerts, D. & H. Cuyckens (eds.) *The Oxford Handbook of Cognitive Linguistics*. 945-988.
- Bye, Patrik. 2007. Allomorphy – selection, not optimization. In Krämer, M. & S. Blaho & P. Bye. (eds.) *Freedom of analysis?* Berlin: Mouton de Gruyter. 63-91.
- Bye, Patrik. 2014. The Nature of Allomorphy and Exceptionality: Evidence from Burushaski plurals. In Mascaró, J. & M.R. Lloret & E. Bonet (eds.) *Understanding allomorphy: Perspectives from Optimality Theory*. London: Equinox. 1-64.
- Cantos Gómez, Pascual. 2013. *Statistical methods in Language and linguistic research*. UK.
- Černyx, Petr Ja. 1999. *Istoriko-étimologičeskij slovar' sovremennogo russkogo jazyka*. Moskva: Russkij jazyk.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1980. *Rules and Representations*. Blackwell, Oxford.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Mahwah, New Jersey/London: Lawrence Earlbaum Associates.
- Cohen et al. 2000 – Cohen, Louis, Laurence Manion, and Keith Morrison. 2000. *Research Methods in Education*. 5th edition. London: Routledge Falmer.
- Cohen et al. 2003 – Cohen, J., Cohen, P., West, S.G., Aiken, L.S. 2003. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. 3rd ed. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Collins et al. 2009 – Collins, Chris & Stephanie N. Guitard & Jim Wood. 2009. Imposters: An Online Survey of Grammaticality Judgments. In *NYU Working Papers in Linguistics, Volume 2: Papers in Syntax, Spring*.
- Comrie, Bernard and Greville G. Corbett (eds.) 1993. *The Slavonic Languages*. Routledge.
- Cowart, Wayne. 1997. *Experimental syntax: Applying Objective Methods to Sentence Judgements*. Thousand Oaks, CA: Sage Publications.
- Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
- Croft, William & D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.
- Crystal, David. 1991. *A Dictionary of Linguistics and Phonetics*. 3rd edition. Blackwell Publishers.
- Čukovskij, Kornej I. 1933. *Ot dvux do pjati*. Leningrad: Izdatel'stvo pisatelej v Leningrade.
- Dąbrowska, Ewa. 2004. *Language, Mind and Brain: Some Psychological and Neurological Constraints on Theories of Grammar*. Edinburgh University Press.
- Dąbrowska, Ewa. 2010. Naive vs. expert intuitions: An empirical study of acceptability judgments. In *The Linguistic Review*. 27. 1-23.
- Dąbrowska, Ewa. 2012. Different speakers, different grammars: Individual differences in native language attainment. In *Linguistic Approaches to Bilingualism*. 2. 219-253.
- Dąbrowska, Ewa & James Street. 2006. Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. In *Language Sciences* 28. 604-615.

- Dadavaeva, P.M. 1978. *Glagoly s prstavkami IZ- i VY- vydelitel'nogo značenija v ruskom literaturnom jazyke vtoroj poloviny XVIII veka (na materiale žurnalov)*. Avtoreferat disertaciji. Leningrad.
- de Bray, R. G. A. 1969. *Guide to Slavonic Languages*. London, NY.
- Dem'janov, Vladimir G. 2001. *Inozajčnaja leksika v istorii ruskogo jazyka 11-17 vekov. Problemy morfoložičeskoj adaptaciji*. Moskva: Nauka.
- Di Sciullo, Anna M. & Edwin Williams. 1987. On the Definition of Word. In *Linguistic Inquiry*. Monograph 14. Cambridge, MA: MIT Press.
- Dickey, Stephen M. 2005. S-/Z- and the grammaticalization of aspect in Slavic. In *Slovene Linguistic Studies*. 5. 3-55.
- Dickey, Stephen M. & Laura A. Janda. 2009. *Xoxotnul, sxitri!*: The relationship between semelfactives formed with -nu- and s- in Russian. In *Russian Linguistics*. 33.3. 229-248.
- Divjak, Dagmar. 2010. *Structuring the lexicon: a clustered model of near-synonymy*. Berlin: Walter de Gruyter.
- Divjak, Dagmar & Stefan Th. Gries. 2006. Ways of trying in Russian: clustering behavioral profiles. In *Corpus Linguistics and Linguistic Theory*. 2.1. 23-60.
- Dixon, R. M. W. & Alexandra Y. Aikhenvald. 2004. *Adjective Classes. A Cross-Linguistic Typology*. Oxford University Press.
- Dobrušina, Ekaterina R. 2009. Metaforičeskaja prstavka ili Periferijnoje vozdejtviye (k voprosu o semantike prstavki o- / ob(o)-). In *Fonetika i grammatika: Nastojaščee, Prošedšee, Buduščee: Voprosy ruskogo jazykoznanija*. 13. Moskva: Izdatel'stvo Moskovskogo Universiteta.
- Dobrušina et al. 2001 – Dobrušina, Ekaterina R., E. A. Mellina, and Denis Paillard. 2001. *Russkije prstavki: mnogoznačnost' i semantičeskoje edinstvo: Sbornik*. Moskva: Russkije slovari.
- Dressler, Wolfgang U. and Lavinia Merlini Barbaresi. 1994. *Morphopragmatics: Diminutives and Intensifiers in Italian, German, and Other Languages*. Berlin, New York: Mouton de Gruyter.
- Eckhoff, Hanne M. & Laura A. Janda. 2014. Grammatical Profiles and Aspect in Old Church Slavonic. In *Transactions of the Philological Society*. 112.2. 231–258.
- Efremova, Tat'jana F. 2000. *Novyj tolkovo-slovoobrazovatelnyj slovar' ruskogo jazyka*. Moskva: Russkij jazyk.
- Efremova, Tat'jana F. 2006. *Sovremennyj tolkovyj slovar' ruskogo jazyka*. Moskva. Vol. 2.
- Efremova, Tat'jana F. 2010. *Sovremennyj slovar' ruskogo jazyka: Tri v odnom: Orfografičeskij. Slovoobrazovatel'nyj. Morfemnyj*. Moskva: Astrel'.
- Endresen, Anna. 2011. Russkije prstavki O-, OB- i OBO-: Raznye allomorphy ili morfemy? Eksperimental'noe issledovanie. In *Proceedings of the XL International Philological Conference at St. Petersburg State University*. Psycholinguistics. St. Petersburg. 46-55. Available at <http://genlingnw.ru/Staff/mart2011.pdf>
- Endresen, Anna. 2013. Samostojatel'nye morfemy ili pozicionnye varianty? Morfoložičeskij status ruskix prstavok o- i ob- v svete novyx dannyx: korpus i èksperiment. (= Distinct morphemes or positional variants? Morphological status of the Russian prefixes o- and ob- in the light of new evidence: corpus and experiment). In *Voprosy jazykoznanija*. 6. 33-69.
