TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. B OSA - TOM. 363
HUMANIORA

Reflexive Space

A Constructionist Model of the Russian Reflexive Marker

by

Aki-Juhani Kyröläinen

TURUN YLIOPISTO
UNIVERSITY OF TURKU
2013

ISBN 978-951-29-5394-3 (PDF)
ISSN 0082-6987

Abstract

This study examines the structure of the Russian Reflexive Marker (*ся/-сь*) and offers a usage-based model building on Construction Grammar and a probabilistic view of linguistic structure. Traditionally, reflexive verbs are accounted for relative to non-reflexive verbs. These accounts assume that linguistic structures emerge as pairs. Furthermore, these accounts assume directionality where the semantics and structure of a reflexive verb can be derived from the non-reflexive verb. However, this directionality does not necessarily hold diachronically. Additionally, the semantics and the patterns associated with a particular reflexive verb are not always shared with the non-reflexive verb. Thus, a model is proposed that can accommodate the traditional pairs as well as for the possible deviations without postulating different systems. A random sample of 2000 instances marked with the Reflexive Marker was extracted from the Russian National Corpus and the sample used in this study contains 819 unique reflexive verbs.

This study moves away from the traditional pair account and introduces the concept of Neighbor Verb. A neighbor verb exists for a reflexive verb if they share the same phonological form excluding the Reflexive Marker. It is claimed here that the Reflexive Marker constitutes a system in Russian and the relation between the reflexive and neighbor verbs constitutes a cross-paradigmatic relation. Furthermore, the relation between the reflexive and the neighbor verb is argued to be of symbolic connectivity rather than directionality. Effectively, the relation holding between particular instantiations can vary. The theoretical basis of the present study builds on this assumption. Several new variables are examined in order to systematically model variability of this symbolic connectivity, specifically the degree and strength of connectivity between items.

In usage-based models, the lexicon does not constitute an unstructured list of items. Instead, items are assumed to be interconnected in a network. This interconnectedness is defined as Neighborhood in this study. Additionally, each verb carves its own niche within the Neighborhood and this interconnectedness is modeled through rhyme verbs constituting the degree of connectivity of a particular verb in the lexicon. The second component of the degree of connectivity concerns the status of a particular verb relative to its rhyme verbs. The connectivity within the neighborhood of a particular verb varies and this variability is quantified by using the Levenshtein distance.

The second property of the lexical network is the strength of connectivity between items. Frequency of use has been one of the primary variables in functional linguistics used to probe this. In addition, a new variable called Constructional Entropy is introduced in this study building on information theory. It is a quantification of the amount of information carried by a particular reflexive verb in one or more argument constructions. The results of the lexical connectivity indicate that the reflexive verbs have statistically greater neighborhood distances than the neighbor verbs. This distributional property can be used to motivate the traditional observation that the reflexive verbs tend to have idiosyncratic properties.

In addition to this, a set of argument constructions, generalizations over usage patterns, are proposed for the reflexive verbs in this study. In addition to the variables associated with the lexical connectivity, a number of variables proposed in the literature are explored and used as predictors in the model. The second part of this study introduces the use of a machine learning algorithm called Random Forests. The performance of the model indicates that it is capable, up to a degree, of disambiguating the proposed argument construction types of the Russian Reflexive Marker. Additionally, a global ranking of the predictors used in the model is offered. Finally, most construction grammars assume that argument construction form a network structure. A new method is proposed that establishes generalization over the argument constructions referred to as Linking Construction. In sum, this study explores the structural properties of the Russian Reflexive Marker and a new model is set forth that can accommodate both the traditional pairs and potential deviations from it in a principled manner.

Contents

Abbreviations

1P	first person plural
1S	first person singular
2P	second person plural
2S	second person singular
3P	second person plural
3S	third person singular
ACC	accusative
ADV	adverb
COMP	comparative
DAT	dative
F	feminine
FUT	future
GEN	genitive
IMP	imperative
INF	infinitive
INS	instrumental
M	masculine
N	neuter
NEG	negation
NOM	nominative
PL	plural
POSS	possessive
PPP	past passive participle
PR	preposition
PREP	prepositional
PRS	present
PRSPP	present passive participle
PST	past
RM	reflexive marker
SUP	superlative

Acknowledgements

This dissertation has been a long process spanning several years and the journey turned out to be less linear than I originally envisioned. Reflecting back, this project would not have been possible without the support and encouragement of numerous people whom I wish to thank. I express my gratitude to everyone who has offered both formal and informal contributions and with whom I have had the opportunity and pleasure to be acquainted during these years.

First, I would like to express my appreciation to my advisors Riitta Pyykkö and Tuomas Huumo. Riitta Pyykkö offered me a stimulating research environment in which she has always encouraged me to pursue my research topic further. Tuomas Huumo introduced me to Cognitive Linguistics and has guided me both formally and informally. I wish to thank Marja-Liisa Helasvuo who has always found the time to listen and discuss my ideas even before they have been fully formulated. I have also had the pleasure of collaborating with her on a separate project which has tremendously influenced my view on linguistics.

I am grateful to both of my preliminary examiners, Laura Janda and Hannu Tommola who both offered invaluable constructive criticism during the final phases of this project. Their expertise and evaluation of the manuscript is very much appreciated. Any errors that remain are mine alone.

The funding I received through Langnet, the Finnish Graduate School in Language Studies, made it possible to concentrate full-time on this dissertation. The seminars organized by the program of Grammar and Theory of Language provided an invaluable forum; thank you to Jussi Niemi and Urho Määttä for establishing such an encouraging environment. I would also like to thank all the postgraduate colleagues I got to know while part of Langnet.

Additionally, I am grateful to the boards of the foundations of Turun Yliopistosäätiö and Emil Aaltosen Säätiö for making it possible to attend the 2009 LSA Summer Institute in Berkeley. The courses held by Adele Goldberg, William Croft and Jennifer Hay at LSA 2009 helped to mold my understanding of usage-based models. Furthermore, DoRa funding made it possible to have research periods at the University of Tartu which were valuable for establishing new contacts and opened new opportunities. I would like to thank Renate Pajusalu, Liina Lindström and Kristel Uiboaed for productive and stimulating periods of research in Tartu.

I thank my friends and colleagues at the University of Turku, especially Emmi Hynönen, Mihail Voronov, and Tuomo Fónsen, who have helped and encouraged me through the years. I would also like to thank the numerous people with whom I have had the pleasure of working together in organizing conferences as part of the Linguistic Association of Finland and Finnish Cognitive Linguistics Association (FiCLA). These opportunities have contributed to my understanding of linguistics as a discipline. I wish to thank Dagmar Divjak and Steven Clancy for organization workshops that turned out to be an essential part of this study; these workshops paved the way for using R statistical programming environment. Additionally, I would like to thank Antti Arppe for methodological discussions which helped me to better understand the nexus between theory and method.

I would like to thank my family and friends who have supported me throughout this long process. I especially grateful to the following persons (in random order): Peter Svensson, Karl Olin, Piia Seppälä, Kevin Fitzpatrick, Pasi Kokkonen, and Vitaly Nikolaev. Without the support of my family this process would not have been possible. I thank my mother Rauli Sormunen, Jussi Sormunen, my sisters Jaana Aarnio and Johanna Söderholm, my brother Timo Kyröläinen, my father Markku Kyröläinen and my cousin Matti Aula. Finally, I would like to thank Vincent Porretta who has supported me through the years of this project.

Aki-Juhani Kyröläinen
Tartu 2013

1 Introduction

How are complex categories formed and maintained? How can surface structures be used to form abstractions? How are constructions, form-meaning pairings, interconnected in a network? These are the questions explored in this study based on the distributional and structural properties of the Russian Reflexive Marker (*-ся*/*-сь*). Herbert's (1962:468) definition of a complex system in his seminal essay serves as a starting point for the exploration: “Roughly, by a complex system I mean one made up of a large number of parts that interact in a nonsimple way.” Importantly, Herbert argues that complex systems are inherently hierarchical but at the same time they consist of subsystems. These notions mesh smoothly with usage-based models of grammar where networks and different levels of granularity are recognized (Bybee, 1985; 2010; Goldberg, 2006; Langacker, 1988b).

This study moves away from the traditional pair account where the reflexive verbs are accounted for relative to the non-reflexive verbs. Instead, the concept of the Neighbor Verb is introduced. A neighbor verb exists for a reflexive verb if they share the same phonological form excluding the Reflexive Marker extending the theoretical basis of the proposed model to lexical network models on morphological structures and (computational) psycholinguistic models on the mental lexicon (Altieri, Gruenenfelder & Pisoni, 2010; Baayen & Moscoso del Prado Martín, 2005; Bybee, 1985; 2010; Chan & Vitevitch, 2009; Geeraert & Kyröläinen, in prep.; Steyvers & Tenenbaum, 2005; Vitevitch, 2008). The concept of the Neighbor Verb is the primary building of the model, opening the path to network structures and to rigorous quantification of the cross-paradigmatic relation in terms of lexical densities, distances, and perceived semantic similarities. Thus, the traditional concept of pair is incorporated by proxy, but it constitutes a subtype in the network. Similarly, reflexive verbs that lack the cross-paradigmatic relation can be accounted for in a natural manner without a priori excluding them. Importantly, properties of the Reflexive Marker can be modeled through a single system without postulating radically different systems.

The data are based on a stratified random sample (N = 2,000) extracted from the Russian National Corpus containing 819 unique reflexive verbs and 717 unique neighbor verbs. The cross-paradigmatic relation holding between the reflexive and the neighbor verbs is modeled through a lexical network defined as the Neighborhood covering over 378,000 verb forms across the two paradigms in the database.

The Reflexive Marker is one of the central morphological categories of Russian verbs in addition to aspect, a generalization holding for all Slavic languages. It covers an impressive array of different functions forming a complex category. The complexity of the category stems from the fact that the Reflexive Marker has penetrated most of the categories associated with verbal semantics and structures in Russian: voice, aspect, personal versus impersonal types, inflectional versus derivational patterns, among other things. Needless to say, the history of the study of the Russian Reflexive Marker is as impressive as the number of categories associated with it. The aim of this study is to offer a contribution to this history through five key points of interest: 1) a usage-based approach built on surface structures, 2) the concept of Neighbor Verb and Neighborhood instead of derivational binary relations, 3) a gradient structure between the reflexive and neighbor verbs, 4) a model of lexical network, and 5) a data-driven approach to form links between argument constructions.

This study operationalizes the concept of the Construction in terms of 25 variables, allowing exploration of the potential interconnectedness of different argument construction types and their slots. Thus, argument constructions are modeled as probabilistic structures, abstractions over verb-specific constructions. Random Forests, (machine learning algorithm), are introduced to model the input. Random forests are built from the input similar to generalizations formed in inductive learning. Thus, a conceptual connection is established between the principles of Construction Grammar and the statistical method exploited in this study, (i.e., inductive learning).

Consequently, the model allows accounting for interconnections in a quantifiable manner and as a gradient structure rather than a priori postulated pairs (Hay & Baayen, 2005). To motivate the patterning that was obtained from the structure of the lexical network, two domain-general principles are argued to be crucial in the formation and maintenance of complex categories. First, the Hypothesis of Connectivity states that the connections between items increase over time. Second, the Hypothesis of Distance states that the distances between items decrease over time. These two domain-general processes are considered to be the motivational pathways behind the structural properties of the Reflexive Marker and the cross-paradigmatic relations, in general.

The following sections discuss previous studies on the Russian Reflexive Marker establishing the background and connection to the proposed constructionist model. Section 1.2 gives an outlook to previous taxonomies of the Russian Reflexive Marker. A more elaborated discussion is offered in Chapters 5–9 that cover the proposed set of the argument constructions supported by the Russian Reflexive Marker. The data and sampling frame are described in Section 1.4. The organization of the study is given in Section 1.5.

1.1 Prerequisites and Objectives

The definition of the label Reflexive Marker follows the morphologically-orientated convention typically upheld in the Russian linguistic tradition (Geniušienė, 1987; Янко-Триницкая, 1962).[1] Thus, the label is attributed to every verb carrying the Reflexive Marker regardless of the semantic range the verb may display in usage. A single label highlights the fact that the various instantiations form a category which in turn requires positing an explanation to

[1] The label marker is used as no stance is taken whether the *-ся* should be analyzed as a postfix, suffix or particle, for example (Paducheva, 2003:174; Князев, 2007:337; Янко-Триницкая, 1962:32). Extensive discussion on the morphological status of the Russian Reflexive Marker is given in Janko-Trinickaya (Янко-Триницкая, 1962:34-36).

account for the various instantiations. Additionally, the various reflexive verbs constitute a system and not a list of some possible range of random types. The traditional label is suitable to underline the principle of interconnectedness advocated in this study.

In order to situate a constructionist approach in relation to previous studies on the Russian Reflexive Marker, and generally on the Russian linguistic tradition, a short exodus is required. A more elaborated picture is given in Section 1.2. A trend arises from the previous studies on the Russian Reflexive Marker. The early Russian tradition was primarily concerned with establishing taxonomy of meanings or functions associated with the Reflexive Marker (Виноградов, 1972; Шахматов, 1925; Янко-Триницкая, 1962). These taxonomies do not, however, have any inherent structure. They are pure listemes. Taxonomies are, nonetheless, a prerequisite for possible future refinements. In contrast, formation of the concept of diathesis; as it is originally laid out in the article by Mel'chuk and Holodovich (Мельчук & Холодович, 1970) and developed further in the Saint-Petersburg Typological School (Храковский, 1974; 1978b; 1981), moved towards establishing systematic connections between verbs and argument structures.[2] At least in the early version, diathesis is considered to be universal (Храковский, 1978a:51). Following Paducheva (Падучева, 2004:51-52), a diathesis can be defined in modern terms as a change in the semantic roles and their corresponding syntactic positions, cf. Section 2.2 (cf. Mel'čuk, 1993; Mel'čuk, 1997).

Certain aspects of diathesis are in close proximity with the constructionist approach as argument constructions are typically viewed in relation to this configuration discussed in Chapter 2 (cf. Barðdal, 2008; Croft, 2001; Goldberg, 2006). A crucial difference is present, nonetheless, as the diathesis is inherently viewed as a binary relation, which is a derivational relation from the unmarked form to the marked form following structuralism and Jakobsonian markedness theory, for example *мыть* 'wash' and *мыться* 'wash oneself' (Падучева, 2002; Якобсон, 1985a).[3] Thus, the derivational relation is established through pairs.

There are, however, several deviations from this pair relation. The first deviation is established with reflexiva tantum, or deponent verbs, which only appear in the reflexive form, for example *бояться* 'be afraid'.[4] The second

[2] Studies published in Cyrillic were transliterated based on the GOST-standard. In other cases, the Romanization given in the study was used.

[3] Russian Grammar by Horálek is an excellent example of applying the principles of markedness. Jakobsonian approach defines the unmarked form as a negation from the marked. Thus, Horálek (1979:267) proposes a distinction between the Passive and Non-Passive (Active) in Russian.

[4] The Russian tradition dominantly uses the term reflexiva tantum whereas the Western tradition uses the term deponent. Traditionally, the term deponent is used to refer to a set of Latin verbs marked with *-r* semantically defined as passives with active meaning. In addition, the term deponent can be used to refer to any defective paradigm. The term media tantum is commonly used by Greek scholars. (Baerman, 2007).

deviation is manifested with reflexive verbs that differ in meaning compared to the non-reflexive verb, for instance *оказаться* 'seem, appear' and *оказать* 'render' cf. Section 3.1.1. Although Geniušienė (1987) has already argued that these types of verbs should be included in the analysis, in reality, they are almost always excluded regardless of the theoretical framework (Fehrmann, Junghanns & Lenertová, 2010; Guhl, 2010; Калашникова & Сай, 2006; Князев, 2007). This issue is related to another imbalance that sample-based studies are virtually non-existent. To the best of knowledge, there is only one published study on the Russian Reflexive Marker, reported in Kalashnikova and Saj (Калашникова & Сай, 2006).[5]

Kalashnikova and Saj (Калашникова & Сай, 2006) show that as a type this deviant set of reflexive verbs is not merely some anomaly, as their sample (N = 4637 reflexive verbs) contained 54% of these "deviant" reflexive verbs in Russian. A similar finding is also presented in Geniušienė. She establishes 700 "deviant" reflexive verbs based on her Lithuanian verb list (Geniušienė, 1987:150). Thus, defining the "deviant" reflexive verbs as a peripheral category is simply an artifact of a priori data selection.

A derivational account is also inherently confined to directionality, (i.e., from the non-reflexive to the reflexive form yielding the distinction between the base and the derived form). However, evidence to support this unidirectional analysis is lacking. Diachronic evidence shows that both directions are possible as discussed in Section 3.1.1 (Кузнецова, М. В., 1984). Furthermore, Hay (2001; 2002) has demonstrated based on corpus and experimental data that the perceived basic form is dependent on the relative frequency, the logarithmic ratio, between the base, (e.g., *iterate)*, and the derived form, (e.g., *reiterate*) in English (cf. Bybee, 2010:46-48).

Hay proposes that parsing is one of the contributing factors to this phenomenon. For instance -*ment* is likely to be parsed in *discernment* due to the higher frequency of *discern* compared to *discernment* and the reverse holds in *government* ~ *govern* (Hay, 2002:534-535). Thus, directionality appears to be relative. Moreover, the vast majority of models of the mental lexicon posit large network structures for the lexicon with a varying degree of connectivity (Altieri et al., 2010; Baayen & Moscoso del Prado Martín, 2005; Chan & Vitevitch, 2009; Geeraert & Kyröläinen, in prep.; Steyvers & Tenenbaum, 2005; Vitevitch, 2008).

This study builds on the previously outlined background and offers a constructionist model on the Russian Reflexive Marker, situating the proposed model against a probabilistic view of linguistic structure. Recent studies in Cognitive Linguistics and Construction Grammar, such as Gries (2007) and Divjak (2009), only to name a few, are connected to a larger body of studies which take usage patterns as a starting point, such as Bresnan et al. (2007) on dative alternation, Bod (2006) on phrasal patterns, Hay (2001; 2002) on morphologically complex words, and Arppe (2008) on near-synonymy of four

[5] They (Калашникова & Сай, 2006:1 footnote 2) cite one other unpublished study by Korolev (Королев, 1968).

Finnish 'think' lexemes. Another possibility is to operate on symbolic manipulation, (i.e., rules). In terms of the Russian linguistic tradition, this approach is most notably present in the Meaning–Text theory initiated by Zholkovskij and Mel'chuk (Жолковский & Мельчук, 1965). Mel'chuk (Мельчук, 1995) has pioneered the theory substantially further.[6] In short, the difference between these two approaches lies in the manner of how language as a system is to be modeled either based on probabilities or symbolic manipulation, (i.e., rules). In this regard, certain properties of these two approaches may yield similar results, but their ontological and theoretical basis are, however, incompatible (cf. Blevins, 2006). On the other hand, the probabilistic approach is easily compatible with corpus data, experimental studies on lexical processing, diachronic change, and language acquisition because probabilistic distributions incorporate variation whereas rules do not (Bybee, 2010; Goldberg, 2006; Tomasello, 2003).

Nonetheless, Construction Grammar is still a fairly young theory. Studies have focused more closely on idiomatic patterns rather than forming a larger network of construction types in language as is also noted by Goldberg (2006:14). Herbst (2010:246) has expressed a similar view in stating that although Construction Grammar appears promising, the majority of the studies have focused on relatively few construction types. Recently, a number of studies have emerged where a larger number of construction types are presented in the network model such as Divjak and Janda (2008), Kyröläinen (submitted) on Russian construction types, and Barðdal (2008) on Icelandic.

The claims laid out in this form may appear at first to depart from the Russian linguistic tradition, especially from the diathesis tradition. This statement is only partly true as the basic principles utilized in this study have been, at least partly, formulated in the diathesis tradition and generally in the Russian Linguistic tradition in various forms. However, the difference is that the principles advocated in this study have not been systematically exploited or incorporated to form a single coherent model. First, the importance of frequency is explicitly proposed by Hrakovskij. He argues that the distinction between the basic and the derived structure is defined in terms of stylistic neutrality and frequency of use. (Храковский, 1974:13). Second, the non-categorical characteristics of linguistic structure are also promoted by Geniušienė (1987:59): "The probabilistic approach to language […] adopted in this study regards language as a continuum of diffuse phenomena (units and categories) merging into one another." Finally, Zolotova (Золотова, Г. А., 2005 [1973]) has already proposed an account which combines basic or core

[6] Although the Meaning–Text theory is poorly known outside the Slavic circle of linguists, there is a substantial syncretism between Meaning–Text, diathesis, and lexical semantics generally known as the Moscow Semantic School (Апресян, Ю. Д. , 2005). Applications range from building computational models of the lexicon to syntactic parsing, for example, the newly implemented Dependence Treebank of Russian (SyntaxRus) which is a subcorpus in the Russian National Corpus.

situational meanings with sentence patterns. For example, the core meaning of Subject of Action corresponds to the pattern N(ominative) V(erb) Acc(usative) in Russian. In constructionist terms, this pattern constitutes the prototypical Transitive Construction combining both form and semantics. A similar position is taken by Leinonen (1985) in her comparative study between the Russian and Finnish impersonal types. Thus, there is a considerable syncretism between different accounts if one is willing to look for them.

At the same time, Knyazev has criticized functional approaches in a sense that a separation between the reflexive verb and the Reflexive Marker is not maintained. For example semantically motivated types such as the motion or the mental event are often evoked in functional approaches (cf. Kemmer, 1993; Manney, 2000). His criticism concerns the relation between the reflexive and the non-reflexive verb as the semantic component of motion does not make a distinction between them, for example *мчаться* 'rush' and *мчать* 'rush' (Князев, 2007). It is possible to agree with Knyazev's statements with certain reservations. First, his position presupposes a perfect separation between the verb and the Reflexive Marker, which is a fully compositional view on semantics. Second, Knyazev achieves this separation by a priori excluding those reflexive verbs that do not form pairs. Third, the distinction assumes that categories are established through necessary and sufficient criteria.

In contrast, the position advocated in this study builds on the idea that multiple cues are available in language. The distinction between the reflexive verb and the Reflexive Marker constitutes a difference in schematicity in this study. The reflexive verbs are associated with the verb-specific constructions whereas the argument constructions are a generalization over them. Lastly, the Reflexive Marker constitutes another level of schematicity established through the links between verb-specific and argument constructions as discussed in Section 10.2.

Another important property of Construction Grammar is that the Argument Construction assigns a clausal meaning to phrasal patterns; albeit abstract, it may be true in certain cases (Goldberg, 2009b). Consequently, this study is also a step towards disambiguating reflexive verbs in text, as such argument construction types as the Motion Construction is an abstract clausal meaning. Thus, the criticism expressed by Knyazev amounts to a different research question altogether - whether or not the same set of the variables can be used to model the difference between usage patterns of *мчаться* 'rush' and *мчать* 'rush'. I refer to such studies as those by Bresnan et al. (2007) and Arppe (2008) on modeling these kind of data where the locus is in differentiating semantically similar usage patterns.

In short, this section outlined the background of this study. Specifically, this study moves away from binary relations towards lexical network models, which posit degrees of connectivity between words. Additionally, words that occupy the basic level of description in lexical network models and linguistic categories are abstractions over them. For example, the category of the Reflexive Marker is an abstraction over lexical reflexive verbs. Similarly, traditional grammatical functions such as the subject are viewed as an abstraction over specific usage patterns. This form of abstraction is referred to as schematicity in this study, cf. Section 2.2.2. The proposed usage-based model is situated against the background of Construction Grammar and Cognitive Linguistics. An introduction to Cognitive Linguistics is given in Section 1.3.

In addition to these assumptions, a probabilistic view on linguistic structures is closely related to utilizing machine learning algorithms. This study builds on classification and regression trees. An introduction to this family of methods is given in Section 1.4.2. Additionally, Chapter 4 introduces an ensemble method called Random Forests, which utilizes classification and regression trees. Importantly, a conceptual connection between the assumptions of Construction Grammar and Random Forests is exemplified. Before turning to the proposed model, an overview of previous studies on the Russian Reflexive Marker is given in Section 1.2. Lastly, Section 1.5 outlines the organization of the present study.

1.2 Previous Studies on the Russian Reflexive Marker

The number of studies dedicated to exploring the Reflexive Marker is voluminous to say the least. An attempt to give a complete survey would constitute an exploration into the historical development of the linguistic tradition in general, a formidable goal deserving a dedicated study of its own. Hence, the following sections are delimited in scope, only outlining the ideas and principles behind the early Russian tradition as these studies constitute the body of the taxonomical approaches. Additionally, previous cognitive and functional studies are briefly illustrated.

Geniušienė (1987:15) outlines two broad approaches in studying the Reflexive Marker. The first is primarily concerned with the attested variation in a language, labeled as the taxonomic approach. On the other hand, the second approach is dedicated to teasing apart the most abstract representation, primarily concerned with the invariance of the Reflexive Marker, labeled as the anti-taxonomic approach. Although the invariant position is prominently present in the formal approaches while the anti-taxonomic figures prominently in the early Russian tradition, the demarcation between them is, nonetheless, less pronounced. The majority of the studies on the Russian Reflexive Marker tend to be top-heavy in a sense that finding an invariant meaning figures are prominently in the literature, partly following the tenets laid out in the works of Jakobson (Якобсон, 1985b). Nonetheless, the possible range of meanings or functions associated with the Reflexive Marker is highly verb dependent, creating a degree of variability in the previous studies. Thus, the following sections are dedicated to the labels identified in the early Russian studies as subsequent works continue from it.

1.2.1 Aspects of the Russian Reflexive Marker

Russian, among other Slavic languages, employs a system of two etymologically related markers to encode the function of reflexivity, the coreference of the

Agent and the Patient.[7] Haiman (1983:781-819) characterizes these as the heavy and the light marker. In the case of Russian, the reflexive pronoun *себя* '-self' constitutes the heavy marker, syntactically independent, while the light Reflexive Marker has two allomorphs. The variation depends on the inflectional form of the verb. The allomorph -*сь* appears after vowels and the -*ся* after consonants. In contrast, the participle forms always appear with the -*ся* form. Additionally, the light marker is attached to the stem after other morphological markers, such as person or number and gender with participle forms, in Russian. Typologically, in a language employing a two-marker system, the light marker tends to be polysemous, covering a range of different functions in addition to the reflexive as is the case in Russian (Geniušienė, 1987; Haiman, John, 1983; Kemmer, 1993). However, there are borderline idiomatic patterns in Russian where both forms are used, for example *замкнуться в себе* 'retreat into oneself', *затаиться в себе* 'hide ones feelings' and *копаться в себе* 'ponder about oneself'.

Diachronically, the light form appeared as clitic and had two cases: the accusative and the dative. Through diachronic changes, the case distinction was lost forming a single marker. Additionally, the clitic fused with the verb in East Slavic languages (Данков, 1981:62-63; Зарицкий, 1961:10-13; Князев, 2007:260). A typologically contrastive study of Slavic languages is presented in Knyazev (Князев, 2007). The clitic form was still used as late as the 17th century according to Sobolevskij (Соболевский, 2004 [1907]:256).[8] In contrast, the function of the dative is poorly documented and, traditionally, it is only mentioned in passing in diachronic studies (Данков, 1981; Зарицкий, 1961:62-63). Sobolevskij (Соболевский, 2004 [1907]) explicitly states that the dative was rarely used. Recently, Kuznecova (Кузнецова, М. В., 1984) considers that the variation between the accusative and the dative case is connected to the animacy of the subject on the one hand and to volitionality or accomplishment of the event on the other based on the diachronic development of the Reflexive Marker.

Generally, the expansion of the light form is connected to the reorganization of voice system in Indo-European languages as a replacement for the ancient Indo-European synthetic mediopassive (Barðdal, Cennamo & Eythórsson, submitted). However, the distinction between the active and middle never had a formal category in Slavic. From a diachronic perspective, the light form came to occupy the functional region to express the previous middle types (Данков, 1981:68; Мейе, 2001 [1951]). Thus, the light reflexive marker constitutes a complex category. Not only is the marker layered with diachronic development,

[7] The Russian tradition typically defines the reflexivity in terms of grammatical function by positing a coreference between the subject and the object. The definition based on semantic roles is typically employed in typological/functional studies as these do not impose a morphological analysis.

[8] Sobolevskij (Соболевский, 2004 [1907]:256) makes some interesting observations. Chronicles from the 15th and 17th centuries contain a number of double usages of the Reflexive Marker, (i.e., both the Reflexive Marker and the clitic).

but also interconnected to the grammaticalization of transitivity in Russian.

Returning to the synchronic perspective, Russian Grammar distinguishes seven categories, and additional subcategories, the status of which is left unspecified (Шведова & другие, 1982:617-618). The position outlined in Russian Grammar continues the work established by Shahmatov (Шахматов, 1925) and Vinogradov (Виноградов, 1975). Table 1.2-1 gives the classification in Russian Grammar. The English labels are taken from Gerritsen (1990).

	Reflexive function	Translation	Instantiation
1)	собственно-возвратное значение	Semantic Reflexive	*мыться* 'wash'
2)	взаимно-возвратное значение	Reciprocal Reflexive	*целоваться* 'kiss each other'
3)	косвенно-возвратное значение	Indirect Reflexive	*строиться* 'build for oneself'
4)	активно-безобъектное значение	Active-Objectless	*кусаться* 'bite'
5)	характеризующе-качественное значение	Passive-Qualitative	*растворяться* 'desolve'
6)	общевозвратное значение	General Reflexive	*сердиться* 'be angry'
7)	побочно-возвратное значение	Secondary Reflexive	*держаться* 'hold'

Table 1.2-1 The classification of the Russian Reflexive Marker in Russian Grammar.

The identified meanings have the following characterizations in Russian Grammar established through grammatical roles rather than on semantic roles:

1) The coreference of the subject and the object
2) The reciprocal action between the subjects
3) The verb denotes an activity performed by the subject for one's own interest
4) The action qualifies a permanent and characteristic property of the subject
5) The action qualifies an inclination of the subject
6) The action is confined to the sphere of the subject
7) The action is directed towards the object

Additional subgroups without explicit definitions are proposed for the verbs which have an intransitive non-reflexive verb pair, for example, *грозить* ~ *грозиться* 'threaten' and *белеть* ~ *белеться* 'whiten.' Additionally, the combination with the Reflexive Marker and a prefix is considered as a separate process of word formation. These are discussed in Section 3.1.1 because these forms tend to break away from the cross-paradigmatic relation between the reflexive and the neighbor verb. Finally, verbs which do not have non-reflexive forms are considered to constitute a separate group, the so-called reflexiva tantum verbs (Шведова & другие, 1982:617-618). In the early Russian linguistic tradition, the reflexiva tantum verbs are typically labeled as общие 'general' reflexive verbs. The label can be traced, at least, back to early studies of Russian grammar, for instance in Fortunatov (Фортунатов, 1899). The classification proposed by Vinogradov (Виноградов, 1972) closely follows Shahmatov (Шахматов, 1925) given in Table 1.2-2.

	Reflexive function	Translation	Instantiation
1)	собственно-возвратное значение	Semantic Reflexive	*защищаться* ‘defend oneself’
2)	средне-возвратное значение	Medial Reflexive	*возвращаться* ‘return’
3)	общевозвратное значение	General Reflexive	*сердиться* ‘become angry’
4)	страдательно-возвратное значение	Passive Reflexive	*освещаться* ‘light’
5)	взаимно-возвратное значение	Reciprocal Reflexive	*целоваться* ‘kiss each other’
6)	косвенно-возвратное значение	Indirect Reflexive	*строиться* ‘build for oneself’
7)	побочно-возвратное значение	Consequential Reflexive	*держаться* ‘hold’
8)	средне-пассивно-возвратное значение	Medial-Passive Reflexive	*представляться* ‘occur, seem’
9)	качественно-пассивно-возвратное значение	Qualitative-Passive Reflexive	*сгибаться* ‘bend’
10)	активно-безобъектное значение	Active-Objectless	*кусаться* ‘bite’
11)	интевсивно-побочно-возвратное значение	Intensive-Consequential Reflexive	*звониться* ‘call’
12)	пассивного обнаружения внешнего признака	Passive acquistion of outer quality	*белеться* ‘whiten’
13)	косвенно-результативно-возвратное значение	Indirect-Resultative Reflexive	*проспаться* ‘sleep it off’
14)	взаимно-моторное значение	Reciprocal-Motorical Reflexive	*срастись* ‘grow together’
15)	безлично-интенсивное значение	Impersonal Intensive	*хотеться* ‘want’

Table 1.2-2 Classification of the Russian Reflexive Marker based on Vinogradov.[9]

The classification exemplified by Vinogradov can be considered as the basic inventory of the different meanings marked by the Russian Reflexive Marker

[9] The label Consequential Reflexive (побочно-возвратное значение) is adapted after Gerritsen (1990).

and subsequent studies build around this inventory as is evident in the classification proposed in Russian Grammar. Nonetheless, three major categories identified in the early Russian tradition are not included in Russian Grammar, namely the Medial-Reflexive, Passive Reflexive, and Medial-Passive Reflexive. Additionally, case marking is not included in the classifications although certain categories are stated as being sensitive for certain case patterns. For example, the Medial-Passive Reflexive is typically attested with the dative case of the subject, cf. Section 3.2.5.

Another important observation is made by Geniušienė related to forming a taxonomy of the Russian Reflexive Marker, namely that the guiding principles are typically based on inconsistent criteria. For example, semantic reflexives are formed based on their relation to non-reflexive verbs and impersonal types on syntactic criterion, such as agreement (Geniušienė, 1987:12-13). To remedy the situation, Geniušienė distinguishes semantic and syntactic taxonomies as separate sets. The division between syntax and semantics is one possible solution to the mismatches in taxonomies, but a verb can display multiple argument constructions leading to a situation where multiple criteria would be applicable to a specific verb. However, multiple argument constructions are rarely discussed in the studies of the Russian Reflexive. For example, these types are briefly mentioned in Geniušienė (1987), Gerritsen (1990), and Zarickij (Зарицкий, 1961). A citation from Geniušienė (1987:141) illustrates the issue at hand: “As we see, polysemy of RVs [reflexive verbs] may be determined by a number of factors, and a more detailed study might lead to a better understanding of the systematic nature of verbal lexical-semantic types and their interrelation.” At least in principle, a constructionist approach avoids these pitfalls as the concept of the Construction is a form-meaning pair by definition.

Vinogradov (Виноградов, 1972:496) makes an important observation regarding the General Reflexive meaning by stating that there are reflexiva tantum verbs, such as *бояться* ‘be afraid,’ which closely resemble these verbs, an observation which is central for the purposes of the present study. The phenomenon in question is not limited to this particular class but extends through the system as a whole. For example, there are reflexiva tantum verbs, such as *касаться* ‘touch,’ which could be considered as instantiating the class of the Secondary Reflexive. This illustrates that the reflexiva tantum verbs are not some isolated set of verbs but are integrated into the system. Thus, the separation between the reflexive and reflexiva tantum verbs is not of meaning, but a division based on theoretical inclinations (cf. Исаченко, 1960:402-403).

There are two later studies which are explicitly built upon the early Russian tradition: Israeli (1997) and Gerritsen (1990). Israeli’s (1997) study leans towards the discourse properties of the Reflexive Marker similar to Glazkova (2011). The merits of Gerritsen's account are considerable. First, she includes morphology in the classification, such as case patterns. Second, multiple patterns of a particular reflexive verb are, albeit briefly, included. Third, both personal and impersonal types are exemplified. Fourth, copula constructions such as *являться* ‘be,’ are considered to be part of the system of the Reflexive

Marker. Generally, this class is disregarded in taxonomies. These four factors make it possible to compare classifications. Both of these studies are used as a frame of reference in establishing the set of the argument constructions used in this study.

However, one important distinction needs to be stated related to previous studies. Isachenko (Исаченко, 1960) posits that the Passive Construction and the impersonal types constitute a form whereas all other reflexive verbs pertain to word formation (cf. Храковский, 1978a:50).[10] A similar position is taken in Russian Grammar (Шведова & другие, 1982:617). Related to this, Israeli (1997) considers that the Passive Construction is a true form and part of the paradigm of non-reflexive verbs, whereas all other types are simply reflexive verbs. Once multiple patterns are factored in, the status of the Reflexive Marker is in flux; depending on the configuration, it is attested. Vinogradov (Виноградов, 1972:496) already illustrated the issue with the Passive Construction, for example *щеки румянились морозом* 'the cheeks turned red because of frost' versus *она румянилась* 'she blushed'. The dichotomy between the Passive Construction and the reflexive verb leads to a situation where the number of homonyms is multiplied in Russian, such as *румяниться* in the Passive Construction versus all other instantiations. Contrary to this mode of analysis, it is argued in this study that the verbs marked with the Reflexive Marker constitute a single paradigm.

There are a small number of studies building on Cognitive Linguistics for describing the Russian Reflexive Marker. The commonality between these studies is found in the identification of the core or central meaning and then extending the other possible meanings through it. In other words, cognitively oriented studies posit a prototype which serves as the basis for the extensions. The Semantic Reflexive is taken as the starting point. The diachronic evidence supports this position although extending a diachronic sense to synchronic description posits a tacit assumption that the core of a category does not undergo changes over time. In the case of the Russian Reflexive Marker, the assumption makes a strong claim, albeit implicit, that regardless of the possible changes in the category, the Semantic Reflexive has retained in status as the prototype.

Generally, the advantage of the cognitive approaches is that the attested types supported by the Reflexive Marker are not presented as constituting a mere list. Instead, the classifications attempt to systematically link the various instantiations building relationships between them. The earliest cognitive study is offered by Janda (1993a) contrasting the Russian and the Czech Reflexive Markers. Enger and Nesset (1999) combine Cognitive Linguistics with Kemmer's (1993) classification and diathesis tradition (Geniušienė, 1987). A third study is offered by Ahn (2005) positing two prototypes: the semantic

[10] Bybee (1985) offers a model to account for the relation between inflection and derivation as continuum rather than a dichotomy, cf. section 10.1.2. However, any possible formulation between these two types is left for future studies on whether a specific argument construction is closer to inflection or derivation.

reflexive and the passive that functions as centers for other meanings of the Reflexive Marker. If a structure established through a prototype is sought after, an analysis based on multiple centers may be considered relevant to the Russian Reflexive Marker as the center might not directly connect all the possible types. Instead, a subtype may establish its own extensions, which would not be any more directly connected to the center (Geeraerts, 1988; 1992; Tuggy, 1993).

In relation to prototypicality, Williams' (1999) case study is an interesting extension and a welcome addition to the cognitively orientated studies. Unfortunately, to my knowledge, the line of investigation has not since conducted. The case study builds on the hypothesis that the Russian Reflexive Marker has a function of prototypicality labeled as Prototype-Accessing. Accordingly, the Reflexive Marker is used in situations which are prototypical denotations for a particular verb compared to the non-reflexive one. This assumption meshes smoothly with the possibility of having multiple centers within a category, as certainly not all verbs can be considered to have this function. Williams constructed a small-scale questionnaire study ($N = 11$) on six verbs depicting facial expressions, such as *щурить* 'squint' ~ *щуриться* 'squint,' to explore the Prototypicality-Access hypothesis. A context which might be considered as illustrating the prototypical situation type for a particular verb was given. Participants were asked to choose a pattern from the option list depicting the transitive construction and two reflexive verb constructions, either a body part occupying the subject position, the object of the transitive scene, or an animate subject occupying the subject position, the subject of the transitive scene. Williams reports that when the prototypical scene was given, the participants preferred the last option and, in the case of less prototypical scenes, the transitive was selected more often (Williams, 1999).

From a typological perspective, Kemmer's (1993) study of the middle voice can be considered as a landmark and the categories established are commonly used in current typological studies (Bostoen & Nzang-Bie, 2010; Moyse-Faurie, 2008). However, as her classification occupies a prominent position in this study, a brief overview of the labels proposed in her study is given here as they are examined more thoroughly in the chapters devoted to the analysis of the construction types. A division is established between a general reflexive situation and grooming and body care. The latter depicts a situation type where actions are performed on one's own body. A similar delimitation is proposed for reciprocal and natural reciprocal events. The latter corresponds to a situation which inherently involves multiple referents, for example, *договориться* 'agree,' and *бороться* 'fight'. In the Russian tradition, these instances coincide with the label lexical reciprocals (cf. Вимер, 2001; Князев, 2007).

Motion events fall into three categories. The first is nontranslational, such as *шататься* 'wobble,' while the second is translational motion, such as *двигаться* 'move.' A third relation is established through a change in the body posture such as *садиться* 'sit down'. Indirect middle is depicted with such verbs as *строиться* 'build for oneself'. Mental events are furthermore divided into several subgroups, for example, emotion middle, such as *волноваться* 'worry,' and cognition middle,such as *удивляться* 'wonder.' A final large group depicts spontaneous events such as *сузиться* 'become narrow.' The Passive Construction, however, is excluded from her classification.

This section illustrated the main taxonomical approaches to the Russian Reflexive Marker. Although a new set of labels may be proposed (cf. Gerritsen, 1990), some verbs may be moved around between different classes (cf. Israeli, 1997) or a certain class may be given a more fine-grained division such as the impersonal types (cf. Галкина-Федорук, 1958). The basic inventory of the Russian Reflexive Marker is organized around the classes established in the early Russian tradition. A more detailed discussion of the previous taxonomies is given in Chapters 5–9 when the proposed set of argument constructions is discussed.

1.2.2 Invariance of the Russian Reflexive Marker

In addition to describing the various instantiations of the Reflexive Marker, establishing a unity or commonality within a category figures prominently in the previous studies. This is perhaps best described as a hunt for the most abstract meaning possible, as the number of proposed invariant meanings is substantial to the point that most accounts try to establish some new invariant meaning. Certainly, an invariant structure corresponds to the principles of establishing the most elegant and parsimonious account for a certain phenomenon. At the same time, the taxonomic and invariant positions are not necessarily in opposition to each other as both positions play a role in previous accounts.

Perhaps the most prominent invariant function of the Russian Reflexive Marker is that of intransitivity, already established in the early Russian tradition (Виноградов, 1972). This invariant function is also upheld, at least in the early formal approaches (Babby, 1975). However, even the concept of intransitivity is labeled differently depending on which aspect is treated as the most prominent. For example, Babby (1975) defines the invariant function as syntactically derived intransitivity, whereas Isachenko (1974:59; Исаченко, 1960:353, 374) defines it as a case of explicit signalization of intransitivity. Nonetheless, it seems that for Babby, a semantic invariant function seems a questionable endeavor. He claims that the function of the Russian Reflexive Marker is to signal that the underlying transitive verb is intransitive in the surface structure (cf. Babby, 1975:298).

Interestingly, intransitivity is viewed as a continuum, if interpreted in modern terms. Already in the early Russian tradition, for example, Vinogradov (Виноградов, 1972) states that the Reflexive Marker signals an increase in intransitivity. A definition is needed allowing to combine both the transitive and non-transitive reflexive verbs, although this definition has never been developed into its full potential compared to the classical paper by Hopper and Thompson (1980) where transitivity is explicitly defined as a continuum.

However, there is a substantial body of reflexive verbs in Russian that do not follow this pattern as was already illustrated in Section 1.1. There are a number of issues of deriving the reflexive verbs directly from the transitive verbs. First, there is no diachronic evidence to support the view that all reflexiva tantum verbs were originally non-reflexive, as is demonstrated by Krys'ko (Крысько, 1984). Second, a problematic property of reflexive verbs is that some of them can appear with accusative object, such as verbs *бояться* 'be afraid' and *слушаться* 'obey,' especially in spoken Russian. This phenomenon is discussed, for instance, by Janko-Trinickaya (Янко-Триницкая, 1962:60, 70-71), while additional examples and references are given in Israeli (1997), and in Knyazev (Князев, 2007), as well. Third, even from a diachronic point of view, Krys'ko shows that certain reflexive verbs have been used in the Transitive Construction, especially verbs of motion, although during the grammaticalization of transitivity these verbs shifted towards the Intransitive Construction type. According to him, the transitive usage was possible till the 19th century, (e.g., *переправиться реку* 'to go across the river'). In contemporary Russian, a prepositional phrase is obligatory in the previous example, *переправиться через реку* 'to go across the river.' (cf. Крысько, 2006:348-365). The same observation is made by Zarickij (Зарицкий, 1961:64-65 and references therein).

The phenomenon in question is not limited to Russian. Certain Latin deponent verbs marked with *-r* can appear in the Transitive Construction, for example *degrassor* 'descend upon,' *furor* 'steal,' and *deveneror* 'exorcise' (Xu, Zheng, Aranoff & Anshen, 2007:134-135). Similarly, certain verbs in Ancient Greek marked with the middle marker *-μαι* appear in the Transitive Construction, for instance ἀποκρίνομαι 'I answer,' *μάχομαι* 'I fight,' and *ἐργάζομαι* 'I work' (Lavidas & Papangeli, 2007:100-101). Thus, the intransitivity as an invariant meaning is highly problematic in three respects: diachronically, synchronically and, typologically in related languages.

From a typological perspective, perhaps the most influential proposal for the invariant function of the Reflexive Marker is established by Kemmer (1993) figuring prominently in recent studies (cf. Bostoen & Nzang-Bie, 2010; Moyse-Faurie, 2008). According to Kemmer's proposal, the Reflexive Marker is used to signal the low-elaboration of the event constituting a structure occupying the position between the transitive and intransitive event types. In the most prominent case, the transitive event depicts the most elaborated event type as the Agent acts upon the Patient, yielding a structure in which the two entities are fully individualized. In the case of the intransitive event type, a single, individualized entity lies in the focus. Thus, a structure occupying the in-between status of these types is labeled as the middle.

In this vein, Kemmer's account is also a semantic definition of voice (cf. Manney, 2000). In a sense, this position is present in Gerritsen's proposal as she considers that the Russian Reflexive Marker is used to signal that the subject is both the starting and terminal point in the event. Thus, the Russian Reflexive Marker assigns an extra role to the subject of the verb (Gerritsen, 1990:5, 278). This proposal closely follows proposals made in Cognitive Linguistics (cf. Croft, 1991; Langacker, 1991) and contemporary approaches on lexical semantics, which rely on establishing the argument structure through the composition of the event structure (cf. Levin & Rappaport Hovav, 2005 for a general overview).

Kemmer's account makes it possible to motivate different usage patterns of the Reflexive Marker. At the same time, certain difficulties arise with the reflexiva tantum verbs, such as *касаться* 'touch.' Certainly, the verb is less transitive in comparison to *убить* 'kill,' for instance, as the Patient does not undergo a change of state. However, by contrasting Russian and English, one would simply have to state that the event structures are carved differently in these languages.

Additionally, the semantic extensions of the Reflexive Marker are sometimes connected to the concept of the Sphere of the Subject or Agent, but its theoretical basis has not been fully explicated.[11] The idea of the Sphere of the Agent is discussed, for instance, by Schenker (1986) and Bakker (1994:35), but rigid definitions are missing (cf. Geniušienė, 1987:14).

Another set of invariant meaning attributed to the Reflexive Marker is related to the event structure and relation between the subject and the object, namely the change in the valency structure of the verb. This perspective is also broader than the one based on the notion of intransitivity. Moreover, it can be used to cover every instance of the Russian Reflexive Marker with the exception of the reflexiva tantum verbs, a position already explicitly expressed in Geniušienė (1987), (e.g., the pair *бросать камни*$_{\mathrm{acc}}$ 'throw rocks' ~ *бросать камнями*$_{\mathrm{ins}}$ 'throw rocks').

Closely related to the change in the valency structure is the concept of the subject-orientation (отсубъектные) and the object-orientation (отобъектные) established by Janko-Trinickaya. This category is a crucial property of contemporary Russian diathesis tradition (Калашникова & Сай, 2006; Князев, 2007).[12] Furthermore, Janko-Trinickaya's observation can be viewed in relation to the development of the notion that predicate structure is a tripartite. It consists of the subject, the predicate and the object, on the one hand, and the position of establishing relationships in terms of binary categories, on the other. The label subject-orientation refers to a change in the argument structure where the argument occupying the subject position is the same as the reflexive and non-reflexive verb. In contrast, the object-orientation is a generalization referring to verbs which have an object of the non-reflexive verb occupying the subject position (Янко-Триницкая, 1962:79-80).

In this manner, the highly influential markedness theory of Jakobson can also be viewed as being assimilated by this concept because the orientation is

[11] The concept of the Sphere of the Agent could be interpreted in the terms of Cognitive Grammar where the concept of dominion is crucial in establishing reference point constructions (Langacker, 1993). As the Reflexive Marker is part of the family of devices in forming anaphoric relations, the sphere and dominion may be perceived as equal labels for the same phenomenon. However, setting up this line of descriptive devices is beyond the scope of this study.

[12] The label orientation is used to avoid any confusion with the unrelated term of subjectivity which figures prominently in Cognitive Linguistics (Langacker, 1990; Verhagen, 2005).

inherently defined in terms of a binary relation, a relation between the marked and unmarked form. Although the orientation may appear as a strong generalization connected to the directionality of the derivation, in reality, it is not. The sole reason is that most reflexive verbs can combine with either type of the orientation depending on the construction type. For example *мыть* 'wash' ~ *мыться* 'wash' can be used to illustrate both orientations. The Semantic Reflexive, 'wash oneself,' constitutes the subject-orientation whereas the Passive, 'be washed,' pertains to the object-orientation.[13]

A change in the syntactic configuration is the primary descriptive device utilized in the diathesis tradition and even commonly in discussion related to valency structure in contemporary linguistics theories (Haspelmath & Müller-Bardey, 2005). A similar binary category, but semantically motivated distinction is proposed by Haiman. He concludes that, typically, the light marker is associated with introverted verbs and the heavy marker with extroverted verbs. The former refers to verbs depicting an activity which is typically acted upon oneself. The latter refers to the opposite relation (Haiman, John, 1983).

The different positions on the possible invariant meaning of function of the Reflexive Marker establish the important facets of the phenomenon in question. The Russian Reflexive Marker cuts through the whole linguistics system transforming an already complex category to an even more complex one, as the descriptive practices are heavily intertwined with the overall view of the system. At the same, the advantage of a particular position is always subjected to the employed research question(s). For example, the reflexiva tantum verbs are commonly excluded from description without any discussion on the motivation behind the delimitation. Another excluded relation is the possibility to display multiple valance frames. Examples *1.2-1*–1.2-3 are specifically extracted from the Russian National Corpus for the purposes of illustration and are not intended as an exhaustive list of all the potential configurations available for the verb *мыться* 'wash.' All the examples are confined to intransitivity. Example *1.2-1* illustrates the Semantic Reflexive.

1.2-1	*Абрам*	*мо-ет-ся*	*в*	*бан-е*	[…].
	NAME.NOM	wash-3S.PRS-RM	PR	sauna-PREP	

Abram is washing in the sauna.
[RNC, Коллекция анекдотов: одесситы (1970–2000)]

The Passive is profiled in 1.2-2. whereas Example 1.2-3 can be considered as an extension of the Passive, occasionally referred to as the Medial or the Potential Passive.

1.2-2	*Посуд-а*	*мо-ет-ся*	*так*	[…].
	Dish-NOM	wash-3S.PRS-RM	like.that	

[13] As Knyazev (Князев, 2007:264-265) points out, the definition of orientation crucially hinges on the definitions of the subject and the object. Specifically, either they are related to semantics or syntax in Russian tradition.

The dishes are washed like that.
[RNC, Кузнецов Алексей Анатольевич. Между Гринвичем и Куреневкой (2002)]

1.2-3 *Во-вторых, ткан-ь легк-о мо-ет-ся* […].
Secondly, fabric-NOM easy-ADV wash-3S.PRS-RM
Secondly, the fabric is easy to wash.
[RNC, Николай Качурин. Mitsubishi Pajero 3.2 DI-D: 14700 км //"Автопилот," 2002.08.15]

Importantly, the subject-orientation (1.2-1) and object-orientation (1.2-2 and 1.2-3) are demonstrated with these examples. If a derivational explanation is sought, the base form would be the Transitive Construction profiled with the verb *мыть* 'wash.' The range of the different configurations is, however, typically disregarded. Exceptions to this are Gerritsen (1990) and Geniušienė (1987), who briefly discuss this possibility. Although the potentiality to display multiple patterns is self-evident, descriptive devices established through a binary relation cannot be used to offer a motivated account without multiplying the actual descriptive devices employed in a given model because the marked forms themselves are not assumed to be related.

The various instantiations do not present a random set or list of types. Instead, these are assumed to be systematically related. The possibility that some of the types are more central in terms of type frequency, for example, however, is almost absent in the literature as most studies are not based on a sample but on a collection of verbs. Similarly, general frequency information about the various types is also sparse. The study by Kalashnikova and Saj (Калашникова & Сай, 2006) includes type frequencies based on the diathesis tradition. Other exceptions would be Engdahl's (2006) study on Scandinavian *-s*, which is diachronically related to the Russian Reflexive Marker, and Kolomackij's (2009) study on passive construction types in Russian.

In sum, the synopsis offered in Sections 1.2.1 and 1.2.2 illustrates the main lines of research already established in the linguistic tradition in investigating the (Russian) Reflexive Marker. Although the Reflexive Marker has been studied intensively, the research questions set for the present study are filling important gaps in the tradition or hitherto, unexplored lines. Any account utilizing invariant meaning must offer an exact model at least on three questions, especially if pairs are posited: 1) How are reflexiva tantum verbs modeled?, 2) How are multiple argument constructions modeled?, and 3) How is the invariant meaning mapped to the exact argument construction of a particular verb?

1.3 Aspects of Cognitive Linguistics

Cognitive Linguistics has become in recent years, a well-established paradigm in linguistics. This is illustrated by the number of introductory textbooks such as Croft and Cruse (2004), Evans and Green (2006), and the Oxford Handbook of Cognitive Linguistics (Geeraerts & Cuyckens, 2007). Cognitive Linguistics started as an alternative approach to the generative paradigm in the late seventies and the early eighties. The most prominent figures of early Cognitive Linguistics are Ronald W. Langacker, George Lakoff, and Leonard Talmy. These key figures contributed to the growth and the theoretical basis of contemporary Cognitive Linguistics with their own research programs. Langacker (1987; 1991) formulated Cognitive Grammar, which is still the most elaborated version of grammatical concepts in Cognitive Linguistics. Lakoff and Johnson (1980) brought (conceptual) metaphor as part of the research program. They argue that the most of human experience is grounded in terms of spatial experience and on our understanding of the human body, which is the bodily basis of conceptual structure (Lakoff, 1987). The latter concept figures prominently in the embodiment research paradigm. An excellent survey on the concept is given by Rohrer (2007). Metaphoric mappings have also been a central role in the Russian linguistics tradition, for example by Apresjan (1974). Additionally, metaphoric mappings are used in the contemporary diathesis tradition by Paducheva (Падучева, 2004) and Knyazev (Князев, 2007).

Talmy introduced the basic principles related to spatial semantics, force dynamics, fictive motion, and figure/ground relations. These central concepts are compiled in two volumes (Talmy, 2000a; b). These concepts are grounded in Gestalt Psychology (Koffka, 1935). The cognitive basis of figure and ground rests on visual perception; one element may be considered to be the most prominent in a visual field yielding the figure (the focus of the attention) while other elements are backgrounded, (i.e., the ground). The process of categorization is perhaps best viewed as a human condition, as tests on illusory pictures, such as the Kanizsa triangle, show that people impose structure to scenes (cf. Spivey, 2007:210-211, 232-236).

Indeed, the role of the spatial semantics, and more broadly spatial experience, can be considered the uniting factor in early Cognitive Linguistics. For example, Cognitive Grammar was originally labeled as Space Grammar. The use of diagrams to capture generalizations may be viewed as a spatial and visual metaphor. Moreover, Lakoff and Johnson (1980) (developed further in Johnson (1987) introduced the concept of image-schema which is postulated to be grounded in spatial experience, such as container, path, and contact. Additionally, early studies were concentrated on the theoretical description of spatial semantics, such prepositions (cf. Zlatev, 2007).

At the time, the generative paradigm was mostly concerned with abstract and formal rules and the lexicon occupied a peripheral position. In this respect, the renewed interest in lexical semantics may be stated as being another dividing line between the two approaches. Similarly, the polysemy of linguistic items can also be interpreted as moving away from structuralism, as meaning was not viewed primarily in relation to contrast or opposition within a system. In contrast, distributional properties, albeit not necessarily binary, have re-entered into Cognitive Linguistics, especially in Radical Construction Grammar (Croft, 2001) and, corpus-based studies; which by necessity rely on distributional properties (Stefanowitsch & Gries, 2003). Indeed, there are a small but growing number of studies which utilize corpus data within Cognitive Linguistics. An extensive survey of these studies is given in Glynn (2010).

In addition to the primacy of semantics, the role of categorization has a crucial role in Cognitive Linguistics formulated by Rosch and others as prototype theory in the seventies. In short, categorization is defined as perceived similarity between objects in Cognitive Linguistics, incorporating prototype structure where category membership is graded and members have unequal status, where some members may be more central in comparison with others (Rosch, 1978; Rosch & Mervis, 1975). Similarity and category membership interact as meaning is not viewed as a separate object of study between semantics and pragmatics (Nosofsky, 1986; Spivey, 2007; Tversky & Gati, 1978). Instead, they are assumed to be based on encyclopedic knowledge of the world, forming a continuum in contrast to a binary category (Bybee, 2010; Goldberg, 2006; Langacker, 1988c).

It follows that language is not viewed as being governed by separate cognitive principles, as it is postulated in the generative paradigm (Chomsky, 1965; 1981; 2002). Instead, Cognitive Linguistics assumes that the same domain-general principles are applicable to the formation of linguistic categories analogously to other perceptual categories (cf. Spivey, 2007 Chapter 6) constituting the basic and central cognitive aspect of the theory (Bybee, 2010). A third assumption made in Cognitive Linguistics is the perspectival nature of linguistic meaning, related to categorization. Meaning is not just an objective reflection of the world, but it is assumed to reflect both the experience of individuals and cultures/communities (Geeraerts & Cuyckens, 2007:5).[14] Although the perspectival nature of linguistic meaning is grounded in categorization, the application of the concept also pertains to linguistic description. For example, a situation can be viewed from different perspectives yielding different linguistic configurations, such as the use of the Transitive or Passive Construction.

Another aspect of perspectivization is that a situation or an entity can be viewed at different levels of granularity, a difference in level of specificity or schematicity, for example, *thing* → *objet* → *vehicle* → *truck* → *pick-up-track* → *battered old pick-up-truck*. In principle, the set of expressions are applicable to the same entity depending on the circumstances. These configurations yield the descriptive tools used in Cognitive Grammar to label different configurations constituting the operations of construal. Every linguistic expression is assumed to be a symbol consisting of the semantic and phonological pole. Furthermore, every linguistic expression is connected to some concept referred to as the base and a usage-pattern highlights certain facets of this base defined as the profile (Langacker, 2009:3, 6-7).

In this vein, the terminology employed by Langacker is in close proximity with the figure and ground used by Talmy, highlighting different facets of

[14] Although Cognitive Linguistics is concerned with usage and generalizations over usage patterns, sociolinguistics factors have generally had a minor role (cf. Geeraerts, Kristiansen & Peirsman, 2010).

operations involved.[15] At the same time, the construal of a situation and broad categorization are fully reflected in cognitively oriented studies. The position is already elaborated in Tuggy (1993), and, construction grammars have begun to fully incorporate this concept. This perspective on meaning is also already present in frame semantics proposed by Fillmore (cf. 1982) and later adapted in a computational model, (i.e., FrameNet). Accordingly, meaning is not described in terms of truth conditions or necessary and sufficient criteria. Rather, it requires the understanding of the situation which amounts to a holistic view of meaning, (Lakoff 1987; Langacker, 1987).

The third tenet highlights another crucial distinction between the Cognitive and the generative paradigms. The latter introduced the dichotomy between competence and performance, mirroring in principle, the position in structuralism, and the dichotomy between language and parole. Accordingly, language is viewed as a social and collective system and parole as an individual activity (de Saussure, 1916). In contrast, the terms in the generative paradigm refer to the knowledge and use of the linguistic system, leaving social aspect outside linguistics proper, reinforcing the assumption of innateness and the autonomy of the linguistic system (Chomsky, 1957; 1965). The debate on innateness is ongoing between the Cognitive and generative paradigms (cf. Goldberg, 2004; Lidz & Gleitman, 2004; Lidz & Waxman, 2004; Tomasello, 2004) even though both have undergone considerable revisions through the course of time. In contrast, Cognitive Linguistics posits a continuum between this dichotomy portraying a language both in terms of categorization and variation. Although variation and sociolinguistic factors are embedded within the Cognitive Linguistics, they figure less prominently. Nonetheless, sociolinguistic factors are present in the studies of Geeraerts et al. (cf. 1994) and his research group (cf. Grondelaers, Speelman & Geeraerts, 2007 for an overview).

In short, the previously outlined introduction to the formation of the Cognitive Linguistics paradigm, and the basic assumption made within the theory, can be grouped under three basic tenets: 1) language is not an autonomous cognitive faculty, 2) grammar reflects conceptualization, and 3) knowledge of language emerges from language use (Bybee, 2010:8-9; Croft & Cruse, 2004:1-2). These assumptions lead to a position where the syntax and the lexicon form a continuum rather than a dichotomy. Consequently, the taxonomy of the Russian Reflexive Marker is not presented as consisting of sharp divisions between form-based and verb-based categories.

1.4 Methodological Considerations and Data

The use of corpora in studying linguistic structure has become increasingly more popular in contemporary linguistics and even in theoretical linguistics. The

[15] The terminology employed in Cognitive Grammar is fairly substantial. It is crucial, however, to distinguish different levels of organization in the grammatical architecture of Cognitive Grammar.

contributing factor for this is close to technical development and to the ease of access to large-scale electronic data. The trend is further visible in the rapid growth of national corpus projects that have been recently undertaken. The increase in use of corpora is also present in recent developments within Cognitive Linguistics, especially in the usage-based approaches (Bybee, 2010; Langacker, 1988b among many others). Underpinnings of this kind, or assumptions, form research questions to the point that methodology and theory are not separated, but constitute a uniform nexus (Geeraerts et al., 1994; Glynn, 2010).

The relevance of this perspective cannot be underestimated as there are critical methodological questions related to forming abstractions, specifically argument constructions. Thus, it is not a matter of offering a definition, but also how the definition can be operationalized. Consequently, a definition becomes connected to the data and the theory. Then again, this has always been the case in linguistics. Even the minimal pair method builds on theoretical assumptions, namely that linguistic data can be analyzed as a pair constituting the relevant fact about the phenomenon in question. Thus, this chapter is devoted to questions related to the data and to the use of statistical methods.

The role of corpora and the status of the data has become an increasingly debated topic in theoretical linguistics across different schools and within schools, as is the case in Cognitive Linguistics. Traditionally, evidence used in linguistics is based on native speaker intuitions. This position is closely related to the late 20th century development in linguistics and the rise of the generative paradigm. The emphasis of data in linguistics shifted towards the view of language and linguistics as a mentalist program leading to the monism of methodological possibilities (cf. Chomsky, 1957; 1965). Traditionally, the label is perhaps more related to the rise of the generative paradigm in the history of linguistics in general (cf. Geeraerts, 2010). For example, the early Russian tradition is mostly based on observed data, albeit with a bias towards literature. Nonetheless, this shift can also be seen as a counteraction to American structuralism, which considered the primary evidence to be observed patterns and induction, the most prominent figure being Bloomfield (1939). The question of the status of introspection is also present in Cognitive Linguistics, for example, Talmy (2000a:5-6) explicitly states that introspection is the primary tool to access semantic information. In contrast, Talmy (2006:xx-xxi) considers that introspection, corpus, and experimental data can offer a better understanding of linguistic structure when they are combined, as every method has strengths and weaknesses.