- Endresen, Anna & Vladimir Plungian. 2011. Russian morphophonemics in a nutshell: The verb *vstat'* 'stand up'. In *Poljarnyj Vestnik*. 14. 22-44.
- Endresen et al. 2012 – Endresen, Anna, Laura A. Janda, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nessel, and Svetlana Sokolova.

2012. Russian 'purely aspectual' prefixes: Not so "empty" after all? In *Scando-Slavica*. 58.2. 231-291.
- Evans, Nicholas & Stephen Levinson. 2009. The Myth of Language Universals: Linguistic diversity and its importance for cognitive science. In *Behavioral and Brain Sciences*. Cambridge University Press. 32. 429-492.
- Evans, Vyvyan. 2007. *A glossary of Cognitive Linguistics*. Edinburgh University Press.
- Evans et al. 2007 – Evans, Vyvyan, Benjamin K. Bergen, and Jörg Zinken. 2007. The Cognitive Linguistics Enterprise: An overview. In Evans, V., B.K. Bergen, J. Zinken (eds.) *The Cognitive Linguistics Reader*. Equinox Publishers. 2-36.
- Evgen'eva, Anastasija P. (ed.) 1999. *Malyj Akademičeskij Slovar'*. Moskva: Russkij jazyk.
- Fábregas, Antonio & Sergio Scalise. 2012. *Morphology. From Data to Theories*. Edinburgh: Edinburgh University Press.
- Faraway, Julian J. 2006. *Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models*. Chapman & Hall/CRC Texts in Statistical Science.
- Féry, Caroline. 1994. Umlaut and Inflection in German. To appear in Kager, Rene & Harry van der Hulst & Wim Zonneveld (eds.) *The Prosody Morphology Interface*. 44. Accessed at Rutgers Optimality Archive ROA 34.
- Filin, Fedot P. 1981. *Istoki i sud'by russkogo literaturnogo jazyka*. Moskva: Nauka.
- Fleischer, W. 1971. *Wortbildung der deutschen Gegenwartssprache*. Leipzig.
- Flier, Michael S. 1985. Syntagmatic Constraints on the Russian Prefix *pere-*. In *Issues in Russian Morphosyntax*. Flier Michael S. & Richard D. Brecht (eds.). Columbus, Ohio: Slavica Publishers. 138-154.
- Fodor, Jerry A. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. MIT Press.
- Forsyth, James. 1970. *A Grammar of Aspect. Usage and Meaning of the Russian Verb*. Cambridge: Cambridge University Press.
- Frank, Stefan L., Rens Bod and Morten H. Christiansen. 2012. How hierarchical is language use? In *Proceedings of the Royal Society B*. Published online 12.09.2012. Available at <http://rspb.royalsocietypublishing.org/content/279/1747/4522.abstract> Accessed 15.08.2014.
- Gabrielatos et al. 2010 – Gabrielatos, C., Torgersen, E., Hoffmann, S. & Fox, S. 2010. A corpus-based sociolinguistic study of indefinite article forms in London English. In *Journal of English Linguistics*. 38.4. 297-334.
- Gallant, James. 1979. *Russian Verbal Prefixes and Semantic Features: An Analysis of Prefix vz-*. Slavistische Beiträge. Band 135. Munich: Otto Sagner.
- Gardner, P. L. 1975. Scales and statistics. In *Review of Educational Research*. 45. 43-57.
- Geeraerts, Dirk. 1997. *Diachronic prototype semantics: A contribution to historical lexicology*. Oxford: Clarendon Press.
- Geeraerts et al. 2010 – Geeraerts, Dirk & Gitte Kristiansen & Yves Peirsman (eds.). 2010. *Advances in Cognitive Sociolinguistics*. Berlin: Mouton de Gruyter.
- Gerd, Aleksandr S. 2004. Morf, allomorf, morfonema. In Verbickaja, Ljudmila A. (ed.) *Teoretičeskie problemy jazykoznanija: Sbornik statej k 140-letiju kafedry obščego jazykoznanija Filologičeskogo fakul'teta Sankt-Peterburgskogo gosudarstvennogo universiteta*. Sankt-Petersburg: Izdatel'stvo Sankt-Peterburgskogo Universiteta. 209-213.
- Golanova, E. I. 1998. O "mnimyx složnyx slovač" (razvitie klassa analitičeskix prilagatel'nyx v sovremennom russkom jazyke). In *Liki jazyka: K 45-letiju naučnoj dejatel'nosti E.A. Zemskoj*. Moskva. 31-39.

- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Golovin, B.N. 1965. Slovoobrazovatel'nye tipy glagolov s prstavkoj vz-. In *Očerki po russkomu jazyku*. Kirov: Izdatel'stvo Kirovskogo pedagogičeskogo instituta. 130-138.
- Gor, Kira & Tatiana Chernigovskaya. 2003. Mental lexicon structure in L1 and L2 acquisition: Russian Evidence. In *Glossos*. 4. 1-31.
- Gorelik, Elena V. 2001. Opisanije glagol'noj prstavki u-. In *Moskovskij lingvističeskij žurnal*. 5.1. Moskva: Izdatel'skij centr RGGU. 37-68.
- Gouskova, Maria. 2012. Unexceptional segments. In *Natural Language and Linguistic Theory*. 30.1. 79-133.
- Gouskova, Maria & Michael Becker. 2013. Nonce words show that Russian yer alternations are governed by the grammar. In *Natural Language and Linguistic Theory*. 31.3. 735-765.
- Greve, Rita & Bärbel Kroesche. 1958. *Russische rückläufiges Wörterbuch* (= Obratnyj slovar' russkogo jazyka). Wiesbaden: O. Harrassowitz.
- Gribanova, Vera. 2009. Phonological evidence for a distinction between Russian prepositions and prefixes. In *Proceedings of FDSL 7*.
- Gries, Stefan Th. 2009. *Statistics for Linguistics with R: A Practical Introduction*. Berlin, New York: Mouton de Gruyter.
- Haas, William. 1954. On defining linguistic units. In *Transactions of the Philological Society*. 54-84.
- Hale, Ken & Samuel Jay Keyser. 2002. *Prolegomenon to a Theory of argument structure*. Cambridge, Massachusetts: The MIT Press.
- Hamp, Erik P. 1988. The spread of *-uno* factitives in Greek. In *Transactions of the Philological Society*. 86.1. 88-91.
- Harrel, Frank E. 2001. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. Springer series in Statistics. NY.
- Harris, Roy. 2001. *Saussure and his Interpreters*. New York University Press.
- Harris, Zellig S. 1942. Morpheme Alternants in Linguistic Analysis. In *Language*. 18.3. 169-180.
- Haspelmath, Martin. 1997. *From Space to Time: Temporal Adverbials in the World's Languages*. München & Newcastle: LINCOM Europa.