The debate is centered on the question about what counts as evidence in linguistics; such questions regard adequacy, reproducibility, and reliability (cf. Kepser & Reis, 2005; Penke & Rosenbach, 2004). On the other hand, the question is also related to what aspect of linguistic data is considered. The answer to this question is typically divided in terms of two camps in linguistics, namely the formal/generative and the functional. Depending on the perspective, Cognitive Linguistics can be viewed as part of the functional (sometimes

Cognitive Linguistics subsumes the functional), although this need not be the case as functional explanation does not necessarily equate with cognitive plausibility (cf. Haspelmath, 2004).

The answer to the type of evidence crucially hinges on which camp is answering the question. The formal approach emphasizes the innateness of linguistic structure, whereas functional and Cognitive Linguistics is primarily concerned with the interplay of language structure and use (cf. Bybee, 1985; Goldberg, 2006; Langacker, 1991). The division between these two camps is clearly illustrated in Newmeyer (2003). He takes a strict stance on the issue by stating that usage is usage and grammar is grammar. The debate is also present in the Russian tradition illustrated by Shapir (Шапир, 2005), Gladkij (Гладкий, 2007), and Kopotev and Mustajoki (Копотев & Мустайоки, 2008). At the same time, linguistic items do not exist in isolation, but in a highly intertwined mixture of contextual properties. The famous statement by Firth (1957) illustrates the point: "You shall know a word by the company it keeps." This appears to be a most fitting statement to the Russian Reflexive Marker, as the context crucially guides the exact semantics of the reflexive verbs.

The increase in availability of large-scale corpora has substantially influenced the view on language, and allowed to support assumptions made in usage-based approaches, especially on lexical and grammatical structures. Bybee (2010:15) states that the belief about limited memory resource was one of the motivations to postulate abstract representations over storage (cf. Dąbrowska, 2004:26-27). Recent findings have substantiated the relevance of the claim that linguistic categories include item-specific properties (Goldberg, 2006; Hay, Nolan & Drager, 2006; Nosofsky, 1988). Nonetheless, the use of corpus data is not without its issues. Questions regarding representativeness and balance are of primary concern when using corpus data (Arppe, 2008:2-3, 67-68; cf. Arppe, Gilquin, Glynn, Hilpert & Zeschel, 2010 for recent discussion).[16]

At the same time, the plurality of data sources has actually increased the awareness on strengths and weaknesses of research designs. Gries (2002:27-28) has suggested a general research strategy by combining individual methods to overcome this issue, as such, supplementing acceptability or grammaticality judgment with corpus data. Thus, there is an increase towards pluralism rather than monism. The implementation of multiple sources of data provides more compelling arguments for or against certain modes of analysis, which is the combination of intuition with corpus and experimental data (cf. Arppe et al., 2010; Arppe & Järvikivi, 2007; Bresnan et al., 2007; Gries, 2010b; Hay, 2001). Although cognitive corpus linguistics is a small branch within Cognitive

[16] Another fundamental issue is related to negative evidence. The absence of a certain type of linguistics phenomenon does not entail that the type in question is ungrammatical. A corpus-based approach to negative evidence is discussed in Stefanowitsch (2008). As the central goal of this study is to establish the typical construction types of the Russian Reflexive Marker, the issue on negative evidence has a marginal role.

Linguistics, it has been gaining popularity in recent years.[17] Instead of portraying language from a monolithic perspective, multiple sources of evidence are increasingly utilized in the studies by Bresnan et al. (2007) and Bresnan and Ford (2010) on the Ditransitive Construction in English, Baayen and Moscoso del Prado Martín (2005) on the structure of the regular and irregular verbs in English, Dutch, and German, and Divjak and Gries (2006; 2008) on verbs denoting 'try' in Russian.

In addition to building corpora, other tools have become more accessible, such as WordNet (Miller, George A., Beckwith, Fellbaum, Gross & Miller, 1990) and FrameNet (Baker, Fillmore & Lowe, 1998). The former is a lexical database containing groups of words based on their lexical relationship, such as, homonymy and antonymy. The latter is another lexical database incorporating semantic information closely mirroring the constructionist perspective. Both of these databases can be considered to be basic tools in contemporary lexical semantics. The increase in availability of electronic resources is apparent; the development is certainly not evenly distributed among languages. Although Russian is the largest Slavic language, it currently lacks both of these resources. In comparison, WordNet is available for such languages as Czech, Estonian, and Finnish.[18]

Although this study is limited to corpus and dictionary-based data, a variable-based analysis increases the complexity of the data, especially when the variables are of mixed type covering both categorical (such as the referent type), and numeric (such as frequency, and variables). The implementation of statistical methods becomes important as it is difficult to establish generalizations when the data at hand consist of a multivariate feature space.

Multivariate methods are required to explore the potential structure present in the data. Furthermore, statistical methods, which can handle mixed types of variables, are required otherwise numeric variables would have to be discretized. For example, frequency might be modeled as low, middle, and high frequency. Discretizing variables causes them to become less sensitive and exploring the possible nonlinearities of the variables are lost (Baayen, 2010b). For the purposes of the present study, the utilized statistical methods serve to fulfill three purposes: 1) exploring the structure in the data, 2) estimating the importance of the variables, 3) evaluating the constructionist model in terms of predictive accuracy, and 4) forming generalizations across constructions.

Typically, statistical methods are divided into two groups: confirmatory and explanatory. Machine learning algorithms combine both aspects of the already

[17] Although converging evidence is not without issues (Arppe et al., 2010), corpus and experimental data do not necessarily target the same processes. A similar position is taken in Bybee, (i.e., for a usage-based approach all sources of evidence counts as evidence). However, data from different sources are influenced by different factors which need to be kept in mind (Bybee, 2010).

[18] A more comprehensive list is given in WordNet of the World (Global WordNet Association, www.globalwordnet.org).

fuzzy division between them. With the confirmatory method, the properties of the model are determined before entering the data, while with the explanatory method the properties of the model are estimated from the data. Classical regression exemplifies the former type of model and clustering, typically used to find structure in the data, illustrates the latter (Hastie, Tibshirani & Friedman, 2009). Most machine learning algorithms blur the distinctions as the models are estimated from the data. This has caused a discussion about the legitimacy of machine learning algorithms, as the functional form of the model cannot be analytically analyzed (Hand, 2006). Friedman (2006:15) illustrates the point by stating that when promising new methods are introduced, a degree of enthusiasm will follow and new gurus are established. At the same time, Friedman (2006:17) argues that certain new methods, such as classification and regression trees and their advancement to random forests, offer genuine improvement in tasks related to classification for practitioners, especially in research fields where the data types do not typically meet the assumptions of the classical regression model.

Strobl et al. take a practical stance on the issue. If classical regression reaches similar accuracy as random forests, the complexity involved in the latter method may be considered unwarranted. In the opposite case, the simpler model established with classical regression may be considered insufficient, missing the complexities in the data (Strobl, Malley & Tutz, 2009b:28). Most linguistic analyses can be divided into three parts: the classification of the data, the assessment of the importance of the proposed variables, and the formation of predictions. All three can be established with random forests.

Section 1.4.1 describes the sampling frame used to obtain the data. The first multivariate method, classification, and regression trees (CART) are described in Section 1.4.2. The principles of CART are important for understanding the ensemble method of random forests, (i.e., a collection of trees instead of operating with a single tree solution). The principles of the random forests algorithm are discussed in Section 4.3. All statistical models are implemented in open source environment for statistical computing R (R Development Core Team, 2011). The majority of the visualizations are done with ggplot2 package in R (Wickham, 2009).

1.4.1 Data and the Sampling Frame

This Section discusses the sampling frame that was used to form the basis of the database. It is important to emphasize that prior to this study, only one previous taxonomic study by Kalashnikova and Saj (Калашникова & Сай, 2006) has been conducted. It was based on corpus data in a sense that data were sampled, at least to my knowledge. Consequently, there is a marginal amount of data available to make direct comparisons and to form guidelines how to obtain a representative sample of the Russian Reflexive Marker. Nonetheless, two random sampling methods were considered: either a repeated measure over verbs or a repeated measure over the Reflexive Marker. Both methods have their advantages and disadvantages depending on the research questions. The basic question for this study is to establish a template for the constructionist model. The value of the two sampling frames is considered against this background.

The first method would require a preselection of the reflexive verbs, influencing the whole structure of the data as a specific verb will display a specific range of construction types in a sample although we do not know the exact range a priori. Moreover, the verbs should be considered to be typical instantiations of the Reflexive Marker, as Cognitive Linguistics and Construction Grammar emphasize the role of typicality. Thus, certain verbs are considered to be more representative of their category. Such information cannot be extracted based on the previous studies. Although a potentially representative inventory of types might be compiled based on the previous studies, the verbs selected to instantiate the types do not necessarily represent the best exemplars of the said category.

There is yet another twist related to the selection of the verbs. The diathesis tradition and structuralism in general do not considering typicality effects, (i.e., instantiations do not display membership in terms of degree, the exclusion of multiple argument structures available for a certain verb introduces another layer of tacit assumptions). Such text book verbs as *строиться* 'build' have a complex structure, for example, *дом строится* ‘the house is being built,’ *мы строимся* ‘we are building (a house for us),’ and *теория строится на чем-*л, ‘the theory is based on something.’[19]

Importantly, none of the instantiations can be considered simply displaying some pragmatic extension that could be used to motivate the exclusion of the multiple instances, as the semantic shifts in the examples are also combined with differences in form. When multiple patterns are excluded, one of these instantiations is implicitly posited as being the basic. This is simply not a descriptive, but also a theoretical assumption which is left unspecified how one of these senses is selected and others excluded, especially if categories do not have degree membership and the various instantiations are described as lexicalized patterns (Israeli, 1997; Князев, 2007; Янко-Триницкая, 1962).

At the same time, the verb-based perspective would offer a “clean” data set because the variability would be considerably reduced. Additionally, a sufficiently large sample for the verbs could be encoded with the variables simplifying the process of utilizing statistical methods (cf. Arppe, 2008; Bresnan et al., 2007; Gries, 2006). However, the assumptions made in preselecting a set of reflexive verbs are numerous. The most important violation would concern random sampling. Generalizations over a small set of preselected items should be stated as such and not as a property of the underlying population (cf. Baayen, 2010b; Chan & Vitevitch, 2009). In my view, taxonomy is a prerequisite for building a fine-grained analysis based on verbs. Thus, the procedure of forming

[19] The last example aligns with the copula constructions in Russian, (i.e., *базироваться* 'be based on' displaying the interconnectedness of the different types). We will return to these instantiations later as they exemplify the theoretical basis of this study.

a sampling frame as a repeated measure over the Reflexive Marker comes with minimal tacit assumptions compared to the alternative; the crucial thing being that all the verbs, which are combinable with the Reflexive Marker, belong to the same category. This assumption rests on the established practice already present in the early Russian Tradition (Янко-Триницкая, 1962).

The Russian National Corpus was chosen as it aims to be a representative sample of Russian covering a wide range of genres/text types over a long period of time (available at http://www.ruscorpora.ru/index.html). At the time of forming the database, the Russian National Corpus contained proximately 140 million words.[20] Additionally, the corpus has a syntactic search function enabling to form a sampling frame, which includes only the reflexive verbs either the *-ся* or the *-сь* form. Additionally, gerundive and participle forms were excluded from the sampling frame. From a constructionist perspective, both of these forms can be considered as construction types in their own right. These components establish the first part of the sampling frame. The second is related to the representativeness of the data, namely, the inclusion or exclusion of specific genres. Generally, there is a tacit assumption in the Russian linguistic tradition that literature is considered to occupy the privileged position compared to other genres. This is implicitly also present in the study by Kalashnikova and Saj (Калашникова & Сай, 2006) as their data are solely based on literature.[21]

Another aspect related to stratifying the sampling frame based on genre is related to the number of instances to be included. As the encoding of the data involves manual tagging, large samples become infeasible, limiting by proxy the number of genres that can be considered. Nonetheless, Biber (1993) has shown that even fairly small samples, proximately 1000 instances, can be representative if typicality effects are sought. For the purposes of the present study, four genres following the principles used in RussNet for written genres were included in the sampling frame: 1) news paper/media, 2) scientific/academic, 3) literature/fiction, and 4) spoken. The selected genres can be interpreted as representing central aspect of language use. Media is a highly varied mode compared to academic compensating each other in this sense. As literature has always occupied a central role in the Russian linguistic tradition, its inclusion is warranted (Азарова, Синопальникова & Яворская, 2004). Finally, the spoken language is self-evidently the primary mode of communication, making its status important for the purposes of the present study (Biber, 1993:181; Chafe, 1992:88-89).

The third and final component of the sampling frame is a limitation based

[20] Russian National Corpus is still being developed. Currently, the corpus contains a fairly large number of subcorpora which were not available at the time of forming the database for this study. Additionally, the word count has been increased substantially to 364 million (20.05.2012). However, exact figures of the composition of the Russian National Corpus cannot be given, for instance word counts, as the summary information was out of date at the time of forming the database.

[21] Goldberg (1995) considers that stylistic aspects are relevant for constructions.

on the time period. Only recent text types were included, ranging between 1995 and 2007, ensuring that the sampling frame reflects contemporary Russian. The news paper/media as a genre has drastically changed during the social changes in Russia, which has included Soviet, Perestroika and finally the contemporary aeras. At the same time, we cannot a priori assume that the recent social changes are reflected in the grammatical construction types of the Russian Reflexive Marker, although changes in the lexicon and style are well-documented (cf. Введенская, Павлова & Кашаева, 2005; Крысин, 2007b). By imposing restrictions on the time period, the possible variation is, nonetheless, reduced.

For each stratum delimited by four genres and the time period, 500 random instances were included in order to keep the proportions balanced across genres amounting to a total of 2000 instances available in the database. If the previous studies are implicitly biased towards literature, the sampling frame used in this study is biased towards written language. Additionally, the data are also biased towards production rather than processing, a general property of corpus data (cf. Arppe, 2008:3).

The possibility to establish generalization hinges on the structure of the sample. The database contains 819 unique reflexive verbs, which is a fairly substantial number of instances considering the sample size (*N* = 2000), and approximately 40% of the data points are covered by them. Another aspect related to the structure of the data points is the frequency of frequency, that is, the frequency distribution of the reflexive verbs based on the sample, given in Figure 1.4-1.

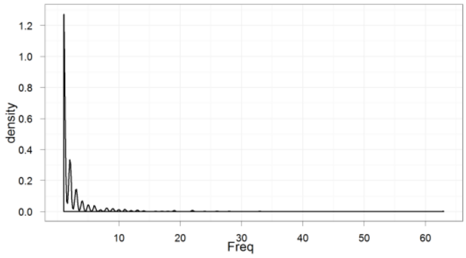


Figure 1.4-1 Density plot of the frequency of the frequency based on the reflexive verbs in the sample.

The distribution of the frequency of the frequency is extremely skewed towards hapax legomena, (i.e., verbs only appearing once, *n* = 509). It is commonly used as a measure of productivity (cf. Plag & Baayen, 2010:126 and references therein). On the one hand, the Reflexive Marker is a productive category; hence, the high number of hapax legomena is fully expected in a small sample. On the other, the high number also indicates that the sample is not biased towards a few frequent verbs offering wider coverage of potential reflexive verbs in Russian, considering the overall sample size. The extreme end is supported with such verbs as *являться* ‘be,’ *оказаться* ‘seem, appear,’ *заниматься* ‘engage,’ and *остаться* ‘stay, remain.’ Thus, the sample appears to display good properties for investigating typicality effect in terms of the structural properties of the verbs.

Nonetheless, the structure of the data imposes certain methodological issues. First, the random sampling introduced data sparseness; the sample size is fairly small, but also extremely wide considering the total number of the reflexive verbs. Effectively, generalizations or typicality effects concerning individual lexical items on different argument constructions cannot be made. Second, the analysis of the data in relation to argument constructions shifts the data towards a repeated measure over verbs. Due to the highly skewed distribution of the reflexive verbs, standard univariate statistical methods, like, χ^2, cannot be used because the aggregated data would be influenced by the frequent verbs attested in the sample (cf. Agresti, 2002:22-23, 78-79, 84-85). For instance, if we were interested in investigating whether two construction types differ in terms of the encoding of the referent type.

For purposes of modeling, the application of a univariate method is typically variable selection (cf. Arppe, 2008:118-119, 148-149). However, Díaz-Uriarte and Alvarez de Andrés (2006) have shown that random forests offer a more reliable estimation of variable importance compared to univariate methods. More importantly, linguistic categories tend to display a high degree of intermediacy to the point where multiple variables influence the structure (Baayen, Feldman & Schreuder, 2006; Bresnan et al., 2007). The results may simply be confounded by some hidden variables. Thus, the results would have to be checked with a multivariate model (Plag & Baayen, 2010:128-129). Yet, a third source of bias might be present in the data. Although this cannot be explored in this study, it is worth mentioning, at least in my view. The composition of texts available in the Russian National Corpus is aimed to cover Russian in its totality. When a fairly narrow restriction on time period is imposed, the number of available texts is effectively reduced. In return, this reduction may lead to artificial over- or under-representation of certain verbs.

The random sample was analyzed into 19 argument constructions, the pairings of form and meaning/function. The argument constructions are discussed through Chapters 5–9 in relation to previously proposed taxonomies. Additionally, a small set of instantiations, *n* = 11, were excluded from the analysis. These were labeled as Other. These instances have considerable deviant patterns in terms of subject marking, for example *сообщаться* ‘communicate’ in 1.4-1.

1.4-1	*В*	*доклад-е*	*сообща-ет-ся*	*о*	*т-ом,*	*что* […].
	PR	report-PRE	inform-3S.PRS-RM	PR	that.PRE	that

It is stated in the report that.
[903, RNC, Радиоэхо // "Поиск," 2003.09.12]

Another pattern is given in 1.4-2 with the reflexive verb *насчитываться* ‘count’.

1.4-2 *Иногда* *насчитыва-ет-ся* *до* *пят-и*
Sometimes count-3S-PRS-RM PR five-GEN
тон-ов *в* *окраск-е.*
hue-GEN.PL PR color-PRE
Sometimes it is counted up to five hues in color.
[163, RNC, Туканы // "Мурзилка," №2," 1999]

A prepositional subject is sometimes used to refer to the impersonal type in 1.4-1 (Gerritsen, 1990) whereas Example 1.4-2 might be considered to have a genitive subject. In Meaning–Text theory, these prepositional genitive instances are considered to occupy the surface subject position called approximative (Мельчук, 1995:237-238). Because subjecthood is not the primary topic of this study, these instances are left for future studies.

In addition, the concept of Neighbor Verb is introduced to model cross-paradigmatic relations in Sections 3.1.1–3.1.4. Based on this, the database contains 717 neighbor verbs and 819 reflexive verbs. Thus, the proposed usage-based model is based on the structural properties of these two types of verbs. The theoretical basis of the model is the topic of Chapters 2 and 3.

1.4.2 Multivariate Methods: Classification and Regression Trees

Classification and Regression Trees (CART) have gained popularity in recent years after being first introduced by Breiman et al. (1984). An excellent introduction to these methods is offered in Strobl et al. (2009b). A technical introduction to CART is given in Breiman (1984), Hastie et al. (2009), and Berk (2008). As Strobl et al. (2009b) point out, CART offers a straightforward interpretation of the results, albeit nontrivial with excellent visualization options for small data sets. With a large number of predictors, the tree solution can become highly pronounced. Another attractive property of CART in relation to corpus data is that they are nonparametric methods and the tree solution is built based on the distribution of the data. They require minimal pre-processing of the data and can handle mixed types of variables. CART is insensitive to outliers, extreme values in the data, compared to the classical regression where even few extreme values can drastically affect the model.

Another property is their flexibility. CART can be used for both classification and regression problems. In classification, the response variable is categorical, for example construction type. The types do not have any inherent order. In regression, the response variable is continuous. The exact type depends on how the variable is measured (Agresti, 2002:2-3).[22] The last attractive property of CART algorithm used in this study is that it can be extended to polytomous classification problem, (i.e., to model a response

[22] Categorical variables can also have ordered scale, such as social class. Even construction types could be perceived as ordered depending on the data available. Two obvious choices would be the age of acquisition and diachronic development. The standard scale types (nominal, ordinal, interval, and ratio) are guidelines.

variable consisting of more than two categorical levels). Polytomous models are rarely used in linguistics although most categories do not come in binary form.

A large number of different algorithms are available and even most of the commercial statistical software packages include several ones. Only the family of recursive binary methods is considered in this study. In this study, a specific algorithm called conditional inference trees is used available in the party package in R (Hothorn, Hornik & Zeileis, 2006b). The 0.9-99992 version is used. A general property of CART models is to divide a given data set into binary subsets with respect to the levels of the response variable given the input. For example, a feature space consisting of five predictors would be divided into increasingly homogenous groups depending on the outcome of the response variable (Hastie et al., 2009:305-306; Strobl et al., 2009b:5-6).

Binary splits are preferred because they decrease the number of required data points. This property of CART will be demonstrated with a subset of construction types and predictors available in the data later in this section. Numeric predictors are treated as ordered, (i.e., their ranks are used instead of values). There are m – 1 possible cut-off points based on their m distinct values that preserve the order. This also results in immunity to standard monotonic transformations of predictors, such as logarithmic transformation of frequency, cf. Section 3.1.7. Ordered categorical predictors also have k – 1 cut-off points similar to continuous ones, (i.e., their k distinct ordered levels). In contrast, unordered categorical predictors or nominal scale have 2^{k-1} – 1 possible cut-off points. For example a predictor with 4 levels has 7 possible cut-off points. Consequently, categorical variables can become computationally demanding (Strobl, 2008:8-9).

In order to establish a split, a measure of node impurity is used. Perhaps the most widely used measure is referred to as the Gini index. The index is smallest when a pure node consists of a single type of the response variable and highest when the types have equal probability in the node (Breiman et al., 1984:104; Strobl et al., 2009b:8). The rationale behind the node impurity measure is that any subsequent split (daughter nodes) should be purer than the previous split (mother node). The Gini index is a nonnegative probability function defined as $\phi(p) = p(1-p)$ (Berk, 2005:7-8).

Nonetheless, some form of stopping criterion has to be implemented otherwise the algorithm would continue until only pure nodes would be present, consisting only of a single value of the response variable. A pure node exhaustively represents the partitions in the data, but such a model would be extremely poor as every data set contains some random fluctuation and an exhaustive model would simply adapt to the quirks of the specific data set, diminishing the capability to generate to unseen data. Thus, some form of pruning method has to be utilized to overcome the overfitting issue where the mode adapts to the data too closely. Typically, cross-validation is performed on the tree structure to reduce overfitting.[23] The conditional inference trees

[23] The basic principle behind cross-validation is to randomly divide the data into

implement a conditional inference procedure for the stop criterion, removing the requirement of utilizing pruning techniques. Another attractive property of CART is that partitioning of the data is established based on the distributional properties of data, such as inductive learning. The classical regression model is a linear function of the predictors. Thus, additivity and linearity are the most important assumptions of the classical regression model.

In linguistic settings, variables can often be correlated, leading to collinearity (Gelman & Hill, 2007:45-46). For example, subject argument and case marking are interrelated in Russian and generally in most languages with the case system. These violations have to be explored and specified in the classical regression model. A number of techniques are available if some of the assumptions are not met in classical regression, ranging from the monotonic transformations of the predictors, such as logarithm, to advanced methods, such as splines, that allow modeling nonlinearities in the data (cf. Agresti, 2002; Gelman & Hill, 2007; Harrell, 2001). In contrast, CART can model nonlinear functions from the input without explicitly specifying them. Additionally, the issue of collinearity is partly alleviated because predictors are modeled one at a time and not in combination. However, the response variable needs to be transformed if required, especially with regression trees, similar to linear regression (Strobl et al., 2009b:6).

The term CART is understood in a narrow sense from now on referring specifically to the conditional inference tree algorithm. The algorithm consists of the following three steps, First, test a global null hypothesis of independence between the predictors and the response variable. Stop if the null hypothesis cannot be rejected, (i.e., the response variable is not dependent of the predictor). If the null hypothesis can be rejected select the predictor with the strongest association to the response variable. Second, carry out a binary split in the selected predictor. Third, repeat steps 1 and 2 recursively. The recursive property of CART ensures that a predictor can appear multiple times in a given tree structure. Thus, interactions between predictors are established in a data-driven manner without specifically stipulating them in the model. The stop criterion, (default setting) is based on a univariate *p*-value[24] (Hothorn et al.,

subsets. A model is trained on one subset (training set) and validated on another set (testing set). Multiple divisions are used and the results are averaged across different runs. The exact procedure of sampling and validation depends on the form of the selected method of cross-validation (cf. Hastie et al., 2009 Section 7.10).

[24] In addition to quadratic term, the Bonferroni and Monte Carlo simulations are available to calculate the *p*-value with conditional inference trees. Bonferroni is one of the several methods available to assess *p*-value when multiple tests are performed on the same data. The Bonferroni corrected *p*-value is calculated by dividing the critical α-value, for example, 0.05, by the number of performed tests (cf. Arppe, 2008:84-85 for discussion on adjustments in general). Monte Carlo is another commonly used method to evaluate the uncertainty in estimations (Davidson & Hinkley, 1997:140-141). The procedure to obtain a *p*-value with Monte Carlo simulation is based on the following formula: r+1/n+1 where n is the number of the replicate samples and r is the

2006b:4-7). The default critical *p*-value is 0.05, an arbitrary, cut-off point convention commonly used in human sciences.

A single predictor is used to test the strength of association. In this vein, the conditional inference trees resemble step-wise regression where predictors are assessed one at the time. Additionally, this procedure overcomes a possible selection bias towards variables with multiple cut-off points. Another important parameter of the model concerns the test statistics. Quadratic forms (the shape of a parabola) are used by default and recommended if categorical variables are used in the model to fasten the computation time because quadratic forms follow χ^2 distribution (Hothorn et al., 2006b:4-7).

Hothorn et al. (2006b) have demonstrated that conditional inference trees meet the required conditions for a successful CART model: 1) the trees are unbiased, 2) the trees do not overfit, and 3) prediction accuracy is equivalent to pruned trees.[25] Strobl et al. (2007) have shown that the default settings of conditional inference tree produces results close to the Random Forests algorithm by Breiman (2001a), cf. Chapter 4. Several parameters influence the fitting and are adjusted with the ctree_control() function. The default parameters are not changed in this study as both individual trees and forests are implemented: 1) test statistics = quadratic, 2) test type = univariate, 3) minimum criterion = 0.95, 4) minimum split = 20, and 5) minimum bucket = 7. The first three parameters were already covered. The last two concern the weights of the nodes. The minimum split with the value 20 requires that the sum of weights is at least 20 in a node before splitting. The minimum bucket with the value 7 requires that the sum of weights is at least 7 in a terminal node (cf. Strobl, Hothorn & Zeileis, 2009a).

To illustrate the conditional inference trees in classification, the Passive (Pa) (n = 208) and Reciprocal (R) (n = 103) Constructions were selected and modeled as a function of the referent type of the subject (L1_Ref) and the frequency of the reflexive verbs on log scale (Ref_Freq_L), cf. Section 3.1.7 for discussion. The variable L1_Ref has four levels: Pe(rson), An(imate), In(animate) and Ab(stract), cf. Section 3.2.2 for discussion. The model formula used in R is given below.[26] The order of the predictors is irrelevant because the model is estimated from the data.

```
ctree(Construction ~ L1_Ref + L_Ref_Freq, data = dat_tree)
```

number of the replicate samples that produce a test statistic greater or equal to the one calculated from the observed data. Numerous variants are available to the above mentioned formula (North, Curtis & Sham, 2002).

[25] Although the different CART methods tested by Hothorn et al. (2006b) appear to be equivalent in terms of predictive accuracy, the actual tree solutions produced by them may vary structurally. The structural variability is influenced by the differences in the stop criterion.

[26] The ctree is the R function used to call the conditional inference tree algorithm. The Construction was modeled as a function (~) of the Referent Type of the Subject Slot and the log Frequency of the Reflexive Verb.

The party package offers extensive visualization options of the conditional inference trees. The classification tree is given in Figure 1.4-2 with two visualization options.

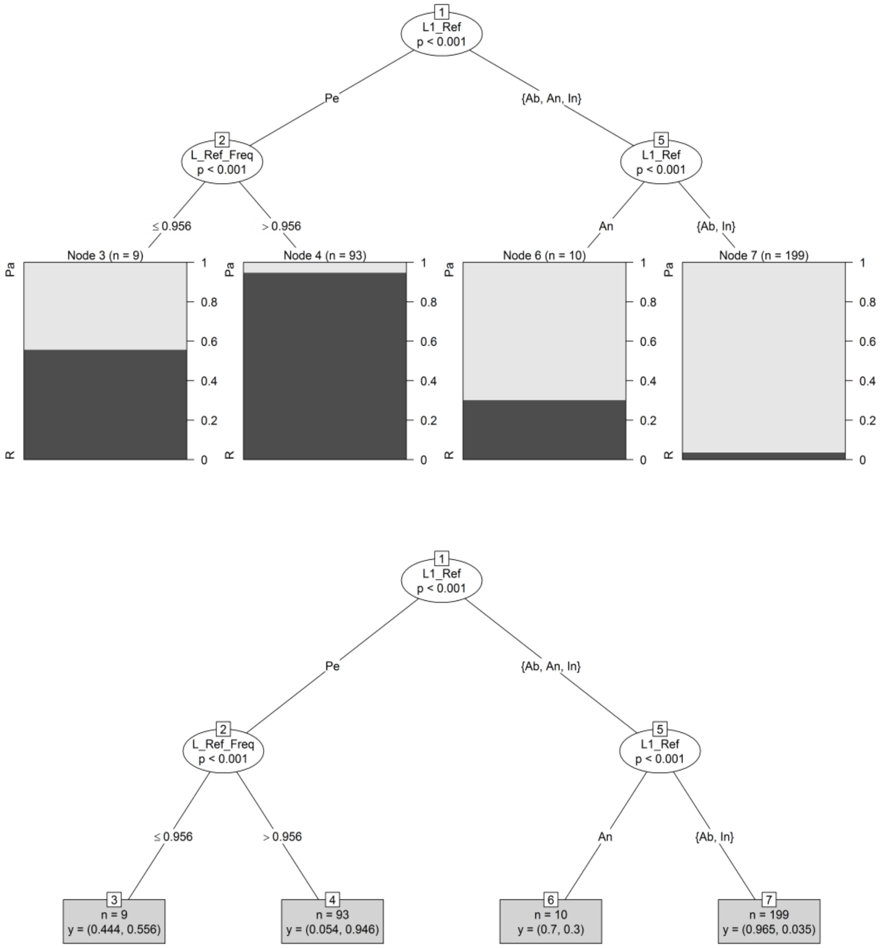


Figure 1.4-2 Conditional inference tree. The Pa(ssive) and R(eriprocal) Constructions were modeled as a function of the Referent Type of the Subject Slot (L1_Ref) and the log Frequency of the Reflexive Verb (L_Ref_Freq). The extended plot (top panel) and simple plot (bottom panel) types are shown.

The predictor variables form a feature space, which is recursively partitioned into sets with increasingly similar values of the response variables. In this illustration, the referent type of the subject partitions the data (node 1) into a binary split between the levels of the predictor. A division between Pe(rson) versus Ab(stract), An(imate), and In(animate) forms the two branches of the tree in node 1. In the right branch, the Pe(rson) is in interaction with the log frequency. A binary split is made splitting the predictor into two: log frequency greater than 0.956 and lesser than or equal to 0.956. After these partitions have been established, the null hypothesis cannot be rejected. Thus, the terminal nodes (3 and 4) display the relative frequencies of the response variable along with the number of data points.

In the right branch (node 5), a split is established between An(inmate) versus Ab(stract) and In(animate). The partitioning with level An(imate) cannot be partitioned any further and a terminal node is given (node 6) along with the relative frequencies of the response variable and the number of data points (*n* = 10). The levels Ab(stract) and In(animate) are partitioned in the node 4 and the terminal nodes are given in node 5 and 6. Additionally, the *p*-values of the partitions are given in the iner nodes. In terms of interpretation, several points need to be clarified.

Even in this small illustration, there is no main effect present in the estimated model based on the selected predictors. The Referent Type of the Subject interacts with the log Frequency of the Reflexive Verb in the right branch. Thus, the left branch cannot be interpreted as constituting a main effect. Considering that a data set always contains some random fluctuation, and the cut-off point can vary, it is unlikely that a CART model can represent main effects, the contribution of a single variable. Instead, the method is capable of displaying complex interactions in a data-driven manner (Strobl et al., 2009b:4, 11-12). In light of corpus data, this property reflects the nature of language as most phenomena are formed through interactions. Essentially, even the definition of a construction as a form-meaning pairing is established through interaction. Finally, the initial partition, node 1, has the strongest association with the response variable. After the initial split, the numbering is, however, arbitrary in the nodes continuing from left to right.

The data-driven procedure of CART is also the source of a number of limitations, which can even drastically affect the results. First, even a small change in the distribution can change the fitted tree, leading to a situation where comparing results in different studies may become difficult. Distributional properties are hardly constant across different data sets. Second, strong predictors can dominate the whole tree structure, potentially masking weaker but important predictors influencing the results of CART. Third, the selection of the initial split determines the whole shape of the tree. For example, a data set consisting of *n* = 100 points could be partitioned into two sets: *n* = 40 and *n* = 60. Once the data are partitioned, the final shape of the tree is influenced by the initial split. Fourth, the splitting procedure itself may be a limiting factor. A binary split creates a sharp decision boundary whereas most phenomena are best viewed as smooth transitions. These factors may lead to a weaker predictive accuracy (Hastie et al., 2009; Strobl et al., 2009b).

The concept of decision surface is crucial for understanding the more advanced methods based on tree-solutions. To demonstrate the sharp decision surface of tree-based methods, two variables *y* and *x1* were simulated, and adapted after Berk (2008). This exemplifies a regression tree, but it also extends to classification problems.[27]

The aim is to model the functional form. The parabolic shape is visualized in Figure 1.4-3 (top panel). To estimate the function, two models were fitted, a linear regression model with the lm() and conditional inference trees with ctree(), given below:

lm(*y* ~ *x1*, data=dat) (linear model)
ctree(*y* ~ *x1*, data=dat) (conditional inference tree)

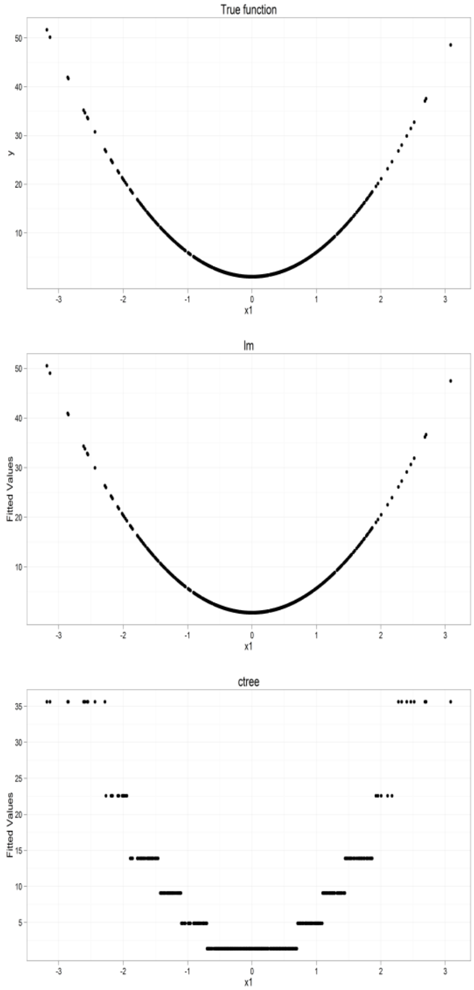


Figure 1.4-3 Scatterplots for estimating target function with two models: Target function (upper panel), fitted values from linear regression (middle panel), and fitted values from conditional inference tree (lower panel).

[27] Arppe (2008) gives an in-depth evaluation of regression models in linguistic setting.

The linear model offers a nearly perfect fit to the data, visible in Figure 1.4-3 (top right panel). The fitted values of the regression model are given on the y-axis. In contrast, the ctree is able to model the parabolic shape inductively from the data, but the surface is sharp because there are 11 binary nodes in the model, visible as clusters of dots in Figure 1.4-3 (bottom panel). The fitted values of the conditional inference trees are given on the y-axis. Because most phenomena tend to be smooth rather than sharp binary decisions, the tree-based model may have lower prediction accuracy to unseen data. However, the classification and regression trees can be extended with random forests to introduce smoothness. This topic is discussed in Chapter 4.

1.5 Organization of the Study

This study offers a probabilistic and constructionist model of the Russian Reflexive Marker. The label Construction Grammar covers a family of different theoretical accounts, albeit related ones. The perspective adapted in this study is a combination of Goldberg's Cognitive Construction Grammar and Croft's Radical Construction Grammar. A similar position is taken, for example, in studies of argument constructions in Icelandic by Barðdal (2008), and Divjak and Janda (2008). This position is discussed in Section 2.1.

The fundamental tenet of the proposed model is that various levels of abstraction are assumed to be readily available in language. Furthermore, it is argued that constructions, pairings of form and meaning, display similar properties. The relation between the verb and the abstraction is formulated in Chapter 2, specifically the interconnection between verb-specific and argument constructions in Section 2.2.3. In general, this chapter establishes the theoretical basis of the model.

Chapter 3 sets forth the proposed linguistic model, building on Bybee's (1985) concept of network structures. Usage-based models assume that the connections between items are dynamic. This property is operationalized as the degree and the strength of connectivity between items. The degree of connectivity is defined relative to a lexical network labeled as the Neighborhood in Sections 3.1.1–3.1.4. The strength of connectivity is discussed in Sections 3.1.5–3.1.9. Additionally, a quantitative perspective on Russian verbs is offered in Sections 3.1–3.1.10 based on the distributional properties that was obtained from the lexical network. Additionally, all the variables used in the multivariate analysis are described throughout this chapter. Finally, Section 3.1.12 gives the summary of the proposed model and anchors it to Construction Grammar.

The proposed set of the argument constructions is discussed in Chapters 5–9. For each argument construction, a connection to previous taxonomies is drawn. Results obtained from the statistical model are given after the theoretical description. Classification and regression trees introduced in Section 1.4.2 are extended to random forests in Chapter 4. This chapter makes a conceptual connection between the usage-based model and the random forests algorithm. Additionally, the fitting process of random forests is illustrated and evaluation measures are given in Section 4.5.

Chapter 10 goes beyond the argument constructions. First, an estimated global ranking of the predictors used in the random forests model is given. The final matter dealt with in this study is the formation of the network of the argument constructions. A data-driven approach is set forth by combining distances obtained from the random forests with clustering. Chapter 11 gives a general discussion on the results and outlines future research prospects in relation to diachronic, synchronic, and experimental studies. Lastly, conclusions are made in Chapter 12.

In sum, Construction Grammar is a simple theory, as minimal assumptions are made that would require independent evidence, such as movement in the generative paradigm (Chomsky, 1965). Additionally, the primary means of forming generalizations in Construction Grammar is through the network model and the slots of the constructions. Thus, this study is an attempt to bring two components together in order to form a coherent model of the Russian Reflexive Marker: 1) the interconnection of the reflexive verbs and the argument constructions and 2) generalizations obtained from the structure of the network. In short, this study is primarily theoretical and programmatic in order to formulate a usage-based model of the Russian Reflexive Marker.

2 Verbs and Constructions

This chapter introduces the basis of the theoretical devices used in this study in relation to Construction Grammar. First, Section 2.1 introduces Construction Grammar. Although the term Construction Grammar covers a number of different theoretical approaches, the shared commonality between them is discussed. In addition, the central developments within the family of Construction Grammar are outlined. Second, the descriptive devices employed in the diathesis tradition and in formal approaches to characterize the argument structure of verbs are discussed in Section 2.2. As this study moves away from the traditional pair account, this section serves as a background and allows us to situate the proposed usage-based model. Third, the concept of schematicity is formulated as it establishes the basis of the formation of different levels of abstractions in Section 2.2.2. Finally, the concept of construction is defined relative different levels of schematicity yielding the schematic relation between the Verb-Specific and the Argument Construction in Section 2.2.3.

2.1 Constructionist Approaches

Similar to the label Cognitive Linguistics, the term Construction Grammar consists of a family of theories and does not constitute a single unified theory of language. Indeed, the concept of prototype is perhaps most suitable to characterize the situation, as different theories under the label Construction Grammar may have different perspective on the structure and the architecture of the theory. Central differences are issues on storage, level of formalism, and cognitive plausibility. The question of storage is both related to descriptive devices employed in a theory, and generally to the issue of redundancy. That is whether all constructions are stored or whether the most parsimonious representation is sought (Dąbrowska, 2008:993-994, 948; Fried & Östman, 2004:6; Goldberg, 2006:214). At the same time, the issues related to storage are not unique to constructionist approaches. Instead, it is an active research question in experimental and computationally oriented linguistics (de Vaan, Schreuder & Baayen, 2007). The family of Construction Grammars certainly did not arise in vacuum. On the one hand, the roots of Construction Grammars are closely connected to the work of Fillmore (1968; 1970), specifically to Case Grammar, which formulated the basic principles of semantic/theta roles.[28] On the other hand, the roots of the family lie in the recognition of idiomatic structures that are typically considered to be located outside the grammar of language in formal approaches.

The basic tenets of Construction Grammar were already figured in the study by Fillmore et al. (1988) on the *let alone* construction and Lakoff's (1987) study on *there*-constructions. Another influential study is presented by Kay and

[28] According to Dowty (1991:548 footnote 3), one of the earliest proposals on semantic roles is given by Blake (1930) covering 87 temporal and locative roles and 26 other roles.

Fillmore (1999) on the *what's X doing?* - construction. In addition to these early proposals, Construction Grammars share the move away from derivations with other monostratal theories, such as Role and Reference Grammar (Van Valin & LaPolla, 1997), and Head-Driven Phrase Structure Grammar (HPSG) (Pollard & Sag, 1994). The former is established through typological studies and the latter is heavily influenced by computational modeling. Several parsers have been developed exploiting the theoretical apparatus of HPSG. At the same time, Construction Grammars are not simply a novelty. The insights of Construction Grammar are also adapted in formal approaches to lexical semantics. For example, Rappaport Hovav and Levin (1998) propose the concept of template augmentation, which captures the essence of a construction. It is a generalization over items which can be used further to motivate extensions. They explicitly discuss this concept in relation to Construction Grammar. Another expansion of the constructionist approach to the domain of formal approaches is the study by Jackendoff (1997) on idiomatic expressions labeled as 'time' –*away*. One example is *Bill slept the afternoon away*.

This study is committed to the so called Cognitive Construction Grammar or CCxG, which in return is perhaps best described as a mix of different approaches fused together under the flag of Cognitive Linguistics. Although the label is proposed by Goldberg (2006:214) and contrasted against typologically oriented Radical Construction Grammar (Croft, 2001) and Cognitive Grammar, recent constructionist studies combine these three approaches. Examples include, Barðdal's (2008) study on argument structures on Icelandic, and Divjak's and Janda's (2008) study on argument constructions in Russian. The commonality in these two studies is the fact that they attempt to cover both central and peripheral aspects of argument constructions. One motivation for this state is the fact that Construction Grammars in general are still in fairly early stages of development. Without a doubt, the Cognitive Grammar set forth by Langacker is the most mature and full-fledged theory of language within Cognitive Linguistics and its theoretical basis essentially covers all aspects of language whereas the state of Cognitive Construction Grammar is not as broad.

Recently, Langacker (2009) establishes the basic descriptive devices of Cognitive Grammar (Langacker, 1987; 1991) in terms of constructions, which can be considered, in my view, as an act of unification to bring separate theories under the common flag. Thus, the term Construction Grammar is used in this study as a general term to cover the family of Construction Grammars. The similarities are far greater than the dissimilarities between them compared to the position in formal approaches (cf. Goldberg, 2009a). Before turning to the description of the basic components utilized in the framework of Construction Grammar, a brief outlook on the development of the family of Construction Grammars is in order.

The most important commitment of Construction Grammar is to describe a language in its totality within a single framework without postulating radically different structures to capture general and productive patterns on the one hand, and idiomatic patterns on the other. From a theoretical perspective, only a single entity is posited, namely the Construction. This commitment is already set forth by Kay and Fillmore (1999:1), and also explicitly stated by Goldberg (2006:18). In short, Construction Grammar grew out from investigations on idiomatic structure and moved to cover broader generalizations in later stages. In this vein, constructionist approaches offer a uniform representation of grammatical knowledge (Croft & Cruse, 2004:255). The central role of idiomatic structures in the early versions of Construction Grammar originates from the idea that if a theory is capable of accounting for peripheral or non-core cases, the same theoretical machinery can also cover the regular types. This transition is evident, for example, in the definition of the Construction. In the early version of Cognitive Construction Grammar, Goldberg (1995:4) defined a construction in the following terms: "Constructions are taken to be the basic units of language. Phrasal patterns are considered constructions when something about their form or meaning is not strictly predictable from the properties of their component parts or from other constructions."

Goldberg (2006:5) modified the earlier definition to capture both idiomatic and general aspect of language: "patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency."[29] Effectively, Goldberg's proposal covers both compositional and non-compositional structures. Michaelis has proposed similar definitions for constructions: 1) Concord Construction which denotes the same kind of entity or event as the lexical expression with which it is combined, and 2) Shift Construction which denotes a different kind of entity or event from the lexical expression with which it is combined (Michaelis, 2004:28-29). The inclusion of frequency can be understood against the general movement within Cognitive Linguistics towards the usage-based framework. Moreover, the family of Construction Grammars has reached the point where the central aspects of language are covered within a single, albeit loosely formed, framework. Recently, Goldberg has concentrated on the psychological plausibility of Construction Grammar and acquisition and learnability of argument constructions. A tenet set forth by Goldberg (1995), and a number of experimental studies are summarized by Goldberg (2006).

Tomasello and his research group are focused on language acquisition based on constructions. The plausibility of usage-based and a constructionist approach to language acquisition as a holistic framework is exemplified by Tomasello (2003). Another branch of Construction Grammars are computationally oriented approaches adapting a unification-based formalism. Closely related to the cognitive approach is UCxG (Unification Construction Grammar) advocated by Fillmore, Fried, and Östman (cf. Fillmore, 1999; Fried & Östman,

[29] There is actually a slight glitch with this definition. A certain string or strings are only considered to constitute a construction if they display sufficient frequency. In order to have a frequency-based definition of construction, a certain degree of accumulation has to take place. Thus, even verbatim form has to be registered in some small manner. This issue is also discussed by Bybee (2010:17-18).

2004; Östman & Fried, 2005). This in return is closely tied to FrameNet established by Fillmore and his research group. A second proponent of this is Sign-Based Construction Grammar extending the framework of HSPG (Sag, 2010). A third unification-based model is Fluid Construction Grammar formed by Steels and his research group. It is a computational model of both parsing and production. Moreover, it is oriented towards research on artificial intelligence and robotics (De Beule & Steels, 2005). Another constructionist model is Embodied Construction Grammar which includes dynamic representations as part of constructions, like simulating motor and perceptual systems. Thus, this approach is more closely connected to the neural theory of language (Bergen, Benjamin & Wheeler, 2010; Bergen, Benjamin K. & Chang, 2005).

This section outlined the major strands of the family of Construction Grammar and general research topics actively pursued within them. Langacker's (2009:60) three point list of the basic operations required for a constructionist model sums up the previous discussion: "To describe a construction fully, one has to specify: 1) the meaning of each component element, 2) how these meanings are integrated to form composite conceptions at different level of organization, and 3) how the construction relates to others (its position in intersecting networks of constructions and constructional variants)." The following sections are devoted to different aspects of Construction Grammars with the focus on the interaction between verbs and generalizations. This study primarily builds on the theoretical basis of Cognitive Construction Grammar and Radical Construction Grammar. At the same time, there is also a strong connection to unification-based Construction Grammar as the concept of construction is modeled with variables. Thus, the relation between a verb and abstraction is gradually built throughout this chapter.

2.2 Interactions between Verbs and Abstractions

The crucial question for any linguistic theory addressing verbal semantics and argument structure involves formulating the basic representation of verbs. Moreover, the question is complicated by the fact most verbs have multiple argument structure realizations available for them. Rappaport Hovav and Levin illustrate this starting position with the verb *sweep*, as in 2.2-1–2.2-5.

2.2-1 *Terry swept.*
2.2-2 *Terry swept the floor.*
2.2-3 *Terry swept the crumbs into the corner.*
2.2-4 *Terry swept the leaves off the sidewalk.*
2.2-5 *Terry swept the floor clean.*

Example 2.2-1 demonstrates the intransitive type and the transitive is given in 2.2-2. Example 2.2-3 is commonly referred to as *into*-construction type in constructionist approaches. Examples 2.2-3 and 2.2-4 illustrate semantically even more extended patterns, namely bringing about a change of location and bringing about a change of state (Rappaport Hovav & Levin, 1998:97-98). The potentiality to combine with multiple argument realizations is not only limited to a few lexical items. On the contrary, this multitude of patterns appears to be the norm.

A list approach could be stipulated to handle this phenomenon that would simply state that each pattern associated with a particular verb describes the possibilities of the argument structure of a given verb. However, this position is simply undesirable because it would multiply the number of lexical entries of a verb to five in the given examples. Additionally, this approach would not be able to state any generalizations over this plethora of lexical entries. This would be counterintuitive considering what is known about categorization. People readily form generalizations over instances. This generalization also holds in Russian demonstrated by lexical studies within the Moscow Semantic School (Apresjan, 1974; Апресян, Ю. Д., 1995b; Падучева, 2004). Although the examples with the verb *sweep* cover only a small portion of possible realizations, they serve to highlight the importance of a well-grounded theoretical model to account for these kinds of patterns. Furthermore, Levin and Rappaport Hovav (2005:190) note that most of these patterns do not even have labels associated with them. Consequently, the following sections are used to establish a usage-based approach to describe the argument structure of verbs and the relations holding between different usage patterns.

2.2.1 Rules and Alternations

A number of theories have been pursued to tackle the role of the verb. The early generative paradigm posited that the argument structure of a verb is determined by its lexical entry, syntactically encoded in the form of a subcategorization frame (Chomsky, 1981). This approach evolved later into the projectionist or linking rule approach. The linking rule approach figures prominently in contemporary formal approaches. The most comprehensive catalog of verb alternations in English is formulated in Levin (1993). The proposed linking rules are considered to be universal or, at least, nearly universal. However, they are not stipulated in purely syntactic terms. Instead, they contain semantic information enabling the linking rule approach to handle differences in meaning between alternative realizations of the argument structure (Pinker, 1989).

The shared commonality between the various projectionist and linking rule approaches is that the lexical entry contains some kind of argument structure, which determines the morphosyntactic realization of its arguments, which is projection (Levin & Rappaport Hovav, 2005:186). Furthermore, this postulated argument structure of a verb is subjected to decomposition, a description akin to semantic primitives (cf. Goddard & Wierzbicka, 2002). In this sense, the argument structure of a verb is described in terms of primitive predicate types which are postulated to reoccur across different verb classes and ultimately across different languages, such as change of state and state.

The roots of semantic decomposition are formulated by Dowty (1979). The semantic decomposition has gained popularity in a number of different approaches to the argument structure of verbs in recent years, for example, in Role and Reference Grammar (Van Valin & LaPolla, 1997) and studies by Rappaport Hovav and Levin (1998). To illustrate the application of this approach to the argument structure of verbs, the verb *sweep* being an activity verb (ACT) can be formulated in the following manner based on the formalism proposed by Rappaport Hovav and Levin (1998:119): [x ACT$_{<sweep>}$ y]. Because the primitive predicates have their own argument structure, this allows to state generalizations over specific lexical items and slots available for them (x and y).[30] This position echoes Paducheva's (2004:30-31, 34) formulations of representing verbal semantics in Russian.[31]

A similar development is also present in the Russian diathesis tradition, an account that starts with a skeletal structure incorporating semantic components in the later stages. The diathesis tradition originates from the studies of the Saint Petersburg typological school in the seventies. In short, diathesis is a marked change in the argument structure of a verb operating on the correspondence between the syntactic and semantic levels. Furthermore, the diathesis tradition assumes that a diathesis is present in every verb because the syntactic and semantic levels can be assigned to them (Храковский, 1974). In this vein, the early version contained only elementary information needed for stipulating the change in the argument structure and was delimited to operate on syntactic and semantic role levels, yielding a two-tiered structure to represent a difference in argument realization.

Like most traditions, the diathesis tradition can be considered as an umbrella for a number of different approaches. For example, Mel'chuk's (cf. 1993; 1997) approach contains both a logical structure as a descriptive device to capture lexical semantics in addition to the diathesis alternation. In contrast, Paducheva's (Падучева, 2004) account builds on the generative paradigm, but also contains elements such as metaphor and metonymy (cf. Князев, 2007:78-79). For the purposes of the present study, the position and the descriptive devices used in the diathesis tradition are presented following Geniušienė (1987) as it is directly related to the subject matter at hand.

This two-tiered structure was further developed into a three-tiered model, which includes the level of referent yielding the classical division between form, meaning, and denotation. The latter position is part of Geniušienė's (1987) typological study of the Reflexive Marker. Examples 2.2-6 and 2.2-7 illustrate a

[30] Rappaport Hovav and Levin (1998:108-109) also posit modification as a second type in addition to predication. This claimincludes such components as manner and instrument directly as part of the basic argument structure of a verb.

[31] Paducheva proposes a set called Taxonomical Category, which is used to describe constant properties of lexical semantics. However, this category does not strictly follow the typically employed set of primitives; for example, causation is not considered as part of Taxonomical Category although such subclasses as tendency and relation are. Instead, causation belongs to the category of Thematic Class of verbs. (Падучева, 2004:30-33, 42-43).

change in diathesis with the Russian Reflexive Maker and Figure 2.2-1 gives its corresponding diathesis.[32]

2.2-6 *Он* *понури-л* *голов-у.*
He.NOM hang-PST.M head-ACC
He hung his head.
(Geniušienė, 1987:54)

2.2-7 *Он* *понури-л-ся.*
He.NOM hang-PST.M-RM
He hung his head.
(Geniušienė, 1987:54)

	$\boxed{\Delta_0}$	→	$\boxed{\Delta_1}$	
Referent level	Person	Part	Person	Part
Semantic role level	Actor	Patient	Actor	Patient
Syntactic role level	Subject	Direct object	Subject	Ø

Figure 2.2-1 Diathesis alternation, adapted from Geniušienė (1987:55)

The representation of Example 2.2-6 is given in diathesis Δ_0 and contains the following components. The referent level contains both Person and Part. The latter corresponds to the body-part (*голову*). The second level contains the semantic roles; Actor and Patient according to her semantic role inventory. The last level is used to establish the syntactic categories of subject and direct object. The change in diathesis from Δ_0 to Δ_1 corresponds to Examples 2.2-6 and 2.2-7. The difference between these formulations resides in the omission or incorporation of the direct object (marked with Ø) in Δ_1 (cf. Geniušienė, 1987:54-55). In this vein, an account based on diathesis links changes in argument structure in terms of correspondences between the semantic, syntactic, and referential levels.

A more recent development in the diathesis tradition is demonstrated by Paducheva (Падучева, 2004). The approach displays the current state of linguistics theories, namely that a number of different positions are fused together. They consist of the incorporation of the semantic class of verbs, such as motion and perception, and the inclusion of the taxonomical category reminiscent of semantic primitives and the diathesis with semantic and syntactic levels. Finally, changes in diathesis are motivated through metonymy and metaphor. The number of different components is, however, substantial. The account contains over 100 different semantic classes and components such as mental and perception verbs. Additionally, over 100 semantic roles and semantic components are included, such as Agent, Causer, and Location. Finally, slightly over 60 different ontological classes as stipulated, such as Process, Activity, and

[32] The glossing is slightly modified to be in line with the notation used in this study. Additionally, the examples are given in Cyrillic.

State (Падучева, 2004:585-589).

Generally stated, the diathesis tradition can be viewed as stemming from the Jakobsonian markedness theory. The descriptive devices formulated in the tradition operate on binary categories, namely between the marked and unmarked verb forms. This has a severe limiting factor if this kind of connection cannot be postulated; there is no theoretical basis to cover these non-paired items. Analogously, this is true for every position which assumes a strict correspondence between different verb forms. The reflexiva tantum verbs are a primary example of a problematic category of this kind.

One solution to this problem is to devise boundaries between different components. This is especially true for formal approaches. For instance, Rappaport, Hovav, and Levin posit a dichotomy between structural and idiosyncratic components. The structural components are assumed to be relevant for grammar, as they are used to relate reoccuring types. In contrast, the idiosyncratic components are used to differentiate verbs sharing the same structural aspects of meaning (Rappaport Hovav & Levin, 1998:106-107).

Interestingly, Rappaport Hovav, and Levin (1998:129) argue that this distinction is also present in constructionist approaches. The idiosyncratic aspect corresponds to the properties of the verb and the structural aspect relates to the argument constructions. Worded in this manner, the commonality between these two approaches becomes more apparent.[33] However, the crux of the matter is not in the descriptive devices posited by these two approaches, but in the mode which ultimately guides the goals of the positions. Levin, and Rappaport Hovav (2005:234) appeal to necessary and sufficient conditions by stating that: "accounts that include both necessary and sufficient conditions are the most successful at meeting this challenge." Although certain facets are shared across different approaches the appeal to necessary and sufficient criteria sets apart usage-based and formal approaches.

Bresnan et al. (2007) and Bresnan and Ford (2010) propose a probabilistic model for the dative alternation, which is one of the most studied alternation types in English. The model achieves over 90% classification accuracy and the results of the model are replicated in different varieties of English. Furthermore, the model is supported by experimental results. Thus, it is a theoretically motivated model, which is supported by converging evidence. Additionally, the variables used in the model are all established through usage and not through necessary and sufficient criteria. Perhaps, formal approaches would not consider this mode of analysis relevant for linguistics. Interestingly, this is also stated in Bresnan et al. (2007:27): "[…] with (traditional) theoretical linguistics regarding the problem of predicting the dative alternation as too difficult to tackle and as

[33] Dividing various theories into two poles is a simplification and stems from the labels used by Levin: projectionist and constructionist approaches (cf. Levin & Rappaport Hovav, 2005:ch. 7ch). The fundamental assumption in the projectionist accounts builds on the lexical entries of verbs, which determine the realization of their arguments.

outside the proper subject matter for linguistic theory. [...] [W]e suggest that by tackling problems of this kind, theoretical linguistics has the opportunity to build collaborative research with psychology, computer science, and allied fields and thereby deepens our understanding of the cognitive foundations of interpretation."

This issue revolves around the question on adequacy whether a probabilistic model capable of handling surface structures or a formal model operating on minimal pairs is preferred. If the position on necessary and sufficient criteria advocated by Levin and Rappaport Hovav is taken, the formulation they present for the dative alternations fulfils the conditions:

to variant:

x cause [y to come to be at (possession) z]

Double object variant:

x cause [z to come to be in STATE (of possession)] by means of
[x cause [y to come to be at (poss) z]]
(Speas 1990 via Levin & Rappaport Hovav, 2005:207)

The recognition of abstractions over verbs, which are assumed to be available in usage similar to concrete lexical items, has recently been adapted in non-constructionist theories. These abstractions are considered as basic event templates and the polysemous nature of verbs can then be attributed to Template Augmentation (Rappaport Hovav & Levin, 1998). Similarly, Paducheva (Падучева, 2004) uses metaphor and metonymy to explain changes in diathesis.

Before turning to the constructionist models, the commonality between the linking rule and diathesis tradition needs to be brought forth. They both explicitly adhere to the claim that one of the senses available for a given verb is the basic and all other possible senses are derived from it. This mode of analysis was implicitly present with the diathesis in Figure 2.2-1. The formulations were given labels Δ_0 and Δ_1. This mode imposes an inherent directionality. Certainly, a diachronic account can be used to motivate the derivational pathway. The discussion on reflexiva tantum verbs will show, however, that this directionality is certainly not as simple as non-reflexive verb → reflexive verb in Section 3.1.1. Derivational models at least in the Russian tradition, consider the relation to be semantic in nature. This is explicitly stated by Paducheva. However, no criteria are offered as to how this connection is established other than it is possible to find the basic one. (Падучева, 2004:149).

This paradox between the basic and the derived form is also discussed by Mel'chuk (Мельчук, 1995:459-460). The mismatches in hunting down the basic form are also illustrated in the previously established linking rule approaches. Examples 2.2-8 and 2.2-9 show the locative alternation, taken from Goldberg (2006:34-35).

2.2-8 *Pat loaded the wagon with the hay.*
2.2-9 *Pat loaded the hay onto the wagon.*

Rappaport, Hovav, and Levin argue that the locative variant is the basic structure and the *with*-variant is derived from it. To support this interpretation, they propose that the locative entails the *with*-variant. The argument functioning as location can be used as direct object with the *with*-variant similar to Pinker's proposal (Levin & Rappaport Hovav, 2005:206-207). However, there are apparent mismatches with this test depending on the selected verbs, (e.g., between *pile* and *stuff)*. For these verbs, the derivational pathway can go either way (Pinker, 1989:38-39, 125). The directionality is supposed to be the strongest form of generalization, but once the number of selected verbs is increased, the inherent directionality becomes aberrant and the derivational pathway can take either form. It is difficult to see the motivation of the strong generalizations, especially if the directionality is supposed to rest on necessary and sufficient criteria.

In contrast to these formal approaches in the early diathesis tradition frequency of use and genre were considered to be factors through which the basic form can be established (Храковский, 1974:13). To my knowledge, this has never been fully developed in the Russian tradition. Nonetheless, recent probabilistic models on word formation support this frequency-based approach. Hay found support for the relative frequency hypothesis. Participants were shown pairs of items such as *inaudible* ~ *audible* and *imperfect* ~ *perfect*, and they were asked to indicate which of the items were perceived to be more complex. In the first pair, the complex form has a higher frequency compared to the base form whereas opposite holds in the second pair. The results indicated that if the base form is more frequent, then the participants were more likely to consider the complex form to be more compositional and the opposite held for the more frequent complex forms relative to the base (Hay, 2001). These findings lead to the situation where the individual components of a morphologically complex word can still be compositional, but the parts are not perceived as analyzable (cf. Langacker, 1987:292). These results indicate that frequency plays an important role both in inflection and word formation (cf. Bybee, 2010).

This section has illustrated previous non-constructionist accounts to verbal semantics and descriptive devices. These accounts are important in framing Construction Grammar. Nonetheless, there is an interesting commonality or trend across different accounts. Most theories on argument structure have moved away from highly abstract representation towards more lexical and elaborated structures.

2.2.2 Schematicity and Schemata

A shared commonality between the Cognitive Construction Grammar and Radical Construction Grammar is the principle of semantic content which separates different constructionist approaches. For example, unification-based and Sign-Based Construction Grammar allows syntactic constructions whereas Cognitive Construction Grammar adheres to content requirement originally proposed by Langacker. Accordingly, the following elements are recognized in Cognitive Grammar: 1) semantic, phonological, and symbolic units that actually occur, 2) schematization, and 3) categorization. In this vein, the principle of the content requirement rules out semantically empty categories (Langacker, 2009:2-3). The same position is taken in this study and only pairs of form and meaning are postulated. The crucial issue is, nonetheless, related to the definition of form and meaning which is gradually defined in Sections 2.2.2 and 2.2.3.

A second important property in Cognitive Construction Grammar is the syntax-lexicon continuum. Instead of establishing a sharp division or separate modules for lexical and syntactical categories, they form a continuum and the concept of construction is used to capture language in its totality. This position is already taken by Langacker (1987:25-27), by Croft (2001:19), and later by Goldberg (2006:5). Table 2.2-1 represents the syntax-lexicon continuum and its basic units.

Construction type	**Traditional name**
Complex and (mostly) schematic	syntax
Complex and (mostly) specific	idiom
Complex but bound	morphology
Atomic and schematic	syntactic category
Atomic and specific	word/lexicon

Table 2.2-1 Syntax-lexicon continuum, adapted after Croft (2001:19).