- Haspelmath, Martin. 2002. *Understanding Morphology*. London: Oxford University Press.
- Heaman, Isabel M. 1984. Baudouin de Courtenay – a pioneer of structural linguistics. In *Working papers of the Linguistic Circle of the University of Victoria*. 4.1. 23-40.
- Heath, Jeffrey. 2008. *A Grammar of Jamsay*. Walter de Gruyter.
- Herman, Louis Jay. 1975. *A Dictionary of Slavic Word Families*. Columbia University Press.
- Hilpert, Martin. 2008. New evidence against the modularity of grammar: Constructions, collocations, and speech perception. In *Cognitive Linguistics*. 19.3. 483-503.
- Hockett, Charles F. 1942. A system of descriptive phonology. In *Language*. 18. 3-21.
- Hockett, Charles F. 1947. Problems of morphemic analysis. In *Language*. 23. 321-343.
- Hockett, Charles F. 1968. *The state of the art*. The Hague: Mouton.
- Hougaard, Christian. 1973. Vyražajet li o-/ob- soveršaemost'? In *Scando-Slavica*. 19. 119-125.
- Ignat'eva, M.V. 1970. Glagoly s prostranstvennymi značenijami prstavki S- (SO-) v sovremennom russkom jazyke. In *Issledovanija po leksikologii i*

- slovoobrazovaniju*. Doklady IX zonal'noj konferencii kafedr ruskogo jazyka vuzov Sibiri. Krasnojarsk. 117-134.
- Isačenko, Aleksandr V. 2003/1965. *Grammatičeskij stroj ruskogo jazyka v sopostavlenii s slovackim: Morfologija*. (The grammatical structure of Russian in comparison with Slovak: Morphology). 2-e izdanie. Part II. Moskva: Jazyki slavjanskoj kul'tury.
- Itkin, Il'ja B. 2007. *Russkaja morfonologija*. Moskva: Gnozis.
- Itkonen, Esa. 2005. *Analogy as Structure and Process: Approaches in Linguistics, Cognitive Psychology and Philosophy of Science*. John Benjamin's Publishing Company.
- Jaccard, James & Choi K. Wan. 1996. LISREL approaches to interaction effects in multiple regression. Thousand Oaks, CA: Sage Publications.
- Jakobson, Roman. 1971. The Kazan' School of Polish Linguistics and its Place in the International Development of Phonology. In Jakobson, Roman. *Selected Writings. Vol. 2. Word and Language*. The Hague: Mouton. 394-428.
- Jamieson, Susan. 2004. Likert scales: How to (ab)use them. In *Medical Education*. 38. 1212-1218.
- Janda, Laura A. 1986. *A Semantic Analysis of the Russian Verbal Prefixes ZA-, PERE-, DO- and OT-* (= Slavistische Beiträge, Band 192). Munich: Otto Sagner.
- Janda, Laura A. 1993. Cognitive linguistics as a continuation of the Jakobsonian tradition: the semantics of Russian and Czech reflexives. In Maguire, Robert A. & Alan Timberlake (eds.) *American Contributions to the Eleventh International Congress of Slavists in Bratislava*. Columbus: Slavica. 310-319.
- Janda, Laura A. 2004. A metaphor in search of a source domain: The categories of Slavic aspect. In *Cognitive Linguistics*. 15.4. 471-527.
- Janda, Laura A. 2007. Aspectual Clusters of Russian Verbs. In *Studies in Language*. 31.3. 607-648.
- Janda, Laura A. 2008. Motion Verbs and the Development of Aspect in Russian. In *Scando-Slavica* 54. 179-197.
- Janda, Laura A. 2010. Cognitive linguistics in the year 2010. In *International journal of cognitive linguistics*. 1.1. 1-30.
- Janda, Laura A. 2011. Metonymy in word-formation. In *Cognitive Linguistics*. 22.2. 359-392.
- Janda, Laura A. 2012. Russkie pristavki kak sistema glagol'nyx klassifikatorov (=Russian prefixes as a system of verb classifiers). In *Voprosy jazykoznanija*. 6. 3-47.
- Janda, Laura A. (ed.) 2013. *Cognitive Linguistics: The Quantitative Turn. The Essential Reader*. Berlin: De Gruyter Mouton.
- Janda et al. 2013 – Janda, Laura A., Anna Endresen, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nettet, and Svetlana Sokolova. 2013. *Why Russian aspectual prefixes aren't empty: prefixes as verb classifiers*. Bloomington: Slavica publishers.
- Janda, Laura A. & Olga Lyashevskaya. 2011. Grammatical profiles and the interaction of the lexicon with aspect, tense and mood in Russian. In *Cognitive Linguistics*. 22.4. 719-763.
- Janda, Laura A. & Olga Lyashevskaya. 2013. Semantic Profiles of Five Russian Prefixes: *po-*, *s-*, *za-*, *na-*, *pro-*. In *Journal of Slavic Linguistics*. 21.2. 211-258.
- Janda, Laura A. & Tore Nettet. 2010a. Taking apart Russian RAZ-. In *Slavic and East European Journal*. 54.3. 476-501.
- Janda, Laura A. & Tore Nettet. 2010b. *Neat theories, messy realities: How to apply absolute definitions to gradient phenomena*. Description of the NRF project available at <http://ansatte.uit.no/laura.janda/> Accessed 15.08.2014.

- Janda, Laura A. & Valery Solovyev. 2009. What constructional profiles reveal about synonymy: A case study of Russian words for SADNESS and HAPPINESS. In *Cognitive Linguistics*. 20.2. 367-393.
- Janssen, Niels & Horacio A. Barber. 2012. Phrase Frequency Effects in Language Production. In *PloS ONE* 7.3: e33202. doi:10.1371/journal.pone.0033202
- Jarceva, Viktorija N. (ed.) 1990. *Lingvističeskij énciklopedičeskij slovar'*. Moskva: Izdatel'stvo Sovetskaja énciklopedija.
- Johnson, Mark. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason*. Chicago: University of Chicago Press.
- Kagan, Olga. 2011. The Scale Hypothesis and the Prefixes *pere-* and *nedo-*. In *Scando-Slavica*. 57. 160-176.
- Kagan, Olga. 2013. Scalarity in the domain of verbal prefixes. In *Natural Language & Linguistic Theory*. 31. 483-516.
- Katyšev, Pavel A. 2005. *Polimotivacija i smyslovaja mnogomernost' slovoobrazovatel'noj formy*. Tomsk: Izdatel'stvo Tomskogo universiteta.
- Kaverina, V.V. 1999. Sovremennaja orfografija pristavok na <z> kak rezul'tat zakonomernogo razvitija jazykovoju komunikacii. In Krasnyx V.V. & A.I.Izotov (eds.) *Jazyk, soznanie, komunikacija: Sbornik statej*. Moskva: Dialog-MGU. 7. 100-114.