Another important distinction between various Construction Grammars is the status of the atomic and schematic units, (i.e., syntactic categories). In terms of Radical Construction Grammar, syntactic categories do not exist independently of constructions. Thus, they are distributional properties (Croft, 2001:18, 159-164). The same principle is adapted in this study, and discussed in Section 3.2.3.

Table 2.2-1 highlights important concepts for purposes of the present study that are intertwined with the role of verbs and the argument constructions in general. Importantly, the concept of schematicity should be kept apart from type frequency and consequently from productivity as is argued and demonstrated by Barðdal (2008:40-45) and Bybee (1995:452-453). The continuum is another aspect which is not necessarily shared among all strands of Construction Grammar. Kay (2002b) argues against the continuum by introducing a separation between constructions which are productive and those which are defined as patterns of coining pertaining to the traditional distinction between grammar and lexicon or productive and idiosyncratic aspect of grammar.

The types of argument construction explored in this study range between productive and fairly unproductive. Extensive discussion related to productivity and patterns of coining is given by Barðdal (2008:26-39). The issue whether a specific argument construction is productive, is not the purpose of this study. Rather, it is related to the concept of schematicity. The Passive Construction serves to illustrate the maximal end of the scale being both productive, and schematic leading to the situation that it is traditionally considered as a separate type in the category of the Reflexive Marker. However, other argument constructions are considered productive. For example, the Experiencer Perspectivization Construction is supported with such verbs as *житься* 'live' and *спаться* 'sleep' (Israeli, 1997), although systematic evidence to support this has not been offered.

To substantiate the productivity claim, Kyröläinen (2008) used the principle of analogical extension proposed by Osherson et al. (1990). If a particular argument construction is productive, language users are likely to extend it to cover novel instances, especially when the items in question are semantically related. This principle is also used by Goldberg (2006) and Barðdal (2008). Additionally, experimental evidence to support this is given by Suttle and Goldberg (2011). To the test productivity hypothesis in terms of analogical extension, Kyröläinen used the closed-class set of verbs of motion in Russian. They obviously constitute a semantically coherent category and due to their close-class status they offer well-delimited type.[34] Based on frequency counts extracted from the Integrum database, a full coverage among the verbs of motion was obtained.[35] However, such specific inflectional forms as *ползется* 'crawl' and *плывется* 'swim, float' indicating that even Integrum, containing a few billion words at the time, was not large enough to establish complete coverage in terms of inflectional forms (Kyröläinen, 2008:185-186).[36]

This particular argument construction type serves to illustrate both general principles and particular facets of the Russian Reflexive Marker. First, even infrequent argument construction types can display, at least, partial productivity. Second, schematicity and productivity, although typically intertwined, need not be as this particular type may be considered to be fairly specific. It is supported by a fairly stable lexical encoding of the modifier demonstrated by Kyröläinen (2008:189) with simple collocation counts. Third, extensions are probabilistic because the extensibility is based on particular items contrasting rules which are inherently binary and categorical. A particular item may be either closer or further from its categorical center leading to a degree membership and probabilistic behavior (cf. Bybee, 2010:73-74).

A third aspect of Cognitive Construction Grammar is the surface generalization hypothesis. Goldberg (2002; 2006) sets forth strong arguments

[34] Gerritsen (1990:179) also discusses the productivity of this type in terms of verbs of motion.

[35] Integrum is a commercial project and not specifically aimed for linguistics. It covers virtually all newspapers published in Russian in an electronic format. The database is updated daily. Thus, it resembles monitor corpora. Introduction to the usage of Integrum is provided by Laguta and Timofeeva (Лагута & Тимофеева, 2007), Nikiporec-Takigava (Никипорец-Такигава, 2006), Mustajoki (2006), and Kyröläinen (2008).

[36] At the time, the Russian Nation Corpus contained proximately 140 million words. Although a fairly large corpus it was insufficient to offer any discussion on coverage related to this particular construction type.

against derivations or alternations, attributing the basic tenet of this position to Chomsky. His claim was based on accounting for the syntax of NPs, which are derived from nouns. Chomsky argued that these NPs have the same syntax, (e.g., *refuse ~ refusal).* Thus, they should not be stated in terms of derivations. Instead, they should be considered to be base generated in generative terminology (cf. Chomsky, 1970:215). Goldberg (cf. 2006:26-28) exemplifies the surface structure hypothesis with the Ditransitive Construction illustrated in 2.2-10 and 2.2-11. Arrows indicate their paraphrases. Traditionally, these patterns are referred to as the dative alternation.

2.2-10 *Mina bought a book for Mel.* → *Mina bought Mel a book.*
2.2-11 *Mina sent a book to Mel.* → *Mina sent Mel a book.*

Goldberg shows that contrary to the derivational approach, the Ditransitive phrases pattern alike contrasting the prepositional paraphrases. Examples 2.2-12 and 2.2-13 extracted from Goldberg (2006:27) illustrate the issue at hand.

2.2-12 **Mina bought Mel yesterday a book.* → *Mina bought a book yesterday for Mel.*
2.2-13 **Mina sent Mel yesterday a book.* → *Mina sent a book yesterday to Mel.*

Barðdal (2008:45-46) provides an analysis of the Ditransitive Construction in Icelandic consisting of seventeen subclasses. In terms of probabilistic models, Bresnan et al. (2007) and Bresnan and Ford (2010) have showed that the dative alternation is predictable based on surface structure. Crucially, the corpus-based results also hold in experimental settings. However, the surface structure hypothesis does not explicitly deny the possible interconnectedness between paraphrases or related patterns.

Interestingly, Zolotova has proposed a model of core sentence patterns for Russian. These patterns are assumed to be formed based on generalizations over a multitude of observed sentence patterns. Additionally, the meaning of a particular component is established in interaction with the other components present in the core sentence pattern. Because the components form a core meaning, the account does not propose pure syntactic categories (Золотова, Г. А., 2005 [1973]:25-26). A stock of the core sentence patterns is conveniently summarized by Leinonen (1985:21-24). For example the core meaning of subject of kinship or social relationship is instantiated with the pattern Noun$_{\text{nom}}$ Copula Noun$_{\text{dat}}$ Noun$_{\text{nom}}$ (*Он мне отец*).[37] Another important property of the core sentence patterns is that they have their own meaning and structure. Thus, Zolotova (Золотова, Г. А., 2005 [1973]:60-61) posits that the dative Experiencer in *мне грустно* 'I am sad' is due to the core sentence pattern of state (Noun$_{\text{dat}}$ Copula Adverb) in Russian and it is not part of the predicate. These properties make the account compatible with constructionist and usage-based models.

[37] The copula 'be' is considered as part of the core sentence, but it is typically omitted in present tense.

Nonetheless, the surface structure hypothesis represents an extreme position similar to the derivational approaches, albeit from a different perspective. Bybee has argued for the existence of two types of general schemata: source- and product-oriented. These account for the formation of categories. The source-oriented schema roughly cohere to the traditional definition of rule, (e.g., for *wait* and *waited).* Thus, the source-oriented schema is a generalization over pairs. In contrast, the product-oriented one corresponds to generalizations over a set of complex forms, like *strung*, *stung*, *flung* and *hung* (Bybee, 1995:430-433). Importantly, the source-oriented schema is not a derivational rule. Instead it is a form of a generalization, such as the productive past tense type -*ed*.

Bybee's schema-based approach contrasts the rule-based model where the source-schema corresponds to a rule and the product-schema to stored lexical patterns advocated by Pinker (1989; 1999). At the same time, the earlier approach has undergone several modifications. The later model proposed by Pinker and Ullman (2002), for example, does not posit that regular forms are never stored, only that they do not have to be. Thus, the form *waited*, for can be stored. It is also worth pointing out that Goldberg's surface structure hypothesis does not deny the existence of cross-paradigmatic relations or paraphrases. Her position is clearly illustrated in the following quotation: "The arguments […] should not be taken to imply that possible paraphrase relations play no role in the learning, processing, or representation of language" (Goldberg, 2006:43).

It follows from this that the Russian Reflexive Marker itself is a form of generalization, removing the requirement to establish a division between the form, which is the Reflexive Marker, and the lexical reflexive verb. This issue was already discussed in Section 1.2.1 in relation to the traditional difference between the Passive Construction, which is postulated to be a form-based contrasting the lexical reflexive verbs. The reflexive verb *строиться* 'build' can be used to demonstrate degrees of connectedness between different abstractions. The Passive Construction once again serves to illustrate the source schema, namely the relation between the Passive and Transitive Constructions amounting to an interconnection between the reflexive verbs *строиться* 'build' and the non-reflexive verb *строить* 'build.' When multiple argument constructions are factored, the combination with the nominative case *строиться* and the prepositional phrase *на* aligns with *базироваться на*$_{\text{prep}}$ 'be based' discussed in detail in Section 9.5. Thus, a specific instance of a particular verb may be more strongly associated with a specific type of generalization rather than being an automatic process (cf. Booij, 2010:4, 88-89).

In sum, a network model can be used to include variation in terms of probabilities rather than resorting to postulating multiple rules and speaker-specific sensitivity to them (Krasovitsky, Baerman, Brown & Corbett, 2011:574; cf. Падучева, 2006). Instead of viewing lexical items as static pairs, the basic tenets of the network model become essential, in a sense, that a specific verb may display a varying degree of strength between different usage patterns and abstractions.

2.2.3 Verbs in Construction Grammar

Goldberg's account on describing the relation between verbs and argument constructions is stipulated in terms of two different sets of roles. Verbs are associated with participant roles while argument constructions are associated with argument roles, which are generalizations over participant roles. The argument roles are another mode for stating semantic roles such as Agent and Patient. The participant roles of a verb are defined in relation to frame semantics (Goldberg, 1995:43-44). The analysis based on frame semantics is further elaborated by Goldberg (2010). In short, frame semantics is intended to capture the encyclopedic knowledge required to understand the meaning of a particular word (cf. Baker et al., 1998; Fillmore, 2007). Thus, the structure of a particular lexical verb may be subjected to a high degree of idiosyncratic properties. However, the slot in an argument construction is a generalization over these particularities (Fried & Östman, 2004:40-43). Although the participant roles are intended as part of frame semantics, in practice, they resemble the proposal made by Dowty (1991:550). These participant roles are labeled as individual thematic roles, for example, the verb *hit* has such subject role as *hitter*.

Goldberg illustrates the analysis with the verbs *steal* and *rob,* as their difference rests in the profiling of the associated frame.[38] According to Goldberg's analysis, *rob* profiles the *target* and the *thief,* while *steal* profiles the *valuables* and the *thief*. When the verb is used with an argument construction, the participant and the argument roles are fused (Goldberg, 1995:45-46). Goldberg's proposal on argument constructions is stipulated in terms of two principles: 1) the Semantic Coherence Principle, and 2) the Correspondence Principle. According to the first principle, only those roles which are semantically compatible can be fused. The second principle, which is also assumed to be the default one, states that each participant role that is lexically profiled and expressed must be fused with a profiled argument role of the construction (Goldberg, 1995:50). Later, Goldberg (2006:40) has widened the scope of the Correspondence Principle to cover cases where it is overridden, for example, by an oblique argument. Generally, the Passive Construction can be characterized as a prime example of overriding the default assumption, the oblique encoding of the Agent argument.

Figure 2.2-2 gives the structure of the Ditransitive Construction in English according to Goldberg (2006:20-21).

[38] Goldberg's analysis is not based on the computational model of FrameNet available for English.

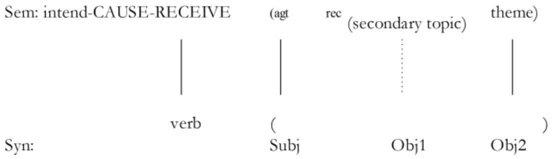


Figure 2.2-2 Structure of the Ditransitive Construction adapted from Goldberg (2006:20-21).

Generally, an argument construction consists of the semantic (Sem) and the syntactic (Syn) pole. The semantics of the verb *give* is described as CAUSE-RECEIVE. The semantic pole covers the semantic roles, Agent (agt), Recipient (rec), and Theme. They are connected to the syntactic pole: Subject, Object$_1$ and Object$_2$. The solid lines are used to indicate the participant roles that must be fused with the semantic roles of the Ditransitive Construction whereas the dashed line is used to indicate the semantic role that can be contributed by the argument construction. In this vein, the Recipient need not be part of the lexical verb. Instead the semantic roles can be provided by argument constructions.

Argument constructions, at least in terms of Goldberg's proposal, resemble the diathesis tradition in a sense that the primary means of connecting verbs and generalizations are through the semantic and syntactic slots. The difference is found that the argument constructions are generalizations over surface structure whereas the diathesis tradition relies on the alternation. Additionally, the argument construction is an abstraction; they primarily encode the linguistic structure in the form of who-did-what-to-whom. Thus, manner is rarely assumed to be encoded with argument constructions (cf. Goldberg, 2006:106; Tomasello, 2003:126).

Recently, certain constructionist accounts have moved away from the "dichotomy" between the verb and the argument construction to more elaborated representation by including, for instance, verb-specific constructions (Iwata, 2008:36-37). This type of abstraction is also considered by Croft (2003), Boas (2011 and references therein) and Barðdal (2008).[39] This move emphasizes

[39] In addition, recent studies have also included verb-class-specific constructions covering traditional semantic classes of verbs (cf. Barðdal, 2008; Iwata, 2008; Levin & Rappaport Hovav, 2005; Падучева, 2004). The semantic annotation available in the Russian National Corpus was extracted for the unique reflexive verbs (*n* = 819) in the database (Кустова, Г. И., Ляшевская, Падучева & Рахилина, 2005). According to the help file, 27 different semantic types are included. Unfortunately, the information is too sparse in its current form; only 38% of the unique reflexive verbs were associated with a semantic tag. Thus, the information could not be exploited for purposes of the present study. The semantic tags also contain inconsistencies related to both coverage and symmetry. First, the verb *драться* 'scuffle, fight' is tagged as an impact verb. In contrast, the verb *бороться* 'fight' does not have semantic information although it can be considered pertaining to same semantic category as *драться*. Second, the tagging is not

the fact that verbs and argument constructions do not reside in isolation, but are always part of some larger structure (Croft, 2003:64). Verb-specific constructions are generalizations over particular usage patterns, (i.e., covering all patterns of a particular verb). If a particular verb combines with two different patterns, the verb has two different verb-specific constructions.

Barðdal's analysis of Icelandic Nominative-Accusative Construction exemplifies the move towards the incorporation of different levels of schematicity, in Figure 2.2-3. The analysis is based on corpus-data, the corpus of Modern Icelandic texts, and covers 303 verb-specific constructions illustrated with the lowest branches. The second level covers the verb-class construction types (*n* = 46). Barðdal also includes even more elaborated account by including an ontological layer (*n* = 6), (i.e., the MAKING and MOVEMENT). The highest level of abstraction is captured with the Nominative-Accusative Construction (Barðdal, 2008:46, 63-67, 68).

always symmetrical. For example, the perfective verb *закрыться* 'close' is classified as a change of state verb whereas the imperfective *закрываться* has no information available. Crucially, the tagging appears to be relative, biased towards certain categories based on the extracted information. The move tag contains 96 instances and the second highest count is with the psych_emot tag, 32. These classes appear to be fairly disproportional in comparison to the number of the tagged reflexive verbs (n=316). An elaborated evaluation, albeit a critical one, of the semantic tagging used in the Russian National Corpus is given by Kretov (Кретов, 2009). Due to these reasons, verb-class-specific constructions are not considered in this study.

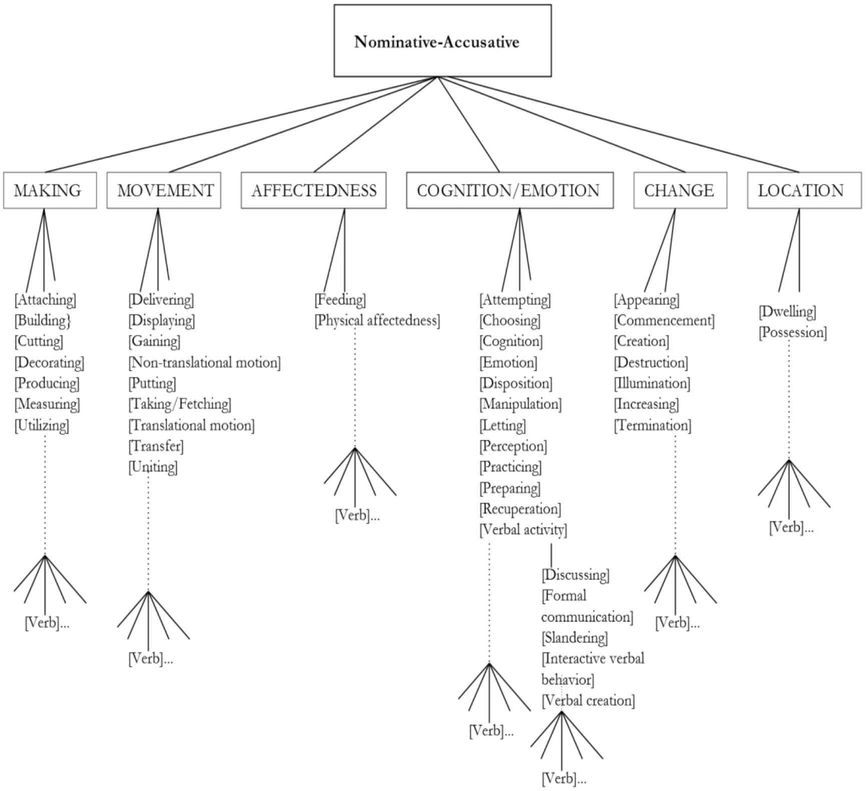


Figure 2.2-3 Structure of the Icelandic Nominative-Accusative Construction, adapted from Barðdal (2008:68).

Similarly, Iwata (2008:88) argues that the meaning of the verb and the meaning of the argument construction cannot be automatically separated. The separation of these different levels of abstractions is also related to the slight bias present in constructionist approaches. The vast majority of the studies are dedicated to the idiosyncratic stock of argument constructions so the number of studies on the compositional argument constructions is small, as reflected also in the evolution of the definition of the construction from idiosyncratic structures to the inclusion of the compositional ones. The commonality of this move towards the different levels of abstraction is a logical conclusion from the assumption of schematicity. Categorization is assumed to represent different levels of abstraction and a similar mode should be reflected in descriptive practices.

For purposes of the present study, the verb-specific constructions are included by default. Simply because (case) patterns are included as a variable, different verb-specific constructions follow from this naturally. Similarly, Faulhaber (2011:282) has criticized how grammatical functions are used to mask differences in usage patterns. This issue is also present in Construction Grammar. For example Iwata uses the syntactic frame of NP V NP PP to capture the locative-as-object variant. Examples 2.2-14–2.2-16 taken from Iwata illustrate the usage patterns of this syntactic frame. (Iwata, 2008:36-37).

2.2-14 *John put the box on the desk.*
2.2-15 *John threw a ball into center field.*
2.2-16 *John sprayed paint onto the wall.*

The use of the syntactic frame abstracts over specific usage-patterns that differ in form, specifically the encoding of the PP as in *on*, *into,* and *onto*. To avoid conflating usage patterns with grammatical functions, verb-specific constructions are defined relative to the patterns and the argument constructions relative to grammatical functions in this study.

The recognition of the verb-specific constructions captures two levels of variability, the within- and between-variability in argument constructions. Examples 2.2-17 and 2.2-18 illustrate the within variation of the verb *оказаться* 'seem, appear' in the Property Construction defined as profiling a property of an entity, as discussed in Section 8.1. The within-variability concerns the Nominative-Nominative and Nominative-Instrumental patterns. Importantly, not all reflexive verbs that appear in the Property Construction display this same variability between the patterns. Instead, they display a degree of connectivity. For instance, *являться* 'be' is associated with the Nominative-Instrumental pattern as in 2.2-19.

2.2-17 *А*   *стенк-а*   *оказа-л-а-сь*   *тоненьк-ая* […].
Also   wall-NOM   appear-PST-F-RM   thinnish-NOM
The wall appeared to be thinnish.
[1985, RNC, Татьяна Рик. Про вредную Бабку-Ёжку // "Мурзилка," №6," 2001]

2.2-18 *А это, скажу я вам,*
[…]
оказа-л-о-сь   *нелёгк-им*   *испытани-ем.*
appear-PST-N-RM   difficult-INS   test-INS
And this, I tell you, appeared to be a difficult test.
[710, RNC. Елена Павлова. Вместе мы эту пропасть одолеем! // "Даша," №10," 2004]

2.2-19 *Термин*   *"потребительск-ие*   *запас-ы"*
Term.NOM   consumer-NOM.PL   stock-NOM.PL
явля-ет-ся   *довольн-о*   *часто*   *употребляе-м-ым.*
be-3S.PRS-RM   quite-ADV   often   use-PRS.PP-INS
The term "consumer stocks" is used quite often.
[145, RNC, Потребительские запасы — сущность и подход к анализу // "Вопросы статистики," 2004]

The between-variability with *оказаться* 'seem, appear' is demonstrated in Example 2.2-20, defined as the Content Construction in this study. The argument construction is used to profile a focus on the content of communication or perception, cf. Section 5.11 for discussion of this type.

2.2-20 *Оказа-л-о-сь,* *что* *кто-то* *похити-л*
Appear-PST-N-RM that someone.NOM steal-PST.M
волшебн-ый *амулет* *принцесс-ы.*
magic-NOM amulet.NOM princess-GEN
It turned out that someone had stolen the magical amulet of the princess.
[1711, RNC, Сергей Седов. Доброе сердце Робина // "Мурзилка," №7," 2002]

Figure 2.2-4 demonstrates the position taken in this study and how the relations between the Argument Construction and the Verb-Specific Construction interact based on the previously given examples. The lexical items are nested within the patterns and, in return, are connected to the schematic argument constructions. The nested structures are used to convey the assumption in usage-based models that every usage of a particular item activates, at least partly, all the instances associated with it.

Every usage of *оказаться* 'seem, appear' in a specific configuration (i.e., Example 2.2-17) will also partly activate all the other patterns associated with *оказаться* 'seem, appear.' This partial activation leads to the accumulation of frequency and, ultimately, network structures (Bod, 2006; Goldberg, 2006; Hay & Bresnan, 2006; Pierrehumbert, 2001). Additionally, the dashed lines indicate connectivity between types without imposing directionality. The dashed lines and the circles are operationalized and defined in quantifiable terms in Chapter 3. The commonality between the reflexive verbs is the shared Nominative-Instrumental pattern. Thus, the verbs display partial overlap in terms of their patterns in relation to the Property Construction.

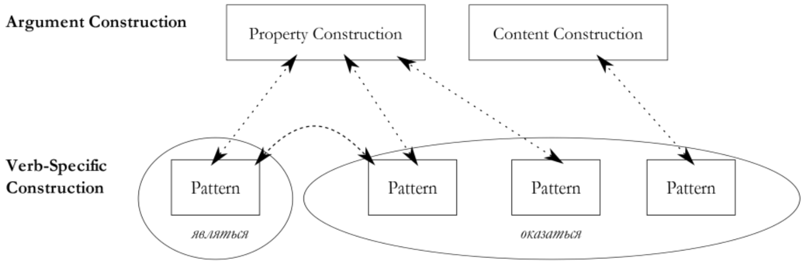


Figure 2.2-4 Idealized relations between the argument constructions and the verb-specific constructions.

If patterns are not recognized the question arises about the process of arriving at the semantic similarity of verbs and, ultimately, to the semantics of argument constructions. In my view, the assumption made here is compatible with the notion expressed by Rappaport Hovav and Levin (1998:99) that similar patterns lead to general verb patterns (cf. Кузнецова, Э. В., 1989; Падучева, 2004). Consequently, these generalizations are considered to constitute the building blocks of the argument constructions.

The modification proposed here, however, implies that the form pole of the argument construction is probabilistic in a sense that it emerges in interaction with the patterns of particular verbs in usage. Similar interpretation might be attributed to Bybee's discussion on constructions: "Constructions also have exemplar representations, but these will be more complex […], they have positions that can be filled by a variety of words or phrases. In addition, many constructions allow the full range of inflectional possibilities on nouns, adjectives and verbs, […]." Bybee assumes that constructions have central members but most constructions vary in the extendibility of the slots.

Related to the discussion on the relationship between verbs and argument constructions, Nørgård-Sørensen (2010) sets forth a strong claim in arguing that Old Russian is a construction-based system contrasting Modern Russian as a valency-based one. Furthermore, he posits that valency and construction are fundamentally different aspects or theoretical types. However, he (2010:50 footnote 2) notes that the label construction is defined in a narrow sense and not as in "certain varieties of so-called Construction Grammar," which define morphological categories also as constructions. As this study belongs to the family of certain varieties of so-called Construction Grammar, the abrupt division between valency and construction is less fundamental. The question is, nonetheless, important. Nørgård-Sørensen claims that the verb determines the shape of the (case) patterns in Modern Russian. The argumentation, more or less, rests on Kris'ko's (Крысько, 2006) extensive studies on Old Russian, specifically on the grammaticalization of transitivity, which is a later phenomenon in Russian. Thus, the claim by Nørgård-Sørensen is fully quantitative in nature; in Old Russian the verb *v* appeared with the valency$_n$ contrasted to contemporary Russian where the valency$_n$ of the verb *v* has substantially decreased or reduced to one.

In order to substantiate a claim of this kind of global magnitude, a random sample of verbs would have to be compiled. Otherwise, the verbs cannot be considered representing, in an optimal situation, the underlying population and show that the valency$_n$ has decreased or at least substantially changed. Such evidence is not provided by Nørgård-Sørensen (2010). Thus, the claim, albeit an interesting one, is hypothetical. It is also worth pointing out that Goldberg discusses the same issue in relation to Turkish and Hindi in contrast to English. It may be the case that in certain languages the verbs may possess a higher cue validity compared to argument constructions. However, it does not follow from this that people would not make generalizations over usage patterns (Goldberg, 2006:120).

At the same time, it seems that there is a certain degree of merit in Nørgård-Sørensen's argumentation. Neighbor verbs, which are verb forms with the same phonological form excluding the Reflexive Marker, were constructed for all the unique reflexive verbs amounting to 717 unique neighbor verbs, cf. Section 3.1.1. The valency types of these verbs were extracted from Efremova (Ефремова, 2000) and supplemented with data from Kuznecov (Кузнецов, 2009 [1998]). Obviously, dictionary-based counts are not equal to a detailed lexicological study but they offer, at least, a systematic approximation. The vast majority of the neighbor verbs are classified as univalent: either transitive (*n* = 533) or intransitive (*n* = 24). Nonetheless, the number of bivalent verbs is fairly large (*n* = 160). Examples are *требовать* 'claim, demand' with accusative, genitive or infinitive, *наблюдать* 'supervise, observe' with *за*$_{\text{ins}}$ or accusative, and *драть* with accusative in the 'break' sense and intransitive in the 'run away' sense (cf. Янко-Триницкая, 1962:74-76). If we follow the dictum of the Moscow Semantic School that two instances constitute a case of regular polysemy, the strong valency-based claim appears to be less potent (Apresjan, 1974). A comprehensive analysis based on the reflexive verbs cannot be provided in this study as the sampling frame was never designed to be a repeated measure over particular reflexive verbs. A more comprehensive probabilistic model on multiple argument structures in Russian estimated with dictionary-based counts is given by Kyröläinen (2012).

In sum, this section introduced the basic design used in this study to model usage patterns of verbs. Verb-specific constructions are assumed to be formed relative to patterns leading to overlapping structures across them. Thus, a generalized argument construction is supported by these partially overlapped verb-specific constructions.

2.3 Summary: Towards Layered Structure

This chapter introduced the basic and central aspects of Construction Grammar. Constructions are considered as the pairings of form and meaning located at different levels of schematicity. Recent advances in Construction Grammar have incorporated both verb-specific and argument constructions allowing to establish more fine-grained analysis.

Two types of relations, the source- and product-oriented schema, were established in this chapter and they anchor the formation of complex categories in usage-based models. In this study, it is referred to as the network model. Importantly, constructions are assumed to form a systematic network structure rather than being a list of unrelated types (Bybee, 2010; Goldberg, 2006; Langacker, 2009). The network analysis sets constructionist models apart, for example, from the diathesis tradition where items are portrayed as binary oppositions and generalizations do not intersect.[40] Thus, Chapter 3 expands the theoretical basis of the Construction Grammar by moving towards a more detailed operationalization of the basic components. Effectively, the questions are centered around defining the slots in the argument construction in measurable terms that would also facilitate comparison of different argument construction types on a global level.

[40] Recently, Paducheva (Падучева, 2004) considers that also derivations in the lexicon form paradigms.

3 Defining the Layered Structure of Argument Constructions

This chapter is devoted to theoretically motivate the properties of the proposed layers of argument constructions. Additionally, the discussion and the motivation of the layers explicitly bring forth the values/levels used to encode the variables. The proposed layered structure can be divided into two sets. The first set covers the verb slot: the lexical verbs and their structural properties and degree of connectivity. Bybee (1985) has proposed that the concept of degree of connectivity is based on three properties: phonological similarity, semantic relatedness and frequency of use. These properties are operationalized and intergraded into the model. In theoretical linguistics, a verb is defined as an inherently relational entity connected to such concepts as process and state, (Croft, 1991; Goldberg, 1995; Langacker, 1987). Thus, the second set covers the argument slots, which are also intertwined with the verb slot.

At the same time, it is worth bearing in mind that when a certain property is operationalized, its definition changes and becomes narrower as is discussed, for instance, by Stefanowitsch (2010). The consequences of operationalization are apparent in this study. For example, the definition of a verb is not a process or state, strictly speaking, but the combination of the proposed variables. Similarly, the definition of the Argument Construction changes from the pairing of form and meaning to the combination of the full set of the proposed variables. A model is the mediator between the theory and the phenomenon in question. Thus, any form of operationalization has to be anchored to some theoretical basis or the meaningfulness of the model is questionable (Suárez, 2004; Suárez & Cartwright, 2008).

The final issue related to operating with variables is the question of the exact number and their status both within language and the proposed model. Stokhof and van Lambalgen (2011:91) note that the ontological status of language is diverse. It is partly connected to physical and biological, social, cultural and historical aspects of language (cf. Levinson & Evans, 2010:2746). This leads to a situation where variables can basically be added ad infinitum, especially in usage-based models where all aspects of language are considered relevant (cf. Arppe, 2008:29). By focusing on the structural properties of the verbs, a model abstracts away from the other possibilities. The exclusion of certain variables, however, should not be understood as deeming them irrelevant to grammar, (i.e., the arguments in the generative paradigm for setting up the division between grammar and usage).

Construction Grammar has moved towards a more fine-grained set of semantic roles compared to the classical two-role system proposed by Dowty (1991:547; cf. Levin & Rappaport Hovav, 2005). If a fine-grained set of semantic roles would be used, such as Mover for the Motion Construction, the semantics of the argument construction and the semantic role would effectively be a one-to-one mapping. In contrast, the contribution of the semantic roles would be valuable if the primary goal was to explore the motivational pathways between different argument constructions and how verb-specific constructions incorporate extensions. Finally, linear order effects are not considered. These are primarily governed by information structure in Russian (Сиротинина, 2006 [1965]) and generally considered as constructions of their own in Construction Grammar (Goldberg, 1995).

The following principles are formulated to function as guidelines for the proposed constructionist model: 1) portability, 2) compatibility, and 3) cognitive plausibility (Fried & Östman, 2004; Goldberg, 1995). First, only those potential variables are considered that allow implementing them in some other setting on argument constructions if they have not been previously implemented. Semantically highly opaque variables are not used in this study, such as the degree of affectedness (cf. Levin & Rappaport Hovav, 2005) or control (cf. Haiman, M. H., 1991). Second, guidelines maintain a dialogue between usage-based models and the diathesis tradition. Third, the guidelines should lead us to formulate testable hypotheses about the structure of the Russian reflexive verbs at least with some of the proposed variables in the model.

In sum, this study focuses on phrasal patterns, or what Zolotova (Золотова, Г. А., 2005 [1973]) labels as core sentence patterns. In this vein, the contextual factors are viewed in a narrower sense than by Divjak and Gries (2006), Arppe (2008) or Bresnan et al. (2007). Compared to the previously mentioned studies, the locus is slightly different. For instance, Arppe offers a probabilistic model of four ‘think’ verbs in Finnish, (i.e., whether it is possible to predict which of the ‘think’ verbs are used in a given context). The possible contextual factors involved in modeling near-synonyms require most likely a fine-grained set of variables compared to a stock of argument constructions, which are situated at a fairly coarse-grained level in general (cf. Goldberg, 2006:43-44). Additionally, the number of argument slots is limited to two. Nonetheless, the superimposed limitation still yields grammatical core sentence patterns in Russian similar to Zolotova's account (Золотова, Г. А., 2005 [1973]). Although the operationalization of usage patterns is narrow compared to the previously mentioned studies, the proposed model contains variables which have rarely been included in studies of argument structures, specifically lexical networks and their structural properties.

3.1 The Layered Structure of Verbs

This section anchors the concept of the verb-specific construction to usage patterns by exploring the structure of the lexicon from the perspective of the verb paradigms and morphological categories. Additionally, the concept of neighbor verbs is formulated that allows moving away from the traditional pair account. The following sections build on the concept of network by introducing neighborhoods, (i.e., lexical networks). The concept of network figures prominently in usage-based models. For instance, Bybee (1985) proposed a network structure for the formation of the past tense in English. A similar position is taken by Brown and Hippisley (2012) where morphology in general is considered to constitute a network. Crucially, the principle of degree of connectivity is augmented by the addition of the lexical neighborhood. At the same time, the traditional pairs are maintained, constituting a cross-paradigmatic relation which then constitutes the source-oriented schema as was argued in Section 2.2.2. Effectively, these factors enable us to posit a global comparison between different verb pairs that is not delimited artificially to minimal and local configurations.

The basic idea of a network model brings constructionist models closer to a larger body of different frameworks both in computational linguistics and computational psycholinguistics. Bybee's (1985) network model for morphology is implicated in connectionist models (Macwhinney, 2001; Marcus, 2001) and to probabilistic models for the mental lexicon on morphology, such as Hay (2001), Bresnan et al. (2007), Baayen and Moscoso del Prado Martín (2005), and Bod (2006). Importantly, the principles of the network model are related to studies on the mental lexicon outside of morphology, such as studies on phonological structures in Pierrehumbert (2003) and phonological densities in Vitevitch (2008). These model relations in the lexicon based on word associations, sense relations extracted from WordNet and Roget's thesaurus in Steyvers and Tenenbaum (2005). The previously mentioned studies are only a small fraction, and certainly all of them do not ascribe to Construction Grammar. However, they share the principles of the network model, a move away from dichotomies to interconnected structures.

Sections 3.1.1–3.1.4 are used to model the degree of connectivity between lexical verbs by introducing the concept of the lexical neighborhood. The strength of connecitvity of lexical verbs is model with frequency of use discussed in Sections 3.1.6 and 3.1.7. Additionally, a new variable is introduced in Sections 3.1.8 and 3.1.9 labeled as the Constructional Entropy. It is used to model the strenght of connectivity between lexival verbs and the argument constructions. Another important semantic component holding between the cross-paradigmatic relation is causation. This variable is described in Section 3.1.10 along with transitivity. Finally, the theoretical basis of the semantic components of verbs covering the paradigmatic relation (aspect, tense, and grammatical functions) are presented throughout this chapter.

3.1.1 Lexical Networks: Reflexiva Tantum and Neighbor Verbs

The status of the reflexiva tantum verbs is a crucial issue to any account on the Russian Reflexive Marker. This refers to verbs that do not have a corresponding non-reflexive verb. This is already noted by Geniušienė (1987:145). Typically, these verbs are simply excluded a priori without any further explications (cf. Fehrmann et al., 2010; Guhl, 2010; Князев, 2007). Thus, the status of reflexiva tantum is the primary data trimming parameter in accounts operating on pairs. Jakobson (1989 [1932]:4) simply states that the reflexiva tantum are non-paired marked forms. The only way to account for these verbs would be either to postulate an unmarked form, which would have been lost due to diachronic changes, or to create a separate set of descriptive devices specifically for the reflexiva tantum verbs. The former is not supported by diachronic evidence. The latter position would be undesirable as it would introduce a more complex account operating on two different systems. Thus, the theoretical adequacy of the account would be questionable. Nonetheless, in recent comprehensive taxonomies of the Russian Reflexive Marker, as in Gerritsen (1990) and Israeli (1997), the reflexiva tantum verbs are included although it is claimed that the reflexive verbs are described in terms of pairs, leaving the reflexiva tantum verbs without any theoretical basis. They are included simply for the sake of completeness.

The standard assumption behind the development of the reflexiva tantum verbs is stipulated in terms of independent evolution unrelated to the base verb (Недялков, 1971:14). However, Kuznecova (Кузнецова, М. В., 1984:64-65) identifies three pathways based on the diachronic development of the reflexiva tantum verbs:

1) Verbs pertaining to the category of Reflexiva tantum only in contemporary Russian, for example *бороться* 'fight' ~ *, *каяться* 'repent' ~ *, and *трудиться* 'work' ~ *.
2) Verbs which are only reflexiva tantum, both diachronically and synchronically, for example *бояться* 'be afraid' ~ *, *гордиться* 'be proud'~ * and *надеяться* 'hope' ~ *.
3) Verbs pertaining diachronically to the category of reflexiva tantum but have a neighbor verb in contemporary Russian, for example *беситься* 'rage' ~ *бесить* 'enrage' , *печалиться* 'mourn' ~ *печалить* 'grieve' and *ругаться* 'swear, curse' ~ *ругать* 'scold, swear.'

Kuznecova's study shows that the assumed derivational directionality from the non-reflexive (base form) to the reflexive verb is weak at best. Lavidas and Papangeli establish similar diachronic mismatches concerning reflexiva tantum verbs in Greek.[41] Based on the mismatches, they go even further by questioning the basis of an invariant meaning to motivate the diachronic development of the Reflexive Marker (Lavidas & Papangeli, 2007).

It is appropriate to agree with Lavida and Papangeli on the status of the invariant meaning of the Reflexive Marker. First, the invariant meaning would have to be mapped in some manner to motivate the loss of the cross-paradigmatic relation between *бороть* 'fight' and *бороться* 'fight,' where *бороть* 'fight' has become archaic in contemporary Russian. Second, the same invariant meaning would have to motivate the emergence of *ругать* 'swear, curse' with *ругаться* 'swear, curse.' Finally, the invariant meaning would have to motivate the maintenance of the reflexiva tantum verbs such as *бояться* 'be afraid.' Thus, an invariant meaning would have to model both directions of changes and also the maintenance of gaps in the cross-paradigmatic relations. Thus, a model is required that can incorporate these kinds of deviations within the category of the Reflexive Marker.

Geniušienė is perhaps the only one to establish definitions in determining the status of pairs labeled as non-reversability. She proposes four types of

[41] They use the term deponent.

derivational (ir)regularities in the formation of the reflexive verbs: 1) morphological non-reversability is the absence of the base form, 2) syntactic non-reversability is an irregular change in the argument structure, 3) lexical non-reversability manifests itself in restricted lexical properties, such as the Latvian non-reflexive verb *meklēt* ‘look for’ versus *meklētie-s* ‘be on heat’, and 4) semantic non-reversability shows a non-related meaning. Additionally, she connects the formation of semantic reflexiva tantum verbs to metaphor and metonymy (Geniušienė, 1987:145-149).

The first criterion, morphological non-reversability, is the canonical case where the reflexive verb lacks the pair. However, reflexive verbs are also intertwined with prefixes, (e.g., such prefix combinations as *в*-Verb-*ся*, *воз*-Verb-*ся* and *раз*-Verb-*ся*). A comprehensive list is given by Vinogradov (Виноградов, 1972:507-508; cf. Янко-Триницкая, 1962:35). The second combination is rare in contemporary Russian being a Church Slavonic loan. Typically, prefixation is accompanied with the loss of correspondence although the base form may still be shared, for example, *звонить* ‘ring, call’ ~ *звониться* ‘ring, call,’ and * ~ *дозвониться* ‘reach by phone,’ or *слушать* ‘listen, obey’ ~ *слушаться* ‘obey, listen,’ and * ~ *наслушаться* ’hear enough, listen for a long time.’ Another complex formation preserves the correspondence of the form but the reflexive verb has acquired an opposite meaning as in *учить* ‘learn’ ~ *разучить* ‘learn gradually’, and * ~ *разучиться* ‘be out of practice.’

Prefixation is a serious issue for pair-driven models whereas construction-based models can motivate these patterns. The lack of the cross-paradigmatic relation constitutes only one of the properties of verbs. Furthermore, prefixation can be considered as a type of construction as it introduces both the form and the meaning, for example ⟨раз|*Verb*|ся⟩ and ⟨на|*Verb*|ся⟩ (cf. Booij, 2010:42-45). Prefixation augments the semantic structure of the verb similar to the semantic component of manner as was already outlined in Section 2.2.3. The schematic Argument Construction is primarily connected to who-did-what-to-whom. Thus, these kinds of mismatches are difficult to motivate in the pair account because the corresponing pair is missing, although the base form such as *звонить* ‘ring, call’ is shared with the verb *дозвониться* ‘reach by phone.’

The second criterion, syntactic non-reversability, is highly problematic. The definition preludes that non-reversability is defined as a violation to the established stock of the diathesis alternations, thus leaving the status of cases which have not been systematically explored open, like the combinations with infinitive. Certain reflexive verbs can combine with an infinitive but this pattern is not necessarily available for the non-reflexive verbs such as *готовить* ‘prepare’ ~ *готовиться* ‘prepare’ and *собрать* ‘gather’ ~ *собраться* ‘gather, intend.’ Thus, the issue with these and similar verbs relates to the question whether the Nominative-Infinitive pattern constitutes a form of regular derivation or not in the diathesis tradition. Another difficulty with these patterns is that the neighbor verb does not necessarily combine with the Nominative-Infinitive pattern. That is, it cannot be derived from the syntactic structure of the neighbor verb. Another difficulty with these patterns is that the neighbor verb does not necessarily combine with the Nominative-Infinitive pattern, that is it cannot be derived from the structure of the neighbor verb. In contrast, certain verbs can combine with the Nominative-Infinitive pattern such as *задумать* 'plan' ~ *задуматься* 'plan.' In terms of the network model, these mismatches are straighforwardly covered with the two types of schemata. The latter type can be considered as pertaining to the source-oriented schema, whereas the former to the product-oriented one as was outlined in Section 2.2.2. In my view, the third and the fourth criteria, lexical and semantic non-reversability, are nearly identical, leaving the semantic similarity once again as the source of the definition in the pair account.

Recently, Kalashnikova and Saj have criticized how the definition of shared meaning has been used to define the pairs. They exemplify the issue with the pair *оправдывать* 'justify, explain' ~ *оправдываться* 'explain oneself.' According to their analysis, certain senses are overlapping and approximate the Semantic Reflexive type, but the non-reflexive verb has idiosyncratic senses which are not attested with the reflexive verb (Калашникова & Сай, 2006:3). An example would be 'generation of speech.'. This same issue is also present with the "pair" *собрать* 'gather' ~ *собраться* 'gather, intend' where only the reflexive verb also contains the 'intend' sense. This raises a serious question about the descriptive practices. Typically, multiple patterns are not considered, masking the potential mismatches between the postulated pairs. If semantic regularity was one of the defining features of the pairs, one would assume a one-to-one correspondence between them.

In my view, the issue raised by Kalashnikova and Saj is, nonetheless, expected if one operates on derivational rules. Only certain parts of the base form need to be used in the derivation. Thus, one would expect that the base form has a more elaborated semantic structure. As WordNet and FrameNet are not available for Russian, a gold standard data source to contrast the exact number of senses for verbs is not possible. Certainly, a dictionary-based perspective could be taken. For the above mentioned pair, *оправдывать* 'justify, explain' ~ *оправдываться* 'explain oneself,' five senses are given by Kuznecov (Кузнецов, 2009 [1998]) and two for the reflexive verb. However, this would only serve as an approximation, as dictionaries can be inconsistent in describing sense structures.

Nonetheless, it is worth pointing out that this same issue is raised already by Janko-Trinickaya connected to metaphorical extensions. An example would be *кипятиться* 'boil' → 'worry.' The latter sense is not part of the semantic structure of the non-reflexive verb *кипятить* 'boil' (Янко-Триницкая, 1962:23). This example goes in the opposite direction where the reflexive verb has acquired a different sense in comparison to the non-reflexive verb. Based on the semantic criteria proposed by Geniušienė, this constitutes a mismatch between the pairs leading non-reversability, although in other usage patterns the senses might overlap. The reflexive verbs *готовить* 'prepare' ~ *готовиться* 'prepare' and *собрать* 'gather' ~ *собраться* 'gather, intend' used to illustrate the Nominative-Infinitive pattern display similar behavior. The Nominative-Infinitive pattern creates a mismatch in relation to the neighbor verb.

In the pair account, an implicit assumption is made that a certain sense functions as the point of reference. When multiple patterns are excluded, the point of reference is a priori selected, hence potential mismatches are masked. The verbs *кипятить* ~ *кипятиться* would constitute a pair in the 'boil' -sense contrary to *кипятиться* in the 'worry' -sense that cannot be reversed back to the non-reflexive verb. Thus, the status of a verb would be dependent on the selected sense, an unintended outcome in the pair account. These mismatches or gaps might be brought "in-line" by evoking metonymy and metaphor as part of the derivation (cf. Князев, 2007; Падучева, 2004). However, even this would leave open the exact status of the lexical items in question, especially in formal approaches where a sharp division between the lexicon and the grammar is postulated.

In order to model the lexical structure of the reflexive verbs from the constructionist perspective, the concept of Neighbor Verb is proposed that straightforwardly motivates both the possible mismatches and also incorporates the traditional pairs. The verbs in the verb-specific constructions are taken as the starting point, constituting a paradigmatic relation. All verbs marked with the Reflexive Marker are considered to belong to the paradigm of the Reflexive Marker. A cross-paradigmatic relation exists between the reflexive and the non-reflexive verb if they share the same phonological form excluding the Reflexive Marker. This is the definition of Neighbor Verb used in this study. This definition follows the tenet outlined by Bybee (1985:117-129) in order to model the degree of connectivity between lexical items: phonological and semantic similarity (cf. Booij, 2010). Thus, the semantic connectivity becomes a property of the cross-paradigmatic relation between verbs (cf. Gerritsen, 1990:23-24). An example is *вить* 'cause to curl' and *виться* 'curl', and *оказать* 'render' and *оказаться* 'seem, appear.' Both of the reflexive verbs have a neighbor verb in Russian, but differ in the degree of semantic connectivity, the latter being semantically dissimilar to its neighbor. The operationalization of the semantic similarity as the semantic degree of connectivity is discussed in Section 3.1.4.

The concept of the Neighbor Verb allows one to model lexical items at various levels of degree and avoids postulating different models for the reflexive verbs that do not confine to the cross-paradigmatic relation. Additionally, the degree of connectivity also enables to motivate, for example, diachronic changes as the model assumes variability over dichotomy. At the same time, the theoretical basis of the cross-paradigmatic relation is faced with a question about what counts as evidence, specifically what counts as evidence in order to state that a certain verb does not have a neighbor verb in contemporary Russian once we move away from such verbs as *бояться* 'be afraid,' *гордиться* 'be proud' and *надеяться* 'hope.' A dictionary-based perspective was taken in this study. The primary sources were Kuznecov (Кузнецов, 2009 [1998]) and Efremova (Ефремова, 2000). Additionally, two-volume word-formation dictionary by Tikhonov (Тихонов, 1985b; a) and Shirshov (Ширшов, 2004) were consulted. Based on the entries in the dictionaries, 717 neighbor verbs were established for the unique reflexive verbs (*n* = 819). Thus, 102 reflexive verbs lack a neighbor verb in the database indicating a strong connectivity across paradigms in Russian.[42]

The other possibility would be to explore the extensibility of a particular verb form, such as using a large database like Integrum to check whether a particular verb form is extended. For example, *корениться* 'root,' *изловчиться* 'contrive,' *спохватиться* 'realize,' and *выситься* 'lift one's head' are classified as lacking the cross-paradigmatic relation. However, *коренить* and *высить* are given in Dal' (Даль, 1903-1909) which is a famous dictionary containing a rich source of dialectal forms, particularly from the 19th century. Because the proposed model is probabilistic, it is expected that some of the reflexive verbs can be extended in usage to have a neighbor verb. The Reflexive Marker is productive in Russian as discussed by Saj (Сай, 2007) and Norman (Норман, 2004). The prediction is that if these forms are extended, they should be extremely infrequent.

As a technical note, for the reflexive verbs that lack the cross-paradigmatic relation, a dummy neighbor verb "None" is used, which avoids missing values in the model input. From a linguistic perspective, the usage of the dummy verb "None" encoding schema presupposes that speakers are sensitive to gaps, in that they are aware either consciously or subconsciously if a particular verb is or can be extended (cf. Boyd & Goldberg, 2011; Goldberg, 2011).

In sum, this section has taken the first step in order to model the gradient structure of verbs across paradigms by introducing the concept of Neighbor Verb which is defined as a shared phonological similarity in the cross-paradigmatic relation. Thus, a reflexive verb has a neighbor verb if and only if the phonological difference is found in the Reflexive Marker. The concept is a prerequisite to account for the fact that the cross-paradigmatic connectivity between items varies in two ways: 1) the degree of connectivity and for establishing lexical networks, and 2) the degree of interconnectedness of verbs in and across paradigms. Subsequent sections continue on formulating the gradient structure in quantifiable terms.

3.1.2 Lexical Networks: Hypothesis of Connectivity

Bybee postulates two schemata for linguistic structure. They are source-oriented and product-oriented schemata cf. Sectio 2.2.2. These can be viewed as functioning in concordance within a complex category. Bybee has proposed that the strength of the connection is a function of semantic and phonological similarity, although the former is claimed to be much stronger (Bybee, 1985; 1995; 2010). Bybee's schemata rest on the idea that linguistic items form larger structures than pairs governed by rules. The principle of the lexical network has become in recent years an active research paradigm (Altieri et al., 2010; Baayen & Moscoso del Prado Martín, 2005; Chan & Vitevitch, 2009; Geeraert &

[42] Nonetheless, the definition can be considered to be biased towards standard contemporary Russian (cf. Крысин, 2007a).

Kyröläinen, in prep.; Steyvers & Tenenbaum, 2005; Vitevitch, 2008).

Bybee's (1985; 2010) position is that morphological relations, for instace, emerge from relations among words based on semantic and phonological similarity. A partial network model of the lexical structure of *unbelievable* is illustrated in Figure 3.1-1 based on morphophonological connections, adapted after Bybee (2010:23).

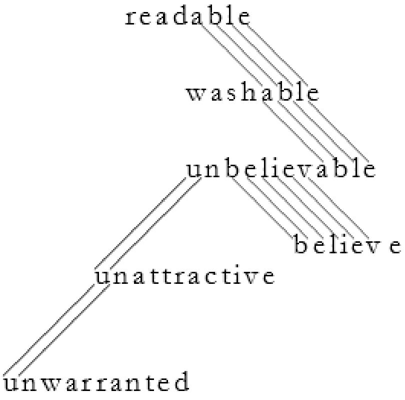


Figure 3.1-1 Internal structure of *unbelievable* in relation to other words, adapted from Bybee (2010:23).

The principle of the network is to bring forth the interconnectedness of lexical items. The phonological connection between items is given with lines. The more connecting lines there are between items, the stronger the connection between them. Thus, *unbelievable* is not only connected to *readable* and *washable* through the suffix *–able,* but also to *unattractive* and *unwarranted* through the prefix *un-*, albeit a weaker connection. The most important aspect of the lexical network model is that it inherently operates on words and, morpholofical connections are abstractions over words (cf. Booij, 2010).

In terms of modeling the mental lexicon, the network model is commonly referred to as neighborhood density in (computational) psycholinguistics. The standard approach to defining the neighborhood density is called Coltheart's *N*. The density estimation based on this measure is a simple operation defined as the number of items of equal length which are produced by changing a single letter or phoneme, for example, *boat* → *float* and *cat* → *bat* (Colthearth, Davelaar, Jonasson & Besner, 1977). According to Yarkoni et al. (2008:971), it is cited nearly 600 times based on ISI Web of Science in 2007. Recently, Pastizzo and Feldman contrast form-meaning (*boat-float*), shared-meaning (*swim-float*) and shared-form (*coat-float*) pairings. They show that the strength of mapping is facilitated for both morphologically unrelated and related items (Pastizzo & Feldman, 2009).

Although the literature on the effects of neighborhood density is substantial, the number of studies on Slavic and, specifically on Russian, is minimal.

Dąbrowska has shown a facilitatory effect for nonce nouns, which populate a high-density neighborhood in Polish. People are more likely to apply dative case ending to nonce nouns when they share a high number of phonologically related items (Dąbrowska, 2008). Kazanina also reports a facilitatory effect on neighborhood density in Russian for prefixed nouns in masked-priming experiments. She found a facilitatory effect for morphologically related prime-target pairs. Examples are *рост* 'growth' and *нарост* 'outgrowth,' and for pseudo-related prime-target pairs, *тон* 'tone' and *притон* 'den,' The *при* is an existing prefix in Russian but the item *притон* is a monomorphemic noun (Kazanina, 2011).[43] The later findings follow the basic tenet of the lexical network model where items are interconnected.

Although neighborhood density has been incorporated in a number of studies, contradictory results in terms of effect, whether inhibitory or facilitatory, are present (cf. Balota, Cortese, Sergent-Marshall, Spieler & Yap, 2004; Ziegler & Conrad, 1998). One possible motivation for the apparent differences in terms of effect is discussed by Yarkoni et al., setting task- and design-specific considerations aside (Balota et al., 2004). They connect the possible differences in the stages of processing, such as global and local similarity. Low-frequency items may benefit from global similarity, leading to increased early processing. For example, *stab* shares such loosely connected items as *station*, *table,* and *stack*. At the local level, a low-frequency item may have a high number of competing items leading to increased processing time. For example, *stab* has tightly connected items like *star* and *slab*, especially if the items in question have a fairly high frequency (Yarkoni & Balota, 2008:977). This highlights the possibility that there is a degree in connectivity between items interacting with frequency and depends on the structure of the lexical item in question.

In addition to behavioral studies, neurolinguistics studies have also been conducted offering converging evidence for neighborhood density. Holcomb et al. provide neurolinguistic evidence for the effects of neighborhood densities in processing linguistic items in visual word perception in English using the event-related potential (ERP) paradigm. Their data show that the ERP component N400 is present based on two experiments (Holcomb, Grainger & O'Rourke, 2002). Furthermore, Laszlo and Federmeier (2009) have shown that the neighborhood effect is present in sentence processing in English based on the ERP paradigm. These findings are significant. The ERP component N400 is a distinctive negative brain electrical activity occurring around 400 milliseconds and has been demonstrated to be sensitive to early lexico-semantic processing (Kutas & Hillyard, 1980). Although the estimations of neighborhood densities are typically established through graphemic operations, the operationalization of the density appears to target semantic processing, also. The presence of the neighborhood effects offers evidence to the lexical network model. Processing

[43] Kazanina does not provide information about the procedure of establishing the neighborhood density.

of an item does not only activate the item in question, but also leads to the partial activation of the connected items.

Additionally, lexical densities are not simply connected to processing and modeling form-based relation. Baayen and Moscoso del Prado Martín (2005) have shown that irregular past tense verbs in English, German, and Dutch have more neighbors that are irregular and more senses (estimated based on WordNet). They offer evidence that irregular verbs have a higher semantic density compared to the regular ones. The hypothesis is supported by corpus, behavioral, and brain imaging data (Baayen & Moscoso del Prado Martín, 2005). At the same time, neighborhood densities are constructed based on some corpora or dictionary. The discussion of the possible direction of the effect, whether inhibitory or facilitatory, is beyond the scope of corpus-based studies. Thus, this matter is not discussed any further and is left for future experimental studies.[44]

Before turning the process of constructing the neighborhood density of reflexive and the neighbor verbs, it is worth noting that the whole concept is intertwined with theoretical and methodological issues. Furthermore, a practical component is also involved as the resources available for English are enormous in compared to Russian. The first issue is the exact structure of the estimated lexical densities; either we opt for a loose or tight solution.

The standard definition of the neighborhood density based on Coltheart's *N* has several limiting factors. First, it is a binary measure. Items are either neighbors or not although it is fairly uncontroversial to state that similarity is a matter of degree. Second, the measure only considers a single operation but longer items require more operations to establish neighbors (Yarkoni & Balota, 2008). Thus, a looser definition of neighborhood density is used in this study, as the reflexive verbs are based on a random sample covering morphologically complex forms. Additionally, the looser density estimation also incorporates the tighter one by definition.

It will be demonstrated in Section 3.1.3 that information is not lost with the looser definition of the neighborhood because the within variation of the lexical

[44] Intuitively, Bybee's hypothesis of the stronger connection based on semantics would seem accurate. However, the situation is far from the hypothesis. Smolka et al. offer an extensive discussion on the matter based on the existing body of experimental research. It appears that Hebrew and Arabic are counter-examples to Bybee's hypothesis. Similarly, Smolka et al. show compelling evidence based on German verbs that morphological relatedness is stronger compared to semantic contrasting English, which displays a stronger effect of semantic relatedness. Smolka et al. partly connect the contradictory results to the morphological structure of the languages. Roots play a pivotal role in Hebrew and Arabic, separating them from the poor morphological structure of English. In contrast, German can be viewed as occupying an intermediary position (Smolka, Komlósi & Rösler, 2009:338-340, 341, 366-367). Russian might display similar behavior to German, as affixation plays a central role in forming aspect and the Reflexive Marker constructions.

neighborhood can be straightforwardly modeled. In addition to operations on units, such as a letter change, Bybee and Moder (1983), and Ramscar (2002) have shown that nonce verbs can be inflected irregularly if they have a similar phonological structure to irregular verbs in English. Pinker (1999:137-148) also considers that even regular verbs that rhyme with irregular ones, like *blink* with *drink* and *stink* and *glide* with *ride* and *stride*, may be more amiable to be used with a novel irregular form. Ziegler and Conrad (1998) have shown that most neighbors are also (orthographic) rhyme pairs in English and facilitate processing. Similarly, Strokel (2002) has demonstrated that rhymes influence the development of the mental lexicon in language acquisition in English. These findings suggest that even larger phonological units, such as rimes, may be part of a lexical network in addition to defining neighborhoods as single unit changes. For the purposes of the present study, the neighborhood densities for the reflexive and the neighbor verbs were constructed as a rhyme space because a large electronic dictionary is available for Russian, titled the Russian Rhyme dictionary (Rhymes, 2011). The dictionary is fairly large containing over 102,000 words and over 3.8 million word forms.[45]

Another methodological issue is the question of the selected inflectional form, as verbs have rich inflectional paradigms in Russian. The selected inflectional form influences the structure of the neighborhood. Certain reflexive verbs have a defective inflectional paradigm, (e.g., *довестись* 'happen,' *захотеться* 'want' and *мечтаться* 'dream.' Defining the rhymes based on the first person, for instance, would create a situation where these verbs would not have any neighbors. Also tense would influence the estimated density. Thus, infinitive form appears to be the most neutral options. The neighborhood of a particular verb is defined based on the infinitive form and with exact match and maximal syllable structure. Additionally, only verb forms were considered excluding for instance, nouns. The Neighborhood Density was constructed for the reflexive and the neighbor verbs in the database.

The estimated Neighborhood Density of a particular verb is the total number of its rhyme verbs. This is the shared phonological body, rime. To illustrate the structure of the estimated densities, the Neighborhood Density (65) of the verb *бояться* [ба^{ъ}јац:а^{ъ}] 'be afraid' contains such rhymes as *устояться* [уста^{ъ}јац:а^{ъ}] 'settle' and *смеяться* [см'и^{э}јац:а^{ъ}] 'laugh.'[46] The phonological transcription of the verbs used in the Rhyme dictionary is given in square brackets. The shared commonality between the verb forms is the rime *-яться*. Thus, the previously used informal discussion of a loose neighborhood is defined and operationalized precisely now as a maximal rhyme space in Russian.

The Neighborhood Density of the verbs is fairly large containing 378, 371 verb forms pairs in total (177, 314 form pairs for the reflexive verbs, and 201,

[45] The lexical structure available for native speakers of Russian might be overestimated due to the large size of the dictionary. This issue is discussed, for instance, by Vitevitch (2008:409-410) in relation to English.

[46] The reflexive verb has also inflected noun neighbors such as *тунеядца* 'parasite.'

057 for the neighbor verbs). The rhyme verbs can be shared across verbs, highlighting the fact that the lexicon is highly interconnected rather than consisting of predefined pairs. The interconnectedness can be brought to light by extracting the unique verbs. Only 6% (n = 11901) of the total density of the neighbor verbs is covered by the unique verb forms. A similar distributional property is present for the reflexive verbs such as 6% (*n* = 10421). The distributional properties of these verbs are visualized with a violin plot in Figure 3.1-2.

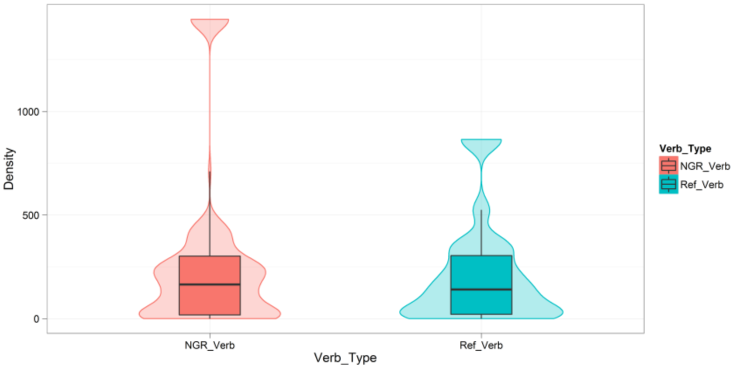


Figure 3.1-2 Violin plot of the Neighborhood Density of the neighbor (NGR) (*n* = 717) and reflexive (Ref) (*n* = 819) verbs with bandwidth 0.6.

A violin plot is a visual and explanatory representation of the distributional properties of a variable, combining both a density and box plot. The estimated density function is visually represented alongside of the box plot on both sides. A density plot is an estimation of the underlying probability density function of a variable. The degree of smoothness of the estimated density is controlled with bandwidth. Larger values may overestimate the underlying density function whereas lower values may underestimate it (Fox & Weisberg, 2011:110-110). The bandwidth was adjusted from the default 1 to 0.6 to decrease the smoothness to highlight the distributional differences.

A box plot is a visual summary representation of a numeric variable conditioned by a categorical variable. The box plot consists of the following quantities: the lower hinge is the first quartile (25% of data points and less) and the upper hinge is the third quartile (75% of data points and greater). The lower and the upper whiskers are the minimum and the maximum values still within the interquartile range (IQR).[47] Values outside this range are potentially considered to be outliers. The outliers are suppressed in Figure 3.1-2, but the estimated density covers them. The median, midhinge (50%), is shown visually

[47] There is a fairly large number of definitions of IQR (Hyndman & Fan, 1996). GGPLOT2 uses the 1.5 rule: the lower whisker as (first quartile -1.5 * IQR) and the upper whisker as (third quartile +1.5 * IQR) where IQR is (the third quartile – the first quartile).

with the band inside the box. The y-axis is given in counts (Figure 3.1-2).

The distribution shows that both types of the verbs have items which do not have any neighbors, such as *взять* 'take' and *выбраться* 'get out,' due to their phonological pole. There are 20 neighbor verbs which do not have a rhyme neighbor, whereas the reflexive verbs have 13. Zero was added for these verbs. Similarly, zero was added to those reflexive verbs that lack the cross-paradigmatic relation.

The data show that the maximum neighborhood density is considerably larger for the neighbor verbs (*n* = 1,445) compared to the reflexive verbs (*n* = 865). The maximum densities are attested with the rimes -*вать* [ват'] and -*ваться* [вац:а^{ъ}]. From a morphological perspective, they are part of such suffixes as -*ов*- and -*ива*-. Both suffixes are productive in Russian. In contrast, the median values indicate that on average the densities between these two types of verbs appear to be in close proximity.