- Keller, Frank. 2000. *Gradience in Grammar: Experimental and Computational Aspects of Degrees of Grammaticality*. Doctoral Dissertation. University of Edinburgh.
- Keller, Frank & Ash Asudeh. 2001. Constraints on linguistic coreference: Structural vs. pragmatic factors. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kilgarriff, Adam. 2005. Language is never, ever, ever, random. In *Corpus Linguistics and Linguistic Theory*. 1.2. 263-275.
- Kim, Jae-On. 1975. Multivariate analysis of ordinal variables. In *American Journal of Sociology*. 81. 261-298.
- King, B.M. & E.W. Minium. 2008. *Statistical reasoning in the behavioral sciences*. Hoboken, NJ.
- Klobukova, L. P. 1985. Struktura veršiny slovoobrazovatel'noj paradigmy ruskogo prilagatel'nogo. In *Aktual'nye problemy ruskogo slovoobrazovanija. Sbornik naučnyx statej*. Taškent. 84-89.
- Knapp, Thomas R. 1990. Treating Ordinal Scales as Interval Scales: An Attempt to Resolve the Controversy. In *Nursing Research*. 39.2. 121-122.
- König, Ekkehard. 2001. Intensifiers and reflexive pronouns. In Haspelmath, Martin (ed.) *Language Typology and Language Universals. An International Handbook*. Vol. 1. Berlin, New York: Walter de Gruyter. 747-760.
- Košev, Aleksej D. 2004a. O glagolax s semantikoju agresivnosti, obrazovannyx pristavkami *o-* i *ob-* (*obrugat'*, *obvešat'*, *osvistat'...*). In Šaronov, I. A. (ed.) *Agressija v jazyke i reči*. Sbornik naučnyx statej. Moskva: RGGU. 126-144.
- Košev, Aleksej D. 2004b. O konceptual'nyx značenijax pristavki *O-/OB-*. In *Voprosy jazykoznanija*. 4. 68-101.
- Krasuxin, Konstantin G. 2008. *Otkuda est' pošlo slovo. Zametki po etimologii i semantike*. Moskva: Nauka.
- Krongauz, Maksim A. 1998. *Pristavki i glagoly v ruskom jazyke: semantičeskaja grammatika*. Moscow: Jazyki ruskoj kul'tury.
- Kustova, Galina I. 2004. *Tipy proizvodnyx značenij i mexanizmy jazykovogo rasširenija*. Moskva: Jazyki slavjanskoj kul'tury.

- Kuznecova A. I. & T. F. Efremova. 1986. *Slovar' morfem russkogo jazyka*. Moskva: Russkij jazyk.
- Kuznecov, Sergej A. (ed.) 2000. *Bol'shoj tolkovyj slovar' russkogo jazyka*. Sankt-Petersburg: Norint.
- Kuznetsova, Julia & Anastasia Makarova. 2012. Distribution of two semelfactives in Russian: -nu- and -anu-. In *Oslo Studies in Language*. 4.1. 155-176.
- Kuznetsova, Julia & Svetlana Sokolova. Forthcoming. Aspectual triplets in Russian: Semantic Predictability and Regularity.
- Kuznetsova, Julia. 2013. *Linguistic Profiles: Correlations between Form and Meaning*. Doctoral Dissertation. University of Tromsø.
- Labov, William. 1984. Intensity. In Schiffrin, Deborah (ed.) *Meaning, Form, and Use in Context: Linguistic Applications*. Georgetown University Press. 43-71.
- Labovitz, Sanford. 1967. Some observations on measurement and statistics. In *Social Forces*. 46. 151-160.
- Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: The University of Chicago Press.
- Lakoff, George & Mark Johnson. 1980. *Metaphors We Live By*. Chicago: University of Chicago Press.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical prerequisites*. Stanford University Press.
- Langacker, Ronald W. 1999. *Grammar and Conceptualization*. Berlin, New York: Mouton de Gruyter.
- Langacker, Ronald W. 2008. *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago, London: The University of Chicago Press.
- Levin, Beth and Malka Rappaport Hovav. 1998. Morphology and Lexical Semantics. In Spencer, Andrew and Arnold M. Zwicky (eds.) *Handbook of Morphology*. Blackwell Publishers. 249-272.
- Levinson, L. 2010. Arguments for pseudo-resultative predicates. In *Natural and Linguistic Theory*. 28: 135-182.
- Lomtev, T. P. 1961. *Sravnitel'no-istoričeskaja grammatika vostočnoslavjanskix jazykov (Morfologija)*. Moskva: Vysšaja škola.
- Lopatin, Vladimir V. 1976 [2007]. Množestvennost' motivaciii i ee otaženie v otglagol'nom imennom slovoobrazovanii. In *Russkij jazyk v škole*. 2. Reprinted in Lopatin, Vladimir V. 2007. *Mnogogrannoe russkoe slovo: Izbrannye statji po russkomu jazyku*. Moskva: Azbukovnik. 177-188.
- Lopatin, Vladimir V. 1977 [2007]. *Russkaja slovoobrazovatel'naja morfemika: Problemy i principy opisanija*. Moskva: Nauka. Reprinted in 2007. Moskva: Azbukovnik.
- Lopatin, Vladimir V. 2000. *Is istorii reformirovanija russkogo pravopisanija*. Available at the electronic portal www.gramota.ru at http://www.gramota.ru/biblio/magazines/gramota/28_15 Accessed 21.12.2013.
- Lopatin, Vladimir V. (ed.) 2007. *Russkij orfografičeskij slovar'*. Moskva: Russian Academy of Sciences.
- Lorenz, Gunter. 2002. Really worthwhile or not really significant? A corpus-based approach to delexicalization and grammaticalization of intensifiers in Modern English. In Wisher, Ilse and Gabriele Diewald (eds.) *New Reflections on Grammaticalization*. Vol. 49 of Typological Studies in Language (TSL). Amsterdam, Philadelphia: John Benjamins Publishing Company. 143-161.

- Luschützky, Hans Ch. 2000. Morphem, Morph und Allomorph. In Booij, Geert et al. (eds.) *An International Handbook on Inflection and Word-Formation*. Berlin, New York: Walter de Gruyter. 451-463.
- Lyashevskaya, Olga N. and Sergej A. Šaroff. 2009. *Častotnyj slovar' sovremennogo russkogo jazyka (na materiale Nacional'nogo korpusa russkogo jazyka)*. Moskva: Azbukovnik.
- Lyons, John. 1968/1971. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.
- Mahootian, Shahrzad. 1997. *Persian*. London: Routledge.
- Mahota, William J. 1993. Romance "re" and Russian "voz". In *Canadian Slavonic Papers/Revue canadienne des slavistes*. 35:3/4. 291-304.