Another important aspect related to the Neighborhood Density is the potential difference between the verbs. They are connected through the cross-paradigmatic relation (*n* = 717) whether the distributions are different for the reflexive verbs (min. = 0, median = 141 and max. = 865) and for the neighbor verbs (min. = 0, median = 165 and max. = 1,445). A two-tailed Wilcoxon rank-sum test was used because the test is less sensitive to extreme values.[48] The difference in the distributions is not statistically significant (*W* = 244488.5, *p*-value = 0.1092). Following previous studies on the effects of semantic attraction based on neighborhoods (Baayen & Moscoso del Prado Martín, 2005; Bybee & Moder, 1983; Ramscar, 2002), the results indicate that the verbs connected through the cross-paradigmatic relation reside in similarly dense neighborhoods. On one hand, the results are unexpected, as the reflexive verbs are typically described being highly lexicalized. On the other, considering that the reflexive verbs are fused with a high number of categories associated with verbs in Russian, the results are actually expected following the tenets of the network model. A type that is associated with a dense (semantic) structure should also be less restricted in terms of categories available in a language.

To motivate the distribution, a domain-general principle labeled as the Hypothesis of Connectivity is formulated. All things being equal, according to the hypothesis, the connections between items increase over time. The temporal dimension is included in order to account for the dynamic aspect of language, leading to a testable hypothesis in diachronic or language acquisition studies. Without the temporal dimension, the hypothesis amounts to a difference in distributions. As Vitevitch has argued that although lexical growth is typically associated with language acquisition, it is hardly controversial to claim that people acquire new lexical items. Furthermore, growth may be connected to preferential attachment. It is more probable that a new item is attached to an item which is already highly connected in a network (cf. Steyvers & Tenenbaum,

[48] A two-tailed test does not assume directionality and tests whether the two distributions are different.

2005:66-68, 72-73; Vitevitch, 2008:415-416). Generally, these principles are an inherent property of the network model based on the seminal work by Barabási and Albert (1999) on self-organizing systems.

To motivate the outlined principles and the temporal dimension, all the English monomorphemic verbs (N = 3,463) available in the English Lexicon Project were extracted along with the estimated number of their phonological neighbors, defined as a one-phoneme difference and the estimated frequency based on subtitles (Balota, Yap, Cortese, Hutchison, Kessler, Loftis, Neely, Nelson, Simpson & Treiman, 2007). Additionally, age-of-acquisition norms were extracted from the mega study by Kuperman et al. (in pres.). Missing values were removed yielding a final dataset of 3085 monomorphemic verbs. The data revealed a statistically extremely significant moderate positive correlation between the log frequency and the phonological neighborhood density: $r(3083) = 0.336$, $p < 0.001$. Similarly, the data showed a statistically significant and moderate negative correlation between the phonological neighborhood density and the age-of-acquisition norms (mean AoA in years): $r(3083) = -0.3244$, $p < 0.001$.[49]

This section introduced the basic theoretical basis of the lexical network model based on rhyme densities. The results indicate that both the reflexive and the neighbor verbs that are connected through a cross-paradigmatic relation, occupy similarly dense neighborhoods. The Hypothesis of Connectivity was formulated to account for the obtained structure. This follows the hypothesis outlined by Bybee (1985:49) in that paradigms have internal structure. Thus, the Neighborhood Density is one property of this structure that influences the formation of complex categories. At the same time, the definition of the Neighborhood Density was formulated to include loosely connected items. Additional theoretical devices are required to bring forth the degree of connectivity within these neighborhoods.

3.1.3 Lexical Networks: Hypothesis of Distance

The previous section introduced the concept of neighborhood density for modeling lexical networks in language. Recently, studies have emerged that incorporate the degree of connectivity within lexical networks. Chan and Vitevitch (2009) have shown that phonological neighbors of an item that are also neighbors to each other are processed slower in an auditory lexical decision task. Such items as *badge* and *log* have the same number of phonological neighbors (n = 13) but the neighbors of *badge* are also more likely to be neighbors to each other compared to *log*. Geeraert and Kyröläinen replicated this result for English irregular verbs (N = 120) in reading using eye-tracking.

The total eye fixation durations were longer for irregular verbs, which also have phonological neighbors that are neighbors to each other, indicating

[49] Brysbaert and New (2009) have shown that frequency information estimated based on subtitles offers, by far, a better fit to lexical decision data compared to a sample complied on written texts.

inhibition in processing while reading. Importantly, Geeraert and Kyröläinen offer tentative support for the basic principle of the network model, namely the degree of connectivity. Additionally, irregular verbs that have phonological neighbors, but their neighborhood is not connected to other lexical neighborhoods of the irregular verbs, had an inhibitory effect on processing based on total eye fixation duration, such as *drive*, *eat,* and *grind* (Geeraert & Kyröläinen, in prep.). This finding offers tentative support for the network model. A global degree of connectivity ensures that information can spread across the network and deviations from this lead inhibition in processing times. These findings indicate that the network model is able to capture, at least, some part of the mental lexicon in a language.

For purposes of the present study, the lexical density of the reflexive and neighbor verbs was estimated based on rhyme pairs. The inner structure of the lexical densities was defined informally in the previous section as containing loosely connected items. Yarkoni et al. have operationalized the loosely connected lexical items with the Levenshtein distance. Their study shows that the Levenshtein distance outperforms the standard Coltheart's *N* measure in lexical decision and pronunciation performance in three large-scale data sets. Levenshtein distance contains a range of properties, which makes it attractive over the standard measure and the properties smoothly mesh with the assumption made, (e.g., in Cognitive Linguistics and Construction Grammar). First, it is a continuous measure. Items can be either more or less closely related to each other, which are known as the principle of continuum in categories. Second, graphemic operations other than substitution can also create closely related items, such as insertion and deletion, for example *widow* → *window* (insertion → *n*) and *planet* → *plane* (deletion → *t*) (Yarkoni & Balota, 2008:971).

Levenshtein distance, a standard string metric in computer science, is also referred to as edit distance proposed by Levenshtein (Левенштейн, 1966). The distance measure has a wide array of applications from spell checking to DNA sequence analysis. An extensive coverage of applications is given, by Kruskal (1983). Levenshtein distance has also been used to analyze distances in linguistics data. Palunčić et al. (2009) analyzed Iranian language family and distances from the reconstructed Old Iranian, and Gooskens and Heeringa (2004) analyzed Norwegian dialectal data.

The Levenshtein distance is the number of string operations, deletion, insertion, or substitution required to transform one string into another. The Levenshtein distance was calculated for the reflexive and neighbor verbs as a graphemic pairwise distance within the neighborhood of the verb, $D(x, \acute{x})$. The costs of the string operations were kept the same, (i.e. 1). The pair-wise distance of the reflexive verb *автоматизироваться* ‘automate’ within its neighborhood (94) was calculated as:

$D(x, \acute{x})$ = автоматизироваться → анестезироваться ‘anesthetize’ = 6

$\vdots$

$D(x,\ \acute{x})$ = автоматизироваться → архаизироваться ‘use archaisms’ = 5

$\overline{D}(x, \acute{x})$ = 5.936

After the pairwise distances were calculated, the average (rounded to three decimal places) was used to represent the internal lexical structure of the verbs within its Neighborhood Density. For instance, the average neighborhood distance of the verb *автоматизироваться* is 5.936. Thus, the reflexive verb requires proximately six graphemic operations on average to arrive at its rhyme neighbors. Figure 3.1-3 visualizes the distributions of the Neighborhood Distance.

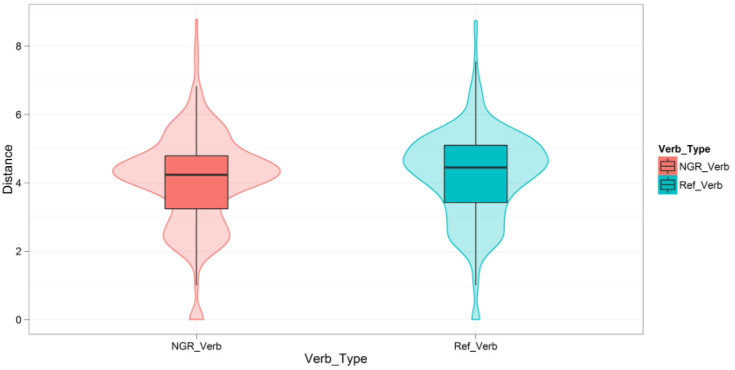


Figure 3.1-3 Violin plot of the Neighborhood Distance of the neighbor (NGR) (*n* = 717) and reflexive (Ref) (*n* = 819) verbs with bandwidth of 0.6.

The results suggest that the reflexive verbs appear to have slightly higher median distances (4.45) than the neighbor verbs (4.235) in terms of the average neighborhood distance. Additionally, the density plots indicate that a larger portion of the neighbor verbs are located around the median in comparison with the reflexive verbs.

The previous section demonstrated that the Neighborhood Density between the cross-paradigmatic verbs was not statistically significant. The degree of connectivity can be used to test whether these densities differ in distances across the paradigms of the reflexive verbs (min. = 0, median = 4.441 and max. = 8.737) and the neighbor verbs (min. = 0, median = 4.235 and max. = 8.778). A two-tailed Wilcoxon rank-sum test was used to compare the distributions based on the cross-paradigmatic verbs (*n* = 717); the difference was statistically significant (*W* = 276519.5, *p*-value = 0.013). The average neighborhood distances are greater for the reflexive verbs in comparison with their cross-paradigmatic neighbor verbs.

Assuming an iconic relation that the average neighborhood distance, at least, partly reflects the properties of a semantic distance (cf. Haiman, John, 1983), the results indicate that the average semantic distance is greater for the reflexive verbs compared to their cross-paradigmatic neighbor verbs, but the commonality is found in the similar densities. Following the principles of the network model, we would assume that the reflexive verbs display a greater degree of idiosyncratic properties compared to the neighbor verbs. One would expect to find more locally constrained verb-specific constructions. Argument structures should lean towards being more lexically specified rather than being highly schematic on average for the reflexive verbs. The opposite should hold for the neighbor verbs (cf. Geniušienė, 1987; Цейтлин, 1978:193-195; Янко-Триницкая, 1962).

The structure of the neighbor verbs is supported by a dense lexical network with shorter distances. Their properties follow these global tendencies. The vast majority of them pertain to the Transitive Construction (n = 527) in Russian. Thus, these structural properties amount to the Hypothesis of Distance according to which the distances between items decrease over time all things being equal. Because the distances are greater the connectivity between items is looser amounting to a stronger item-specificity and, a locality effect (cf. Baayen & Moscoso del Prado Martín, 2005:668, 695; Bybee, 1985:131). In addition, as a type, argument constructions marked with the Reflexive Marker should also display a lower degree of productivity in general compared to the neighbor verbs due to the locality effect (cf. Barðdal, 2008; Bybee, 2010).

Finally, through the concept of distance we arrive at the proper definition of the cross-paradigmatic relation. A cross-paradigmatic relation is defined as a constant difference in unit distance between items. The definition borrows the notion of unit from Cognitive Grammar, where a unit is defined relative to entrenchment. Langacker (1987:59) gives a general definition of entrenchment: "Every use of a structure has a positive impact on its degree of entrenchment, whereas extended periods of disuse have a negative impact. With repeated use, a novel structure becomes progressively entrenched, to the point of becoming a unit; moreover, units are variably entrenched depending on the frequency of their occurrence."

In order to count as a cross-paradigmatic relation, the unit constituting the difference must be entrenched in language. There is very little doubt whether the Russian Reflexive Marker constitutes a unit. First, it is used frequently. Second, it is applied to novel items. Third, it covers a fairly substantial number of items. Consequently, the constant difference leads to a cross-paradigmatic relation between them and the non-reflexive verb constitutes the neighbor verb of the reflexive verb if and only if the distance between them is constant. The reflexive verbs form their own system within the grammar of Russian, albeit interconnected (cf. Янко-Триницкая, 1962:30).

In sum, this section introduced another important theoretical device to capture the structure of the lexicon. The lexicon is not an unstructured reservoir of items. The proposed variables of the Neighborhood Density and the Neighborhood Distance enable us to model this structuring. At the same time, it is important to keep in mind that the proposed variables are an operationalization for the concept of structured paradigm. Nonetheless, the cross-paradigmatic relation between the reflexive and the non-reflexive verbs is quantified. For example, *оказаться* 'seem, appear,' and *оказать* 'render' are connected through cross-paradigmatic relation. In order to model the full cross-paradigmatic relation, the strength of the perceived semantic similarity between words is discussed in the following section.

3.1.4 Lexical Networks: Semantic Similarity

The fundamental tenet of the pair account is the shared semantic similarity between the items. Once multiple argument constructions are included, the stability of the shared semantics becomes increasing difficult to maintain and establish. Paducheva (Падучева, 2004) simply states that it is possible to find a core meaning, but no indication is given how exactly this is to be achieved. The concept of stable meaning is also embedded in usage-based models, as in Goldberg's (2006:39) Argument Construction Grammar. Another complication is the very essence of stable meaning and how to arrive at it (cf. Ramscar, 2002). If usage patterns are used as such, the exact profile of the pattern may guide the interpretation. In my view, the principle of the stable meanings lies in the relation between a particular reflexive and the neighbor verb and the semantic similarity need not cover all the potential senses associated with them individually. Furthermore, Budanitsky and Hirst point out that semantic similarity is a narrower concept compared to relatedness. For example, *bank* is semantically similar to *trust company,* but dissimilar items can be semantically related through lexical connections. In usage, *money* and *river* may be used as cues to disambiguate the meaning of *bank* in English (Budanitsky & Hirst, 2006).

Accordingly, the verbs *кипятить* 'boil' ~ *кипятиться* 'boil, worry' illustrated in Section 3.1.1 may still be perceived to be semantically similar. The 'boil' -sense may be enough to facilitate higher similarity, although the 'worry' -sense is only part of the semantic structure of the reflexive verb. Thus, the question becomes to what extent the verbs, such as *кипятить ~ кипятиться,* are perceived to be semantically similar. Sections 3.1.2 and 3.1.3 introduced the concept of the Neighborhood Density and the Neighborhood Distance to model connectivity between items. This section introduces the final component of connectivity used in this study, namely the perceived semantic similarity of the cross-paradigmatic verbs.

A semantic similarity task is the obvious choice to offer some answers to the question. Mass rating tasks would have to be set up due to the large number of verbs that form the cross-paradigmatic relation (*n* = 717) in the sample. Another complication is the fact that the verbs vary in complexity, ranging from morphologically simple verbs, such as *двигать* 'move' ~ *двигаться* 'move,' to prefixation, such as *напитать* 'feed' ~ *напитаться* 'eat enough.' Furthermore, a small number of verbs have either homographs or homonyms based on the infinitive forms leading to ambiguity. Homographs have identical spelling, but differ in pronunciation, (e.g., *разбегаться* 'start to run a lot' versus *разбега́ться* 'scatter'). Primarily, homographs have a difference in stress and the previous pair also differs in aspect in Russian, the latter being imperfective verb. Homonyms have the same spelling and pronunciation. For example, the verb *находиться* has three homonyms according to Kuznecov: 'be located,' 'walk plenty of,' and 'find' (Кузнецов, 2009 [1998]).

The potential ambiguity of the verbs was checked based on the information provided in three dictionaries: Evgen'eva (Евгеньева, 1999), Kuznecov

(Кузнецов, 2009 [1998]), and Efremova (Ефремова, 2000).[50] Items were considered either as homonyms or as homographs when they had two or more separate entries in the dictionaries. However, the dictionaries do not give consistent results for lexical entries. For instance, *раздаться* has two entrie. They are the 'sound' and 'widen' meanings in Kuznecov (Кузнецов, 2009 [1998]), but not in Evgen'eva. Additionally, a separate entry is given for the verb *сдаться*, for example, with the modal meaning, 'be necessary' compared to the 'give in' (Кузнецов, 2009 [1998]). It is worth pointing out that the Reflexive Marker is associated with modal semantics (Geniušienė, 1987; Gerritsen, 1990; Князев, 2007). Whether the modal sense should be given a separate entry is an interesting question, but it is not pursued in this study. Generally, homonymy is considered to be the last resort. Thus, the conflicting entries were resolved based on the information provided in Evgen'eva (Евгеньева, 1999) because fewer separate entries to the reflexive verbs are consistently given.

For the reflexive verbs, the database contains eight verbs that have a homograph and 26 verbs with homonyms. For the neighbor verbs, the figures are nine homographs and 33 homonyms. However, the ambiguities do not necessarily extend across paradigms, (e.g., *стро́иться* 'build' ~ *стро́ить* 'build' versus * ~ *строи́ть* 'triple' and *завяза́тьься* 'tie up' ~ *завяза́ть* 'tie up' versus * ~ *завяза́ть* 'stall'). The homonyms also differ in aspect the latter being imperfective verb. The data contained 15 overlapping homonyms. The semantic similarity ratings are biased for these verbs, as contextual information was not provided, but effect is minor considering the total number of cross-paradigmatic verbs (*n* = 717) in the database.

An expert rating task was conducted with two native speaker linguists. The cross-paradigmatic verbs (*n* = 717) were divided into four data sets and encoded using an Access form which was provided for the raters.[51] The raters worked independently of each other, and resubmitted the Access database once the data sets were completed. There was no feedback. The raters were instructed to finish one set at a time. Additionally, the raters were instructed to decide based on their initial perception. A pair, a neighbor, and a reflexive verb, appeared on the screen one at a time and the pairs were rated on a binary scale either similar or dissimilar. The items always appeared in the same order, a neighbor and reflexive verb, to avoid potential differences in judgments due to point of reference because similarity judgments are known to be context-dependent and asymmetrical (Spivey, 2007:272-274; Suttle & Goldberg, 2011:1257; Tversky & Gati, 1978; Whitten, Newton Suter & Frank, 1979).

The two raters had an 84% agreement indicating fairly high consistency.

[50] Another issue is related to the concept of contemporary Russian. Certain verbs have a homonym but they were tagged as archaic, for example *свари́ться* 'cook' versus *сва́риться* 'argue.' The latter is considered to be archaic. These cases were excluded.

[51] Considering that the pairs are a dominant method in linguistics, expert raters may have different conceptualizations about what counts as a pair compared to naive native speakers.

Related to this is the role of the items upon which the raters disagreed. Assuming that the 84 percentage-wise agreement reflects the true agreement among the raters, the probability of agreement on a wrong label by chance is a small 2.56%, (1 - .84)*(1 - .84) = .00256. Although native speaker ratings are taken to be "correct and unbiased" in a weak sense (cf. Budanitsky & Hirst, 2006:43), the disagreements might be due to the perceived saliency of a particular sense of one of the verbs leading to a situation in which the senses are weighted differently among individual raters, such as *казаться* 'seem, appear' ~ *казать* 'show' and *деться* 'disappear, escape' ~ *деть* 'place, leave.' This offers a way to implement a more fine-grained categorization of the similarity ratings given in Table 3.1-1.

	None	Dissimilar	Intermediate	Similar	Sum
Reflexive Verb	102	38	114	565	819

Table 3.1-1 Distribution of the variable Reflexiva Tantum measured as the degree of semantic similarity based on the unique reflexive verbs (*n* = 819) in the databse.

The reflexive verbs, which do not have a neighbor verb in the database, are labeled as "None," the traditional definition of reflexiva tantum verb as was outlined in Section 3.1.1. The label dissimilar refers to the verb pairs that both raters perceived to be semantically dissimilar, like *получиться* 'happen, become' ~ *получить* 'receive, get,' *разобраться* 'grasp' ~ *разобрать* 'dismantle,' and *оказаться* 'seem, appear' and *оказать* 'render.' The label intermediate is assigned to verb pairs for which the raters disagreed on the status of semantic similarity, for example *стремиться* 'strive, aim' ~ *стремить* 'direct' and *перебираться* 'get across' ~ *перебирать* 'sort, handle.' As indicated previously, the chance agreement on the wrong label is small; therefore we might anticipate fluctuation with these verb pairs in future studies. Finally, the label similar refers to verb pairs which were perceived to be semantically similar by both raters, like *проводиться* 'be underway, be conducted' ~ *проводить* 'lead, conduct' and *мечтаться* 'dream' ~ *мечтать* 'dream, wish.' The concept of Reflexiva Tantum is now operationalized as the measure of the degree of perceived semantic similarity of the cross-paradigmatic relation.

In terms of the structure of reflexive verbs, there appears to be strong cross-paradigmatic semantic connectivity. First, based on the unique reflexive verbs (*n* = 819), 88% of them are connected to a neighbor verb. Second, the data indicate that this connectivity is further strengthened by the perceived semantic similarity between them, as 79% of these reflexive verbs are also perceived to be semantically similar. This structure leads to a strong clustering of verbs with regard to the phonological and semantic similarity. These reflexive verbs can be construed as forming a bridge across paradigms with varying degrees of strength. Finally, by combining the measures of the Neighborhood Density, the Neighborhood Distance, and the perceived semantic similarity, the degree of connectivity of items is now fully operationalized.

3.1.5 Frequency and Frequency Effects

Frequency is considered to be one of the best predictors in lexical decision tasks, whether or not a certain string constitutes a word a language. Murray and Forster (2004:721) state: "Of all the possible stimulus variables that might control the time required to recognize a word pattern, it appears that by far the most potent is the frequency of occurrence of the pattern […] Most of the other factors that influence performance in visual word processing tasks, such as concreteness, length, regularity and consistency, homophony, the number of meanings, neighborhood density, and so on, appear to do so only for a restricted range of frequencies or for some tasks and not others." Cognitive and constructional models follow the usage-based approaches attributing a significant importance to frequency, considering it to be one of the fundamental properties of a linguistic system. This perspective sets them apart from the structural and generative paradigms. Especially, the Chomskian paradigm follows the position that usage patterns and frequency are irrelevant for the study of grammar. The strongest opposition to the basic tenets of usage-based models is expressed by Newmeyer (2003).

Bybee's studies on frequency and frequency effects are pioneering in cognitive and functional approaches to language and the basic properties of frequency effects identified in her studies. Her studies cover the significant domains of linguistic systems. These include morphology (Bybee, 1985), phonology (Bybee, 2001), the formation of categories (Bybee, 2007), and more recently, the interaction between categories and constructions (Bybee, 2010). At the same time, frequency of use has, for the most part, always been part of the functional linguistic paradigm. For instance, Paul (1989) has already argued that frequency of use is one of the primary factors influencing changes in inflectional paradigms.

Generally, morphology is one of the research areas where frequency and, consequently, probabilistic models are increasingly utilized compared to the traditional rule-account or symbolic manipulation, where morphological structures are modeled as gradient categories rather than a dichotomy (cf. Hay & Baayen, 2005 for an overview). Frequency and pattering are the basic properties of usage-based grammars even if they are fully compositional. Frequency and skewed input have also been shown to be essential parts of language acquisition (Goldberg, 2006; Tomasello, 2003).

The subsequent sections introduce the basic effects associated with frequency closely following Bybee's studies. They are used to formulate the different effects in terms of token and type frequencies. As frequency in itself is a distributional property and its meaningfulness is applicable only in some context, token and type frequency are used to differentiate levels of granularity. In terms of verbs, a token frequency constitutes the frequency of an individual form attested in the sample while the type frequency is used to refer to the summed frequencies of the inflectional forms of a particular verb, also commonly referred to as lemma.

3.1.6 An Outlook on Frequency Effects

In terms of frequency effects, as Bybee formulates it (2007:17-18), there is always a question of directionality, namely whether frequency is a cause or an effect. Because frequency is only observable in relation to some distribution, its unique contribution is in the formation of a certain category. For example, the Reflexive Marker is certainly an interesting question. Bybee does not take a categorical position in this question. Instead, she considers both positions to be present. When a certain pattern is observed along with its frequency, it constitutes an affect. In contrast, repeated experiences can impact mental representation constituting a cause for the observed frequency effects (Bybee, 2007:18; cf. Chesley & Baayen, 2010:1367). This ambivalence or duality of frequency and frequency effects is prevalent in cognitive and functional approaches. Related to this, Baayen has demonstrated that lexical properties of monosyllabic and monomorphemic words (*N* = 1,042) account for 91% of the variance of their frequency. This indicates that frequency in itself is partly co-determined by a number of other factors because a measure includes both frequency-as-repetition and frequency-as-contextual-experiencer, (e.g., for noun—verb ratio, written—spoken ratio and the information richness of the inflectional paradigm of the word) (Inflectional Entropy cf. Section 3.1.8 and 3.1.9) (Baayen, 2010a:436, 444-446).

The best example of iconicity is the one-form-one-meaning principle. In contrast, economy is a competing factor of iconicity and can be considered as the motivational pathway influencing the formation of polysemy. Economy amounts to the reduction of the inventory of linguistic items as much as possible yielding polysemous items deviating from the principle of iconicity (Cristofaro, 2005:8-9, 289-290). Importantly, Köhler (1986) is one of the first to demonstrate that (log) frequency is correlated with the number of meanings of a word in German. This was replicated by Baayen and Moscoso del Prado Martín (2005) for verbs in English, Dutch, and German. Similar results for Russian verbs are demonstrated by Kyröläinen (2012). Considering that reduction and polysemy are typically associated with frequency, the question of cause and effect is still present. However, this question is only pressing if a linguistic category is deemed to be formed based on some single factor. In this study, frequency as a variable is part of a multivariate approach alleviating the issues of cause and effect.

Bybee's studies on frequency have identified three major effects: 1) the conservative effect, 2) autonomy effect, and 3) reducing effect. All of them are associated with high frequency tokens. Bybee attributes the conservative effect on the accessibility of a given token, which is also supported by experimental evidence, like lexical decision tasks. Frequency is typically negatively correlated with reaction times, as the higher the frequency, the faster the reaction time (Baayen et al., 2006; Forster, 2004; Hino, J & Pexman, 2002). Repeated usage strengthens the association of the memory representation. Additionally, the conservative effect of frequency is typically used as a motivational factor in usage-based studies of irregular verb forms. For example, *keep* → *kept*, has resisted the past tense formation with the regular suffix *-ed* (Baayen & Moscoso del Prado Martín, 2005; Bybee, 2010). Similar examples can be drawn from Russian, verbs of motion, such as *идти* 'go' and *нести* 'carry,' have maintained their paradigms, which can be considered to be deviant in contemporary Russian. The nexus between frequent verbs and argument constructions is demonstrated by Goldberg. Few high frequency verbs dominated the structure of the argument constructions in mother's speech based on Child Language Data Exchange System database, for example *go* in the Intransitive Motion Construction and *give* in the Ditransitive Construction. The observations support the view of usage-based models and the role of frequency in language acquisition (Goldberg, 2006:75-78). However, it does not follow from this that frequency of use is the single and the only contributing factor (cf. Baayen, 2010a; Bybee, 2010:48-53).

The reducing effect is typically related to the reduction in articulation affecting both high frequency phrases and items. The prominent evidence for this effect is, the reduction of *don't* in American English. According to Bybee, the reduction is prevalent in conjunction with the first person pronoun (Bybee, 2010:43-44 and references therein). The reducing effect is phonological in nature. Thus, this effect is not considered in this study.

3.1.7 Frequency: Practicalities and Distributions

The absence of frequency information in previous studies can be partly attributed to the fact that large scale frequency information, in general, has not been available for Russian. For example, the frequency dictionary by Zasorina (Засорина, 1977) contains proximately 40,000 words based on a sample of one million words. More recently, a new frequency dictionary by Lyashevskaya and Sharov (Ляшевская & Шаров, 2009) was made available. The dictionary is based on a sample of the Russian National Corpus (100 million words). Nonetheless, a caveat associated with frequency information is the quality of the sources and the sample size crucially affecting the value of the lexical frequency information. At the same time, large scale corpora are becoming more widely available, making it possible to evaluate the quality of different sources of frequency information, and validation studies on frequency information are a growing topic.

Recent validation studies have shown that the sample size is an important factor, for example 10 million words, offering better estimates for frequency in comparison smaller corpora. However, the effect of the sample size does not follow the assumption more is better. Instead, a larger sample size only makes the effect of low frequency items more stable. Another important factor is the composition of the data used to build the frequency counts. Brysbaert and New state that the standard assumption of the primacy of the media (newspapers and magazines) and literature consisting of edited text may lead to a slight bias in frequency estimations because repetition is typically avoided in these genres

(Brysbaert & New, 2009:978-979).[52] In this regard, the frequency dictionary by Lyashevskaya and Sharov (Ляшевская & Шаров, 2009) offers a sufficient sample size and may be considered to be fairly stable. Any reference to frequency information used in this study is based on the estimations provided by Lyashevskaya and Sharov (Ляшевская & Шаров, 2009).

Another important aspect in utilizing frequency is, yet again, the level of granularity, whether to utilize type or token frequency. The frequency dictionary gives the type frequency, which may lead to cumulating the frequency effect in certain cases. Type frequency is invariant to possible token effects. Reflexive Verbs, which have multiple valency patterns, can also be sensitive to person marking, amounting to a difference in token frequency, (e.g., *остаться* 'stay, remain'). The impersonal usage pattern is delimited to third person; whereas no such restrictions apply to personal types. Brysbaert and New compared type frequencies against token frequencies in English. They ran several regression models where lexical decision latencies were modeled as a function of log frequency (+1). The type frequency offered only a slight advantage over the token frequencies in the range of 1%-2%. However, the advantage diminished even further when the number of letters and syllables was added to the models. Their data were extracted from the English lexicon Project, which contains word processing latencies for over 40,000 English words (Brysbaert & New, 2009:982-984). The difference between the token and type frequency may well be dependent on the phenomenon and the exact research questions.

The frequency dictionary includes the most frequent lexical items in the Russian National Corpus. The frequencies are normalized to one million words. The occurrence of an item is divided by the total number of items in the corpus and then multiplied by one million. Highly infrequent items are not included, leading to missing values. A common practice to compensate missingness is to impute them (Gelman & Hill, 2007:530-531). A simple and straightforward method is to add some constant to all values commonly referred to as the start (Agresti, 2002:397-398; Mosteller & Tukey, 1977). It is a common method used to handle missing values with frequency counts (Manning & Schütze, 1999). Thus, a constant +1 was added to all frequency counts to compensate missingness.

There are certain issues, nonetheless, related to imputing missing values. The sample contains 31 unique reflexive verbs with missing frequency information, for example *поймагься* 'catch', *моргаться* 'blink' and *детализироваться* 'specify'. A slightly higher number of missing values is attested with the unique neighbor verbs, n=65, for example *базировать* 'base,' *случить* 'breed,' and *попытать* 'torture for a while.' This is not an issue, as the missing values are still separated for these types of verbs. However, another issue of missingness is related to the reflexive verbs lacking the neighbor verb. For these verbs, the missingness is related to their structural properties and not to the sampling procedure, a

[52] As Brysbaert and New (2009:987) conclude: "knowing which frequency measure is the best is one thing; having access to it is another."

distinction between a sampling zero and structural zero. The sampling zero refers to the verbs, which are not included in the frequency dictionary, contrasting the structural zero (cf. Agresti, 2002:392-393). The data contains 102 structural zeros. To avoid missing values in the data, these verbs have a dummy neighbor verb "None" with the frequency of one. In terms of interpretation, the used imputation procedure factors in the possibility that these verbs may have a neighbor verb in certain contexts. At the same time, the procedure treats the possible effect of the cross-paradigmatic relation between these two types as equal.

Another case of amplification, an artifact of the assembly method of the frequency dictionary, is present when items are not disambiguated. Ambiguity leads to a slight amplification of the frequency effect on a particular item. As was outlined in Section 3.1.4, the source of amplification is homonymy in this data set (cf. Baayen & Moscoso del Prado Martín, 2005:670). At the same time, Budanitsky and Hirst (2006:24 footnote 6) consider that the usage of a non-disambiguated corpus is a trade-off between accuracy and size, but the usage of a non-disambiguated corpus is a more general approach. Additionally, disambiguated corpora are not available for Russian and, as was pointed out earlier, the frequency estimations for low-frequency items are likely to be biased as a disambiguated corpus would certainly be smaller than un-disambiguated.

Table 3.1-2 gives the frequency distribution of the reflexive and neighbor verbs on the original scale (+1).