- Makarova, Anastasia. 2009. *Psycholinguistic Evidence for Allomorphy in Russian Semelfactives*. Master's thesis. University of Tromsø.
- Makarova, Anastasia. 2014. *Rethinking diminutives: A case study of Russian Verbs*. Doctoral Dissertation. University of Tromsø.
- Makarova, Anastasia & Laura A. Janda. 2009. Do It Once: A Case Study of the Russian -Hy-Semelfactives. In *Scando-Slavica*. 55. 78-99.
- Marinova, Elena V. 2010. Vopros ob analitičeskix prilagatel'nyx v otečestvennoj i zarubežnoj lingvistike. In *Lingvistika. Vestnik Nižegorodskogo universiteta im. N.I. Lobačevskogo*. 4. 2. 628-630. Available at http://www.unn.ru/pages/issues/vestnik/99999999_West_2010_4%282%29/60.pdf Accessed on 27.09.2013.
- Markarjan, N. E. 1974. Iz istorii upotreblenija glagol'nyx obrazovanij s pristavkami PERE- i PRE- (na materialax letopisnogo jazyka 15 v.). In *Voprosy slovoobrazovanija i leksikologii drevnerusskogo jazyka*. Moskva: Nauka. 96-109.
- Markov, Vitalij M. 1970. Problema formirovanija samostojatel'nyx morfem na osnove protivopostavljenija fonetičeskix variantov. In Markov, Vitalij M. (ed.) *Voprosy grammatičeskogo stroja russkogo jazyka*. Kazan': Izdatel'stvo Kazanskogo universiteta. 3-21.
- Mascaró, Joan. 2007. External Allomorphy and Lexical Representation. In *Linguistic Inquiry*. 38.4. 715-735.
- Matthews, Peter H. 1972. *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation*. Cambridge: Cambridge University Press.
- Matthews, Peter H. 1974. *Morphology. An introduction to the theory of word-structure*. 1st edition. Cambridge: Cambridge University Press.
- Matthews, Peter H. 1991. *Morphology*. 2nd edition. Cambridge: Cambridge University Press.
- Matthews, Peter H. 1993. *Grammatical theory in the United States from Bloomfield to Chomsky*. Cambridge: CUP.
- Matthews, Peter H. 2003. Morpheme. In Frawley, William J. (ed.) *International Encyclopedia of Linguistics*. 2nd edition. Vol. 3. Oxford: Oxford University Press.
- Matushansky, Ora. 2002. On formal identity of Russian prefixes and prepositions. In: Csirmaz, A. and Z. Li, A. Nevins, O. Vaysman and M. Wagner (eds.) *Phonological Answers (and their corresponding questions)*. MITWPL 42, MIT Working Papers in Linguistics. Cambridge MA. 217-253.
- Matveeva, N. P. 1983. Razvitie abstraktnyx značenij pristavki vz- (vzo-, vs-) v ruskom jazyke v sopostavlennii s ukrajskim. In *Russkoe jazykoznanie*. 7. Kiev. 10-17.
- McGranahan, William J. 1975. Russian initial consonant clusters: A comparative approach. In *Russian Language Journal*. 29.102. 13-20.

- Mel'čuk, Igor' A. 1998. *Kurs obščej morfologii* (A Course of General Morphology). *Tom 2. Morfoložičeskiye značenija* (Morphological meanings). Moskva, Vena: Jazyki ruskoj kul'tury.
- Mel'čuk, Igor' A. 2001. *Kurs obščej morfologii* (A Course of General Morphology). *Tom 4. Morfoložičeskiye znaki* (Morphological signs). Moskva, Vena: Jazyki ruskoj kul'tury.
- Miklosich, Franz R. von. 1970. *Etymologisches Wörterbuch der slavischen Sprachen*. Reprint. Amsterdam: Philo Press.
- Mizoe, Tatsuhide. 2011. *Verbal Prefixation in Slavic. A cognitive semantic analysis of s-prefixed Russian loan verbs from English and other languages*. Doctoral Dissertation. Université Laval.
- Nefed'ev, Mixail V. 1994. Semantičeskaja èvoljucija glagol'nyx pristinok *na-* i *ob-* v istorii ruskojazyka XI – XVIII vv. In *Voprosy jazykoznanija*. 4. 73-83.
- Nefed'ev, Mixail V. 1995. Zametki o razvitanii slovoobrazovatel'nyx tipov (na primere glagolov s pristinokoj *ob-*). In *Voprosy jazykoznanija*. 6. 90-96.
- Neff, David. 2002. Interview: Eugene Nida on Meaning-full Translations. From October 7, 2002. In *Christianity today*. Available at <http://www.christianitytoday.com/ct/2002/october7/2.46.html>; accessed 16.08.2014.
- Nemčenko, Vasilij N. 1979. *Sposoby ruskoj slovoobrazovanija: Diachroničeskij aspekt*. Gor'kij: Izdatel'stvo Gor'kovskogo universiteta.
- Neset, Tore. 2000. Iconicity and Prototypes: A New Perspective on Russian Verbs of Motion. In *Scando-Slavica*. 46. 105-119.
- Neset, Tore. 2008. *Abstract Phonology in a Concrete Model. Cognitive Linguistics and the Morphology-Phonology Interface*. Berlin and New York: Mouton de Gruyter.
- Neset, Tore. 2009. Is the choice of prefix arbitrary? Aspectual prefixation and Russian verbs of perception. In *Slavic and East European Journal*. 54.4. 666-689.
- Neset, Tore. 2010. The art of being negative: metonymical morphological constructions in contrast. In *Oslo Studies in Language*. 2.2. 261-279.
- Neset, Tore. 2013. The History of the Russian Semelfactive: The Development of a Radial Category. In *Journal of Slavic Linguistics*. 21.1. 123-169.
- Neset et al. 2009 – Neset, Tore, Anna Baydimirova (Endresen), Laura A. Janda. 2009. Getting out: A preliminary analysis of *vy-* and *iz-*. In *Poljarnyj Vestnik*. 12. 72 - 85.
- Neset et al. 2011 – Neset, Tore, Laura A. Janda, Anna Endresen. 2011. Two ways to get out: Radial Category Profiling and the Russian Prefixes *vy-* and *iz-*. In *Zeitschrift für Slawistik*. 4. 377-402.
- Nichols, Johanna. 1993. Transitive and Causative in the Slavic Lexicon: Evidence from Russian. In Comrie, B. and M. Polinsky (eds.) *Causatives and Transitivity*. John Benjamin Publishing Company. 69-87.
- Nida, Eugene. 1948. The identification of Morphemes. In *Language*. 24. 414-441.
- Nida, Eugene. 1964. *Toward a Science of Translating: With Special Reference to Principles and Procedures Involved in Bible Translating*. Leiden: Brill.
- Nida, Eugene. 1975. *Componential Analysis of Meaning: An Introduction to Semantic Structures*. The Hague: Mouton.
- Nida, Eugene. 1982. *Translating Meaning*. San Dimas, California: English Language Institute.
- Nikitevič, Aleksej V. 2006. Glagoly-otad'ektivy v sostave nominativnyx rjadov. In Krys'ko, Vadim B. (ed.) *Slavistika: Sinxronija i diachronija. Sbornik naučnyx statej*. Moskva: Azbukovnik. 222-233.

- Nørgård-Sørensen, Jens. 2011. *Russian Nominal Semantics and Morphology*. Bloomington: Slavica Publishers.
- Offord, Derek. 1996. *Using Russian: A guide to contemporary usage*. Cambridge: Cambridge University Press.
- Ožegov, Sergej I. & Natalija Ju. Švedova. 2001. *Tolkovyj slovar' russkogo jazyka*. Moskva: Russkij jazyk.
- Pacjukova, Ol'ga A. 2009. Protjažennyj prefiks *obez-* (*obes-*): formirovanije i funkcionirovanije. In *Lingvistika. Vestnik Nižegorodskogo universiteta im. N.I. Lobačevskogo*. 6.2. 308-313.
- Panov, Mixail V. 1971. Ob analitičeskix prilagatel'nyx. In *Fonetika. Fonologija. Grammatika*. K 70-letiju A.A. Reformatskogo. Moskva: Nauka. 240-253.
- Paster, Mary E. 2006. *Phonological Conditions on Affixation*. Doctoral dissertation. University of California, Berkeley.
- Paster, Mary. 2009. Explaining phonological conditions on affixation: Evidence from suppletive allomorphy and affix ordering. In *Word Structure*. 2. 18-37.
- Pater, Joe. 2004. *Exceptions and Optimality Theory: Typology and Learnability*. Paper presented at Conference on Redefining Elicitation: Novel Data in Phonological Theory, New York University. Available at https://files.nyu.edu/ld43/public/Abstracts/pater_abs.pdf Accessed 15.08.2014.
- Pater, Joe. 2007. The Locus of Exceptionality: Morpheme-Specific Phonology as Constraint Indexation. In Bateman, Leah, Michael O'Keefe, Ehren Reilly and Adam Werle (eds.). *University of Massachusetts Occasional Papers 32: Papers in Optimality Theory III*. 259-296.
- Pater, Joe. 2010. Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Parker, Steve (ed.) *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox. 123-154.
- Pell, Godfrey. 2005. Use and misuse of Likert scales. In *Medical Education*. 39.9. 970.
- Percov, Nikolaj V. 2001. Russkij vid: slovoizmenenie ili slovoobrazovanie? In Percov, Nikolaj V. *Invarianty v russkom slovoizmenenii*. Moskva: Jazyki russkoj kul'tury.
- Pesetsky, David. 1979. *Russian morphology and lexical theory*. Ms. MIT. <http://web.mit.edu/linguistics/people/faculty/pesetsky/russmorph.pdf> (accessed 12.08.2014).
- Pilipenko, Natal'ja Anatoljevna. 2001. *Razvitie abstraktnyx značenij u glagolov s prstavkoj včz- (voz-, vz-) v istorii russkogo jazyka*. Dissertacija kandidata nauk. Rostovskij gosudarstvennyj universitet. Rostov-na-Donu.
- Pozolotina, Tatiana M. 2009. Emptiness is not so empty: The meaning of the Russian prefix *vy-* in the natural perfective verbs. In *Poljarnyj Vestnik*. 12. 85-105.
- Preobražensky, Alexandr G. 1964. *Etymological Dictionary of the Russian Language*. NY, London: Columbia University Press.
- Pugh, Stefan M. and Ian Press. 1999. *Ukrainian: A Comprehensive Grammar*. London: Routledge.
- Rappaport Hovav, Malka & Beth Levin. 2002. Change of State Verbs: Implications for Theories of Argument Projection. In *Proceedings of the 28th Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Field Linguistics*. 269-280.
- Raxilina, Ekaterina V. 1998. Kognitivnaja semantika: istorija, idei, rezul'taty. In *Semiotika i informatika*. 36. 274-323.

- Roberts C. B. 1976. Lexical differentiation of the Russian prefixal allomorphs O-, OB-, OBO-. In *Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationsforschung*. 29. 64-76.
- Roberts, C. B. 1981. The origins and development of O(B)- prefixed verbs in Russian with the general meaning 'deceive'. In *Russian Linguistics*. 5.3. 217-233.
- Rodina, Yulia. 2007. Semantics and morphology: the acquisition of grammatical gender in Russian. Doctoral dissertation. University of Tromsø.
- Rogoza, Natal'ja R. 1992. *Otymennye glagoly s suffiksom -i- v russkom jazyke 11-17 vekov*. Kandidatskaja dissertacija. Moskva.
- Romanova, Eugenia. 2004. *Superlexical vs. lexical prefixes*. In *Nordlyd*. 32.2. 255-278.
- Rosch, Eleanor. 1973. Natural categories. In *Cognitive Psychology*. 4. 328-350.
- Rosch, Eleanor. 1975. Cognitive representations of semantic categories. In *Journal of Experimental Psychology: General*. 104.3. 192-233.
- Rosch, Eleanor. 1977a. Human categorization. In Warren, N. (ed.) *Studies in Cross-linguistic Psychology*. London: Academic Press. 1-49.
- Rosch, Eleanor. 1977b. Classification of Real-World Objects: Origins and Representations in Cognition. In Johnson-Laird, P.N. & Wason, P.C. *Thinking: Readings in Cognitive Science*. Cambridge: Cambridge University Press. 212-222.
- Rosch, Eleanor. 1978 [1999]. Principles of categorization. In Lloyd, B. & E. Rosch (eds.) *Cognition and Categorization*. Hillsdale, NJ: Erlbaum. 27-48. Reprinted in Margolis, E. & S. Laurence (eds.) 1999. *Concepts: Core Readings*. Cambridge, MA: MIT Press. 189-206.
- Rosch, Eleanor. 1983. Prototype Classification and Logical Classification: The Two Systems. In Scholnick, E. K. (ed) *New Trends in Conceptual Representation: Challenges to Piaget's Theory?* Hillsdale: Lawrence Erlbaum Associates. 73-86.
- Rosch, Eleanor & Carolyn B. Mervis. 1975. Family Resemblances: Studies in the Internal Structure of Categories. In *Cognitive Psychology*. 7.4. 573-605.
- Rozental', D. & M. Telenkova. 1999. *Slovar' trudnostej russkogo jazyka*. Moskva.
- Sanders, Ted & Eve Sweetser. 2009. Causality in Language and Cognition – what causal connectives and causal verbs reveal about the way we think. In Sanders, Ted & Eve Sweetser (eds.) *Causal Categories in Discourse and Cognition*. Berlin, NY: Mouton de Gruyter. 1-19.
- Sangster, Rodney B. 1982. *Roman Jakobson and beyond: Language as a system of signs*. Berlin: Mouton.
- Šapošnikov, Aleksandr K. (ed.) 2010. *Ètimologičeskij slovar' sovremennogo russkogo jazyka*. Moskva: Flinta, Nauka.
- Šarić, Ljiljana. 2012. A cognitive account of the prefix *uz* in Croatian. In *Jezikoslovlje*. 13.1. 191-218.
- Šarić et al. 2013 – Šarić, Ljiljana, Anna Endersen, Agnieszka Będkowska-Kopczyk, Svetlana Nedelcheva, Liljana Mitkovska, Eleni Bužarovska, Wojciech Lewandowski. 2013. *Cognitive Linguistics and Typology: Comparing Polysemous Prefixes O- and OB- across Slavic*. Paper presented at 9th Mediterranean Morphology Meeting in Dubrovnik, Croatia.
- Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. (= Course in General Linguistics). 1st ed. Edited by Charles Bally and Albert Sechehaye, with the collaboration of Albert Riedlinger. Lausanne & Paris: Payot. (2nd ed. in 1922).
- Schmid, Hans-Jörg & Helmut Küchenhoff. 2013. Collostructional analysis and other ways of measuring lexicogrammatical attraction: Theoretical premises, practical problems and cognitive underpinnings. In *Cognitive Linguistics*. 24.3. 531-577.

- Schupbach, Richard D. 1975. *Russian Desubstantival Derivation. A paradigmatic view*. In *Studia Linguistica et philologica*. 2. California: Anma Libri.
- Schütze, Carson T. 1996. *The Empirical Base of Linguistics: Grammaticality judgements and linguistic methodology*. Chicago, London: The University of Chicago Press.
- Semenov, Anton V. 2003. *Slovoovrazovatel'nyj slovar' russkogo jazyka*. Moskva: Junves.
- Shull, Sarah. 2003. *The Experience of Space. The Privileged Role of Spatial Prefixation in Czech and Russian*. München: Verlag Otto Sagner.
- Słoński, Stanisław. 1937. *Funkcje prefiksów werbalnych w języku starosłowiańskim (starobułgarskim)*. Warszawa: Nakład Towarzystwa Naukowego Warszawskiego.
- Slovar' russkogo jazyka 11-17 vv.* 1981. Vol. 8.; 1976. Vol. 3. Moskva: Nauka.
- Smiskova-Gustafsson, Hana. 2013. *Chunks in L2 development: a usage-based perspective*. Doctoral dissertation. University of Groningen.
- Smith, Carlota S. 1997. *The Parameters of Aspect*. Dordrecht: Kluwer Academic Publications.
- Song, Jae Jung. 1996. *Causatives and causation: a universal-typological perspective*. London, NY: Longman.
- Sorace, Antonella & Frank Keller. 2005. Gradience in Linguistic Data. In *Lingua*. 115. 1497-1524.
- Sorokina, T. N. 2012. K istokam pravopisanija bezudarnyx glasnyx v korne. In *Russkaja reč'*. 5. 87-93.
- Soudakoff, Dorothy. 1975. The prefixes *pere-* and *pre-*: A definition and comparison. In *The Slavic and East European Journal*. 19.2. 230-238.
- Spagis, Anna A. 1968. *Parnyje i neparnyje glagoly v ruskom jazyke*. Moskva: Prosveščenie.
- Spencer, Andrew. 1991. *Morphological theory: An introduction to word structure in generative grammar*. Oxford: Blackwell.
- Sprouse, Jon. 2007. *A program for experimental syntax*. Doctoral dissertation. University of Maryland.
- Stankiewicz, Edward. 1972. *A Baudouin de Courtenay Anthology: The Beginnings of Structural Linguistics*. Bloomington, London: Indiana University Press.
- Steels, Luc & M. Hild (eds.) 2012. *Language Grounding in Robots*. New York: Springer Verlag.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collocations: Investigating the interaction of words and constructions. In *International Journal of Corpus Linguistics*. 8.2. 209-243.
- Steriopolo, Olga. 2007. Yer vowels in Russian prepositions. In *Proceedings of Formal Approaches to Slavic Linguistics 15: The Toronto Meeting*. Ann Arbor, MI: Michigan Slavic Publications. 365-385.
- Sternefeld, Wolfgang. 2006. *Syntax. Eine morphologisch motivierte generative Beschreibung des Deutschen*. Band 2. Tübingen: Stauffenburg Verlag.
- Stokhof, Martin & Michiel van Lambalgen. 2011. Abstractions and idealizations: The construction of modern linguistics. In *Theoretical Linguistics*. 37.1-2. 1-26.
- Strobl et al. 2009 – Strobl, Carolin, Gerhard Tutz, and James Malley. 2009. An introduction to Recursive Partitioning: Rationale, Application, and Characteristics of Classification and Regression Trees, Bagging, and Random Forests. In *Psychological Methods*. 14.4. 323-348.
- Sussex, Roland & Paul, Cubberley. 2006. *The Slavic Languages*. Cambridge, New York: Cambridge University Press.

- Švedova, Natalija Ju. et al. 1970. *Grammatika sovremennogo russkogo jazyka*. Moskva: Nauka.
- Švedova, Natalija Ju. et al. 1980. *Russkaja grammatika*. Vol. 1. Moskva: Nauka.
- Svenonius, Peter. 2004. Slavic prefixes inside and outside VP. In *Nordlyd*. 32.2. 205-253.
- Tabačenko, Ljudmila V. 2011. *Pristavočnye pozicionnye glagoly v istorii russkogo jazyka: poliaspektnyj analiz*. Dissertacija doktora filologičeskix nauk. Južnyj federal'nyj universitet. Rostov-na-Donu.
- Tadžibaev, B. B. 1985. Semantika otad''ektivnyx slovoobrazovatel'nyx cepoček (na materiale imen prilagatel'nyx so značenim cveta). In *Aktual'nye problemy russkogo slovoobrazovanija. Sbornik naučnyx statej*. Taškent. 138-145.
- Tagliamonte, Sali A. & R. Harald Baayen. 2012. Models, forests and trees of York English: *Was/were* variation as a case study for statistical practice. In *Language Variation and Change*. 24.2. 135-178.
- Talmy, Leonard. 1985. Force dynamics in language and thought. In *Chicago Linguistic Society* 21.2. 293-337.
- Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. Vol. 1.: *Concept structuring systems*. Cambridge, MA: MIT Press.
- Tatevosov, Sergei G. 2008. Intermediate prefixes in Russian. In Antonenko, Andrei et al. (eds.) *Annual workshop on formal approaches to Slavic linguistics: The Stony Brook meeting*. Ann Arbor: Michigan Slavic Publications. 423-445.
- Tatevosov, Sergei G. 2013. Destruktivnye zametki o raspredelitel'nom sposobe dejstvija. In *Russkij jazyk v naučnom osveščanii*. 2. 24-44.
- Taylor, John R. 1995. *Linguistic Categorization: Prototypes in Linguistic Theory*. 2nd edition. Oxford: Clarendon Press.
- Thomas, Lawrence L. 1969. *The History of the Russian Literary Language from the Seventeenth Century to the Nineteenth*. A condensed adaptation into English with an introduction by Lawrence L. Thomas. Madison: The University of Wisconsin Press.
- Timberlake, Alan. 2004. *A Reference grammar of Russian*. Cambridge: Cambridge University Press.
- Tixonov, Aleksandr N. 1985. *Slovoobrazovatel'nyj slovar' russkogo jazyka*. Moskva: Russkij jazyk.
- Tixonov, Aleksandr N. 1996. *Morfemno-orfografičeskij slovar'*. *Russkaja Morfemika*. Moskva: Škola-press.
- Tolstoy, Nikita I. 2002. Cerkvornoslavjanskij i russkij: ix sootnošenije i simbioz. In *Voprosy jazykoznanija*. 1. 81-90.
- Tomasello, Michael. 1992. *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press.
- Townsend, Charles E. 1968. *Russian word-formation*. Slavica Publishers. Reprinted in 2008.
- Townsend, Charles E. & Laura A. Janda. 1996. *Common and Comparative Slavic: Phonology and Inflection*. Columbus, OH: Slavica.
- Trask, Robert L. 1996. *A Dictionary of Phonetics and Phonology*. London, New York: Routledge.
- Tucker, Elizabeth. 1981. Greek factitive verbs in -οω, -αινω and -υνω. In *Transactions of the Philological Society*. 79.1. 15-34.
- Uluxanov, Igor' S. 1974. Slavjanizmy i narodno-razgovornaja leksika v pamjatnikax drevnerusskogo jazyka 15-17 vv. (Glagoly s pristavkami *pre-*, *pere-* i *predž-*). In

- Voprosy slovoobrazovanija i leksikologii drevnerusskogo jazyka*. Sbornik statej. Moskva: Nauka. 58-122.
- Uluxanov, Igor' S. 1977. *Slovoobrazovatel'naja semantika v rusckom jazyke i principy ee opisanija*. Moskva: Nauka.
- Uluxanov, Igor' S. 1996. *Edinicy slovoobrazovatel'noj sistemy russkogo jazyka i ix leksičeskaja realizacija*. Moskva: Russkie slovari.
- Uluxanov, Igor' S. 2004. *Slavjanizmy v rusckom jazyke (glagoly s nepolnoglasyymi prstavkami)*. Moskva: OOO "Upravlenie texnologijami".
- Ušakov, Dmitrij N. (ed.) 2008. *Bol'šoj tolkovyj slovar' sovremennogo russkogo jazyka*. Moskva: AST, Astrel'.
- Uspenskij, Boris A. 2002. *Istorija russkogo literaturnogo jazyka (11-17 vv.)*. Moskva: Aspekt Press.
- Valedinskaja, O.R. & Elena I. Golanova. 2007. O psixologičeskom i lingvističeskom aspektax termina "so-krizis". In Zemskaja, Elena A. and M.L. Kalenčuk (eds.) *Žizn' Jazyka*. Moskva: Jazyki slavjanskix kul'tur. 177-181.
- van Oss, Charles. 1989. *Aspekte der Intensivierung im Deutschen*. Tübingen: Narr.
- Vasmer, Max [Fasmer, Maks]. 1971. *Ètimologičeskij slovar' russkogo jazyka*. Moskva.
- Verhagen, Arie. 2010. Construal and Perspectivization. In Geeraerts, D. & H. Cuyckens (eds.) *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press. 48-82.
- Vinogradov, Viktor V. 1972. *Russkij jazyk*. Moskva: Vysšaja škola.
- Vinogradov, Viktor V. 1994/1999. *Istorija slov*. Moskva: Tolk.
- Vinogradov, Viktor V. et al. 1952. *Grammatika russkogo jazyka*. Vol. 1. Moskva.
- Vinogradova, Valentina N. 1972. Stilističeskije sredstva slovoobrazovanija. In *Stilističeskije issledovanija na materiale sovremennogo russkogo jazyka*. Moskva: Nauka. 175-245.
- Vinogradova, Valentina N. 1984. *Stilističeskij aspekt russkogo slovoobrazovanija*. Moskva: Nauka.
- Voronova, N. A. 1985. Slovoobrazovatel'naja struktura pristavočnyx otymennyx glagolov. In *Aktual'nye problemy russkogo slovoobrazovanija. Sbornik naučnyx statej*. Taškent. 154-157.
- Wade, Terence. 1992. *A Comprehensive Russian Grammar*. Oxford: Blackwell.
- Xaburgaev, Georgij A. 1974. *Staroslavjanskij jazyk*. Moskva: Prosveščenie.
- Yearley, Jennifer 1995. Jer vowels in Russian. In Beckman, J. & S. Urbanczyk & L. Walsh (eds.) *Papers in Optimality Theory*. Amherst, Mass.: GSLA. WEB. 533-571.
- Zaliznjak, Andrej A. 1980. *Grammatičeskij slovar' russkogo jazyka*. Moskva: Russkij jazyk.
- Zaliznjak, Anna A. & Aleksej D. Šmelev. 2000. *Vvedenie v russkuju aspektologiju*. Moskva: Jazyki rusckoj kul'tury.
- Zaliznjak, Anna A. 2001. Semantičeskaja derivacija v značenii rusckoj prstavki U-. In *Moskovskij lingvističeskij žurnal*. 5.1. Moskva: RGGU. 69-84.
- Zemskaja, Elena A. 1972. Okkazional'nye i potencial'nye slova v rusckom slovoobrazovanii. In *Aktual'nye problemy russkogo slovoobrazovanija. Sbornik statej*. Samarkand. 19-28.
- Zemskaja, Elena A. 1975. K probleme množestvennosti morfonologičeskix interpretacij (spornye slučai členenija proizvodnyx slov v sovremennom rusckom jazyke). In *Razvitie sovremennogo russkogo jazyka. Slovoobrazovanije. Členimost' slova*. Sbornik statej. Moscow: Nauka. 69-88.
- Zemskaja, Elena A. 1992. *Slovoobrazovanie kak dejatel'nost'*. Moskva: Nauka.
- Zemskaja, Elena A. 2006. *Sovremennyj russkij jazyk. Slovoobrazovanije*. Moskva: Nauka.

Žučenko, N. K. 1969. *Glagoly s pristavkami pre- i pere- v istorii ruskogo jazyka i ego dialektax*. Avtoreferat dissertacii. Saratov.