Reflexive verbs					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	4.10	10.30	27.15	26.45	532.90
Neighbor verbs					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.00	2.85	10.50	46.90	34.50	2398.00

Table 3.1-2 Frequency distributions of the reflexive (*n* = 819) and neighbor (*n* = 717) verbs on the original scale. A constant +1 was added to all values.

A drastic difference is found in the mean value. In skewed distributions, the mean is sensitive to skewness and typically lies in that direction indicating that the neighbor verbs have a longer right tail. This property is visible in Figure 3.1-4. Another question related to frequency effects is the interpretation. Typically, frequency is discussed in relation to a discretized scale of low, intermediate, and high frequency. These notions are connected to the size of the sample and cannot be interpreted as representing some absolute threshold value. Because the sample size contains a fairly large number of verbs, the frequency effects can be tied to the distributional properties. Thus, values centered on the first quartile and less might be viewed as pertaining to the category of low frequency. Similarly, the intermediate frequency can be interpreted as reflecting the values around the median. Finally, the third quartile and greater can be taken as the discrete version of high frequency.

When larger samples are used, frequency is commonly transformed to a logarithmic scale. Word frequency distributions tend to follow a log-normal distribution (Howes & Solomon, 1951). Additionally, log transformation reduces the effect of possible outliers, smoothing frequency distributions, especially if the classical regression methods are employed. In this study, natural logarithms are used in the statistical models and graphs. The log transformation is possible only for strictly positive values. Hence the addition of constant +1 is required as log(1) = 0. The frequencies on log scale are given in Table 3.1-3.

Reflexive verbs					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.000	1.411	2.332	2.381	3.275	6.278
Neighbor verbs					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0	1.047	2.351	2.41	3.541	7.782

Table 3.1-3 Frequency distributions of the reflexive (*n* = 819) and neighbor (*n* = 717) verbs on log scale. A constant +1 was added to all values.

For the reflexive verbs, the three highest frequency verbs are *оказаться* 'seem, appear' (6.278 freq), *являться* 'be' (6.261 freq), and *остаться* 'stay, remain' (6.181 freq). The same order of the neighbor verbs is *сказать* 'say' (7.782 freq), *говорить* 'speak' (7.470 freq), and *хотеть* 'want' (6.900 freq).

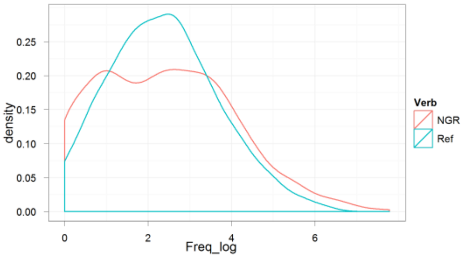


Figure 3.1-4 Density plot of the frequency distributions of the reflexive (*n* = 819) and neighbor (*n* = 717) verbs on log scale. A constant +1 was added to all values.

Figure 3.1-4 gives the density plot of the frequencies of the reflexive and neighbor verbs on log scale. The neighbor verbs appear to have a slight bimodal distribution shown by the two peaks. The graph brings forth the distributional differences. A fairly large number of the reflexive verbs are located around the median (2.332) and mean (2.381) values. In contrast, the neighbor verbs tend to have a higher number of verbs located around the third quartile (3.541) and the maximum value (7.782). Similarly to the density and distance measures, the distributional differences on log scale between the verbs connected through the cross-paradigmatic relation (*n* = 717) was tested for reflexive verbs (min. = 0, median = 2.332, and max. = 6.278) and neighbor verbs (min. = 0, median = 2.351 and max. = 7.782). A two-tailed Wilcoxon rank-sum test was used and the difference was not statistically significant (*W* = 256459.5, *p*-value = 0.9405). The results are surprising considering that more complex forms are typically less frequent and already observed by Harwood and Wright (1956). Similar to the results obtained with the Neighborhood Density, the reflexive and neighbor verbs appear to have similar distributions and the differences are found within the distances in the neighborhoods. From a usage-based perspective, the results offer support for the view that the reflexive verbs form a system of their own and the cross-paradigmatic relation is a connection and not a derivational relation.

Another issue is related to the sampling frame used in this study and whether it created artifacts. To estimate this possibility, all the verbs were extracted from the frequency dictionary (*n* = 12 328). Figure 3.1-5 gives a density plot of the frequencies of the reflexive and non-reflexive verbs attested in the frequency dictionary.

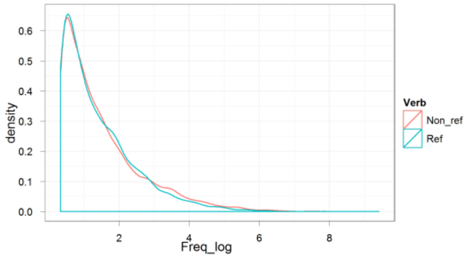


Figure 3.1-5 Density plot of the frequency distributions of the reflexive (*n* = 3545) and non-reflexive (*n* = 8783) verbs on log scale based on the frequency dictionary. A constant +1 was added to the values.

The distributions of the reflexive and non-reflexive verbs are almost similar. The difference between them is at the maximum range. Additionally, the distributions closely follow well-known frequency distributions (curvilinear shape). Linguistic categories tend to have a high number of low frequency items and only a relative few high frequency items (Biber, 1993; Zipf, 1965 [1935]).

Reflexive verbs					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3365	0.5878	1.0990	1.4010	1.9020	6.2780
Non-reflexive verbs					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.3365	0.6419	1.1310	1.4960	1.9880	9.4060

Table 3.1-4 Frequency distributions of the reflexive (*n* = 3545) and non-reflexive (*n* = 8783) verbs on log scale based on the frequency dictionary. A constant +1 was added to the values.

The verb *быть* 'be' (9.406 freq) is the most frequent verb of the non-reflexive verbs, a drastic difference between it and other high frequency verbs. The second most frequent verb is *мочь* 'can' (7.977 freq), and the third most frequent verb is *сказать* (7.782 freq). It seems that the sampling frame introduced some artifacts. First, the ratio of the verbs with an intermediate frequency, around log 2, is overrepresented in the sample. However, this artifact shows that the sample is confined to typicality. As Biber (1993) has demonstrated that even a small sample can be representative if typicality effects are sought after. Second, the bimodal distribution of the neighbor verbs may be another artifact, although the matter is less straightforward. For instance, not all non-reflexive verbs have a phonologically corresponding reflexive verb and vice versa. Thus, the bimodal distribution might be a property of the neighbor verbs that does not generalize to the non-reflexive verbs as a whole.

3.1.8 Entropy: Uncertainty and Information Content

This section introduces the concept of entropy from information theory and its application to linguistic data and the rationale behind it; borrowing its foundation from studies on morphological processing. Section 3.1.9 spells out the interpretation of an entropy-based measure and demonstrates its application to quantifying the strength of connectivity between the verb-specific and argument constructions. Lexical variation is one of the basic properties of language. In addition to describing lexical variation in terms of taxonomic hierarchies, studies on lexical processing have typically attempted to quantify this variation. Perhaps the most obvious quantification is frequency as was already established in Section 3.1.5. Additionally, Sections 3.1.2 and 3.1.3 established measures of lexical connectivity in terms of densities and distances. Another crucial aspect of verbal semantics is the connection between the verb and the abstraction (argument constructions). Goldberg takes a similar stance, namely to tease apart verb-specific meaning from sentence meaning. She utilizes conditional probability estimations to bring forth the possible support originating from verbs (Goldberg, 2006:105, 117-119). This section introduces a new measure called Constructional Entropy aligning with studies on morphology where the information content of a paradigm is quantified (cf. Milin, Durdevic & Moscoso del Prado Martín, 2009; Moscoso del Prado Martín, Kostic̀ & Baayen, 2004).

Generally stated, entropy is a measure of uncertainty associated with a random variable. In this sense, the concept of entropy is attributed to the mathematical theory laid out by Shannon (1948), and it is referred to as Shannon's entropy, now commonly used in telecommunication and compression algorithms. The basic application of entropy is to evaluate the predictability of information contained in a message in some finite set. Thus, the entropy value is quantification of the expected value in a message. For example, a finite set of coin tosses with a fair coin has the maximal entropy value because there is no predictability. Heads and tails have equal probability (0.5). It follows from this that adding more information to the finite set will decrease entropy and increase predictability. The knowledge that the coin is rigged, and always yields tails in the finite set, the entropy value is 0 and, thus, maximally predictive (Shannon, 1948:10-11).

The previously outlined description gives the basic components of modeling a discrete random variable in terms of entropy yielding Shannon's entropy (Shannon, 1948:11) H(X), where H is entropy and it is defined as $H = -\sum p_i \log p_i$. Generally, the basic components of Shannon's entropy measure are probability *(p)* and logarithm (log). Thus, *p* is the probability of a value *i*. The second component is the logarithm, which can be calculated differently depending on the selected base. By selecting a different base, it does not, however, change the interpretation of the entropy value. The different bases are just measurements in different unit size. Perhaps the most commonly used measure of entropy is bits corresponding to log unit with base 2. There is an intrinsic connection between the log2 and probabilities. The log2 encodes a string in a binary form, a sequence of 0s and 1s. Similarly, probability ranges between 0 and 1 (Shannon, 1948:1-2).

Applications of entropy-based models in linguistics are currently increasingly popular, especially in computational and experimental linguistics.[53] The motivation behind this wide array of applications is perhaps two-folded. On the one hand, the increasing number of corpora facilitates the application of probabilistic models to linguistic structure. On the other hand, instead of trying to model context directly, probabilities of a random variable can be used if we have a known set reducing the problem of defining context considerably (Levy, 2008; McDonald & Shillcock, 2001; Milin et al., 2009; Moscoso del Prado Martín et al., 2004).

Recently, Genzel and Charniak (2002) have shown that the entropy value of a sentence, taken out of context, is correlated with the sentence position in discourse in English. The results are also replicated in Russian. These results are intuitively in accordance with discourse structure; out-of-context sentences are harder to understand and discourse initial sentences can never be out-of-context (Genzel & Charniak, 2003). Moreover, Keller (2004) has shown that the corpus-based results of Genzel and Charniak are confirmed with eye-tracking data. These findings demonstrate that entropy-based measures can, at least partly, be used to model the influence of context in terms of entropy.

[53] In addition to this recent expanse of entropy-based models, Goldsmith discusses the early status of information theory in the works of Jakobson, Trubetzkoy, and Hocket in phonology (Goldsmith, 1998). In contrast, Chomsky (1986:342) questions the relevance of information theory in linguistics: "There was a lot of euphoria about such approaches to language. In part, it came from the prestige and achievements of information theory, which involved similar notions; in part, the statistical approaches to linguistics, and, in part, it had a kind of technological air to it. There was a lot of euphoria at that time in the area of linguistics in general, about the potential great achievements that lay head along these lines. It was thought that they were already partly real."

For the purposes of the present study, the argument constructions are interpreted as the known set. Argument constructions profile the sentential semantics on an abstract level. For a constructionist account, the question relates to the relation between the verb and the argument construction. Some verbs may function as better cues for the sentential semantics while others might be more dependent on the argument construction. Thus, there is strength of connectivity between them and the following section is used to offer a new measure to quantify this relation.

3.1.9 Constructional Entropy

This section introduces an application of entropy-based measure to argument construction. For example, Moscoso del Prado Martín et al. (2004) have proposed a model where the amount of information carried by word and its paradigm are quantified, (e.g., *think* and *thinker).* This approach is an application of Shannon's entropy measure to evaluate in probabilistic terms the type- and token-based effects in morphological processing. Thus, it is a measure which incorporates two well-established frequency effects in the lexicon.

Following Milin et al. (2009:53), the Inflectional Entropy is defined as the relative frequency of the inflected variants of a given word. It combines both the type and the token frequencies (cf. Moscoso del Prado Martín et al., 2004:4). Similarly, we can postulate that the informational entropy of an argument construction is related to the verbs which instantiate it.[54] This follows the basic tenets in constructionist approaches to studies of language acquisition, which have shown that abstractions (argument constructions), are formed over instances. At the same time, arriving at the concept of Constructional Entropy contains both practical and theoretical implications. Before turning to the theoretical aspects, the practical side is demonstrated.

To demonstrate how the Constructional Entropy was calculated, one of the infrequent types is used, the Stimulus Extension Construction. Following the tenet of the network model, the usage of an item will partly update the whole network of that specific item, the type frequency of the reflexive verbs was used. Table 3.1-5 gives the reflexive verbs supporting this construction type in the sample and the frequency of the verbs. Given the formula of Shannon's entropy, $H = -\sum p_i \log p_i$, the f(w_1) corresponds to the frequency of a verb and the f(w) is the cumulative frequency of the verbs supporting the construction type. The p(w_i) is the probability for a verb given the distribution in the sample. Thus, it is the relative frequency of a verb in the construction type where the frequency is divided by the cumulative frequency.

[54] Certainly, it is conceivable that a language has argument constructions which might have lost their verbal element. However, I will not speculate with this matter in this study because the argument constructions always contain a verb in the data.

Reflexive verb	f(*w*$_i$)	p(*w*$_i$)=f(*w*$_i$)/f(*w*)
ощущаться 'be felt'	13.9	0.337
забыться 'forget'	11.8	0.286
ощущаться 'be felt'	13.9	0.041
почувствоваться 'feel'	1.7	0.041
f(*w*)	41.30	

Table 3.1-5 Stimulus Extension Construction and its frequency distribution. The columns give the verbs supporting the construction, the frequency of the reflexive verb, and the relative frequency of the reflexive verb given the distribution of the argument construction in the sample. Based on this information, we can calculate the Constructional Entropy (*CH*) for the Stimulus Extension Construction:

–(0.337 * log2(0. 0.337)) + … + (0.041 * log2(0.041))
= 1.764

Thus, the Constructional Entropy of the Stimulus Extension Construction is 1.764 bits in the sample. As the Constructional Entropy for a particular construction type is constant in the sample, the contribution of a particular verb to the Constructional Entropy is used. For example, *ощущаться* 'be felt' contributes 0.529 bits to the Constructional Entropy of the Stimulus Extension Construction. The sum of these contributions is the Constructional Entropy of a particular construction. The Constructional Entropy is creased by the number of verbs and also when the probabilities of the verbs are similar.

A final clarification needs to be made in calculating the *CH* of a specific construction type. It is self-evident that certain verbs can appear multiple times in a construction type. This gives two options, consider only unique verbs or the whole distribution attested in the sample. The latter option is implemented. The verbs appearing multiple times have a higher influence on the cumulative frequency. Theoretically, this can be motivated by assuming that items appearing multiple times impact the formation of a category (Bybee, 2010; Hay & Baayen, 2005). Goldberg (2006:85-89) offers evidence that skewed input influences argument constructions whether a specific verb dominates the distribution or not. In this regard, these measures can be considered to highlight certain facets of entrenchment.

This has both practical and theoretical implications. From a practical side, this allows us to calculate the informational entropy of an argument construction based on a sample. In most cases, it is infeasible to have a total sample of a certain argument construction in a corpus. For example, the Reflexive Passive Construction is not tagged separately in the Russian National Corpus and it is impossible to automatically derive instances of this construction type from a corpus. From a theoretical point, the measure operates on type frequency. Type frequency in itself is a cumulative frequency over instantiations. Thus, the measure assumes that the type frequency of a particular verb is an important factor for a specific argument construction. As was argued in Section

2.2.3, verbs can have different strengths of connectivity to the argument constructions. Examples with the verb *оказаться* ‘seem, appear’ are repeated here for convenience, as in 3.1-1 and 3.1-2.

3.1-1 *А стенк-а оказа-л-а-сь тоненьк-ая* [...].
Also wall-NOM appear-PST-F-RM thinnish-NOM
The wall appeared to be thinnish.
[1985, RNC, Татьяна Рик. Про вредную Бабку-Ёжку // "Мурзилка," №6," 2001]

3.1-2 *Оказа-л-о-сь, что кто-то похити-л*
Appear-PST-N-RM that someone.NOM steal-PST.M
волшебн-ый амулет принцесс-ы.
magic-NOM amulet.NOM princess-GEN
It turned out that someone had stolen the magical amulet of the princess.
[1711, RNC, Сергей Седов. Доброе сердце Робина // "Мурзилка," №7," 2002]

In 3.1-1, the verb *оказаться* ‘seem, appear’ contributes 0.0598 bits of information to the Property Construction, whereas the contribution to the Content Construction is 0.1795 bits in 3.1-2. Based on the Constructional Entropy, there is less uncertainty when the verb *оказаться* ‘seem, appear’ appears in the Property Construction contrasting the Content Construction. Thus, the semantics of the verb and the argument construction align in the Property Construction for this particular verb. This prediction is in concordance with the semantic description of this verb, for instance, given in Paducheva (Падучева, 2004).

3.1.10 Causation and Transitivity

In addition to surface structure generalizations, cross-paradigmatic relations are another important factor for the formation and maintenance of linguistic categories. For the neighbor verbs, two properties are typically regarded to be central, namely causation and transitivity. Causation is another dominant property attributed typically to transitivity situated at the core of syntactic theories in general, as described by Croft (1998), Langacker (1999), Lakoff (1987), and Talmy (2000a).[55] Langacker defines the prototypical Agent as a person who volitionally initiates physical activity, through physical contact, in the transfer of energy to an external object. The definition of the Patient role is relative to Agent. It is an inanimate object that absorbs the transmitted energy and undergoes a change of state (Langacker, 1991:285).[56] The concept of

[55] In Talmy’s account, causation is part of force-dynamics, the manifestation of force in an event. Talmy makes fine-grained distinctions and offers a list of 20 features, for example, whether the force is present or absent, generic or particularized and pushing or pulling, among other things (Talmy, 2000a:462-463).

[56] Hopper and Thompson (1980:252-253) also posit similar argumentation for

causation figures prominently also in the descriptions of the Russian Reflexive Marker (Geniušienė, 1987; Gerritsen, 1990; Князев, 2007; Падучева, 2001). In terms of the standard pair account, certain verbs confine to the causative ~ decausative alternation, although Russian does not have a morphological causative marker. From a semantic point of view, a causative transitive verb profiles a situation type where one entity undergoes a change of state caused by another entity. Such transitive causative verbs *разбить* 'cause to break,' *истощить* 'exhaust,' *создать* 'create,' and *испугать* 'startle, make afraid'.

Croft distinguishes several different types of causation following Talmy's earlier work that was later reanalyzed under the category of force-dynamics. Action verbs, like *hit* and *break* are defined as indicators of physical or volitional causation. According to Croft, this distinction depends on animacy and control of the subject argument. Mental verbs covering emotion, cognition, and perception are defined as affective causation. The last type is inducive causation that holds between mental and social event types. Such verbs are *persuade* and *convince* (Croft, 1991:166-167). Nonetheless, Croft (1991:213) notes that the difference between physical and inducive causation is minor. These fine-grained distinctions have not been utilized in descriptions of the Russian Reflexive Marker. Crucially, the different shades of causation are actually stated relative to the semantic classes of verbs. As semantic classes are not used as a variable, these fine-grained distinctions are not established.

Mel'chuk considers that the non-causative reflexive verbs are semantically basic ones compared to their causative pairs, although formally they may be considered to be derived from the causative type. According to him, the directionality of the derivation is confined in both formal and semantic directions, creating complications for descriptive devices (Мельчук, 1995:467-468; cf. Янко-Триницкая, 1962:167-168). Related to the question on directionality, Mel'chuk postulates that the causative type is semantically more complex supporting this claim on the formation of causative verbs in Japan with the suffixes -(*a*)*s*-/-*sas*-. According to him, the causative is always derived both formally and semantically in these cases. Certain reflexive verbs are primary and also diachronic, such as *отпочковаться* 'gemmate' ~ *отпочковать* 'cause to gemmate' and *приземлиться* 'land' ~ *приземлить* 'cause to land' (Мельчук, 1995:459-460, 471). The basicness of the reflexive verb compared to the causative non-reflexive verb is also posited by Jahontov (Яхонтов, 1981). Similar verb pairs are present, for example, in German where the intransitive verb is diachronically the basic one and the causative variants are later formations, such as *liegen* 'lie' ~ *legen* 'lay,' *sitzen* 'sit' ~ *setzen* 'set,' *fallen* 'fall' ~ *fällen* 'fell,' *sinken* 'sink' ~ *senken* 'cause to sink' and *stehen* 'stand' ~ *stellen* 'put.'[57]

Babby argues that the decausatives are categorically derived from the

agentivity in terms of degree. The participant high in agency affects the mode of transfer.

[57] This is also visible in the inflectional paradigm of the verbs. The intransitive ones are strong verbs.

causatives. Otherwise, the Reflexive Marker would have a dual function. First, it is used to mark a decrease in valency with transitive verbs to derive the intransitive verb. Second, the formation of causation would have to be stated as an increase in valency because the removal of the Reflexive Marker from the intransitive verb would derive the causative verb, (i.e., an increase in valency). However, his native informants considered certain reflexive verbs to be basic in comparison with the causative, (e.g., *простудиться* 'catch cold' ~ *простудить* 'cause to catch cold'). Interestingly, Babby attributed this discrepancy to frequency of use (Babby, 1983:71-73). Babby is certainly correct that by accepting the possible dual function, the enterprise on obtaining the invariant meaning of the Reflexive Marker is compromised. A possible solution to the matter at hand, however, is a theory-internal issue in formal approaches bearing no consequences to constructionist accounts that do not operate on derivational relations.

Currently, the most exhaustive account on Russian decausative verbs is offered in Paducheva (Падучева, 2001). According to her characterization, the reflexive verb is semantically derived if agentivity can be regarded as the primary meaning of the non-reflexive verb compared to happening. In cases where this definition does not hold, the reflexive verb is semantically the base form (e.g., *растворить* 'cause to solve, melt down' ~ *раствориться* 'solve, melt down' and *обрушить* 'cause to rain down'~ *обрушиться* 'rain down' (Падучева, 2001:66-67, 69).[58] The division proposed by Paducheva appears to subsume Talmy's fine-grained properties, such as whether the causation is agentive and intentional (Talmy, 2000b:158, 167-168). This is yet another source that undermines the common practice of describing the reflexive verbs as derived from the non-reflexive verbs. However, by acknowledging this discrepancy, the explanatory power of derivation diminishes as the derivation can go in either direction, and most likely, is an item-specific property. At the same time, the definition hinges on defining the semantic content of happening for all the neighbor verbs and, once again, establishing the primary or basic sense for a particular verb.

In order to maintain replicability, the causative component of the neighbor verbs was extracted from the manually disambiguated subcorpus of the Russian National Corpus.[59] As expected, different approaches to causative verbs yield different lists. The semantic class of the lexical phasal verbs, such as *начать* 'begin' and *закончить* 'complete,' are not tagged as causative in the Russian National Corpus, whereas they are considered to be one of the main semantic subclasses of the Russian causative verbs by Paducheva (Падучева, 2001:69) and Knyazev (Князев, 2007:538). Table 3.1-6 gives the distribution of the semantic component Causative based on the unique the neighbor verbs.

[58] Although it seems that Paducheva has reconsidered her position (2003:174): "Decausatives are derived from causatives."

[59] Gerritsen (1990; cf. Янко-Триницкая, 1962:167) imposes a strict paraphrasing test in order to determine whether a pair constitutes a causative ~ decausative alternation, the ability to combine with *заставить*$_{\text{imp}}$/*заставлять*$_{\text{perf}}$ 'force, make.'

	Causative	Non-causative	Sum
Neighbor Verb	219	498	717

Table 3.1-6 Distribution of the semantic component Causation of the unique neighbor verbs (*n* = 717).

The number of the causative neighbor verbs appears to be fairly high, substantiating Paducheva's (Падучева, 2001) claim that the causative component is an important aspect of the semantics of the cross-paradigmatic relation, derivation in her position, in Russian and should not be ignored. Additionally, the variable Causation also involves a practical component, namely the role of the reflexive verbs that lack the cross-paradigmatic relation. For these verbs, the Causative component was tagged as "None," highlighting the departure from the cross-paradigmatic relation.

Importantly, the semantic component of causation is considered to be an inherent property of the neighbor verbs and it is part of the cross-paradigmatic relation and not a derivational component. From a usage-based perspective, serious mismatches appear if the causative ~ decausative is considered as an alternation type of its own. Table 3.1-7 demonstrates the issue with mental reflexive verbs.

Reflexive verb	Frequency	RT	Neighbor verb	Frequency
бояться 'be afraid'	266.5	None	Ø	0
пугаться 'become frighten'	10.2	Similar	*пугать* 'frighten'	32.6
волноваться 'worry'	49.9	Similar	*волновать* 'agitate'	33.1
насторожиться 'become concerned'	8.9	Intermediate	*насторожить* 'alert'	5.3
стесняться 'be ashamed'	30.3	Dissimilar	*стеснять 'embarrass'*	2.4

Table 3.1-7 Cross-paradigmatic relations of mental verbs based on Frequency, on the original scale, and perceived semantic similarity (RT = Reflexiva Tantum).

The reflexive verb *бояться* 'be afraid' lacks the cross-paradigmatic relation. Hence, it would constitute some separate type in terms of the pair account. The verbs *пугаться* 'become frightened' ~ *пугать* 'frighten' might be considered to follow the directionality of causative → decausative based on the frequency of use. This directionality is contrasted with the verbs *волноваться* 'worry' ~ *волновать* 'agitate,' suggesting a relation of decausative → causative. Finally, the verbs *насторожиться* 'become concerned' ~ *насторожить* 'alert' and *стесняться* 'be ashamed' ~ *стеснять* 'embarrass' display mismatches in perceived semantic similarity. Rather than assuming a binary relation cutting across paradigms, the cross-paradigmatic relation displays degrees of connectivity. However, the difference of the log frequency distributions of these two subgroups (*n* = 219), the causative neighbor verbs (min. = 0, median = 2.565 and max. = 6.267) and the reflexive verbs (min. = 0, median = 2.416 and max. = 5.261), was not statistically significant based on a two-tailed Wilcoxon rank-sum test (*W* = 23426, *p*-value = 0.6755).

The cross-paradigmatic verbs do not have a statistically significant difference in the distributions based on their Neighborhood Density or log Frequency. These reflexive verbs also have a higher number of semantically similar neighbor verbs. The causative component establishes a large subtype within the cross-paradigmatic relation and the data support the view that even subtypes follow the global properties of the reflexive and the neighbor verbs.

In addition to causation, transitivity is another component attributed to the neighbor verbs in the pair account. Transitivity and, subsequently, intransitivity are interconnected with both the semantic and syntactic aspect. The former connects particular instantiations as directed activity towards the object. The latter illustrates the form pole of this with the direct object encoded in the accusative case without a preposition (Князев, 2007; Храковский, 1974). Recently, Janda (2008b) explores the interconnectedness of transitive verbs with other case patterns, such as genitive and accusative prepositional phrases, (e.g., example *хотеть* ‘want’ and *надеяться на*$_{\text{acc}}$ ‘hope for’). In comparison, the semantic intransitive position is already echoed by Fortunatov (Фортунатов, 1899) in relation to transitive verbs. He considers that the Reflexive Marker signals a change in transitivity. In Fortunatov's taxonomy, the (in)transitivity is defined as the relation of the activity depicted by a lexical verb towards the subject. A notion which later was refined by Shahmatov (Шахматов, 1925) to include the relation between the subject, the predicate, and the object.

These patterns of the neighbor verbs were extracted from Efremova (Ефремова, 2000) and supplemented with data from Kuznecov (Кузнецов, 2009 [1998]). Because argument constructions are understood as generalizations over usage patterns, the term Valency is used to refer to the different types of the neighbor verbs. Obviously, dictionary based counts are not equal to a detailed lexicological study but they offer, at least, a systematic approximation.

	Transitive	Intransitive	Bivalent	Sum
Neighbor Verb	533	24	160	717

Table 3.1-8 Distribution of the Valency of the unique neighbor verbs.

Table 3.1-8 summaries the distribution of the variable Valency of the neighbor verbs. The label Bivalent was assigned to verbs which were tagged as transitive and intransitive, such as *жаловать* ‘accord, like,’ *загибать* ‘bend,’ and *треснуть* ‘burst, break.’[60] The distribution illustrates that only a small number of

[60] The three labels aggregate over usage patterns of the neighbor verbs, although verbs are typically associated with multiple patterns. The verb *жаловать* serves to

the neighbor verbs are intransitive, (e.g, *мечтать* 'dream,' *попасть* 'fall into,' and *светить* 'shine'). In terms of the cross-paradigmatic relation, the reflexive verbs appear to be highly associated with transitive neighbor verbs, estimated on the basis of the information provided in the dictionaries. Thus, the argument constructions of the neighbor verbs follow the prediction of the Principle of Distance. Shorter distances across densities ensure faster spread of information. The situation is, nonetheless, slightly different with the neighbor verbs because the bivalent verbs also include the Transitive Construction corresponding to multiple patterns. A classification tree was fitted to test three distributional differences: Valency of the neighbor verbs as a function of Neighborhood Density, Neighborhood Distance, and log Frequency. Frequency was included in the model because it is generally a strong predictor for the number of patterns or the number of senses (Baayen & Moscoso del Prado Martín, 2005; Köhler, 1986; Kyröläinen, 2012). Additionally, the intransitive neighbor verbs were excluded due to their infrequency in the sample. Figure 3.1-6 gives the fitted classification tree.

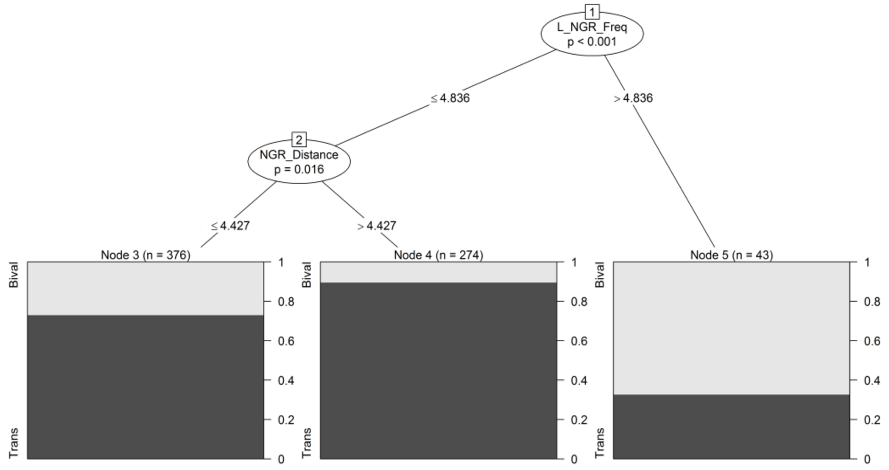


Figure 3.1-6 Classification tree of Neighbor Verb Valency as a function of Neighborhood Density, Neighborhood Distance, and log Frequency.

Following the domain-general principles, the results are expected. First, the Neighborhood Density is not statistically significant as predicted by the Hypothesis of Connectivity. The neighbor verbs are assumed to form a paradigm. The log Frequency appears to be the strongest predictor, yielding the first split: log Frequency greater than 4.836 or lesser than or equal to 4.836

illustrate the issue at hand. The 'like' sense corresponds to the Transitive Construction, whereas the Intransitive Construction is attested with the 'accord' and 'arrive' senses. Both of these are tagged as archaic in Kuznecov (Кузнецов, 2009 [1998]). In contrast, only the 'like' and 'accord' senses are given in Efremova (Ефремова, 2000). The latter is also tagged as archaic.

(node 1). The terminal node 5 contains the highest proportion of bivalent verbs, 67.4%. The other branch is in interaction with the Neighborhood Distance (node 2). A split was formed at the Neighborhood Distance greater than 4.427 or lesser than or equal to 4.427. The largest number of bivalent verbs (n = 102) was located in the terminal node 3. These results support the Hypothesis of Connectivity and the Hypothesis of Distance. Information spreads more freely across shorter distances, facilitating the formation of semantic densities, (i.e., multiple patterns in this case, all other things being equal).

In sum, this section introduced the semantic components of the neighbor verbs used in the model. In this study, Causation is understood as an inherent semantic property of the neighbor verbs. Additionally, the basic valency patterns of the neighbor verbs are included in the model in order to establish schematic cross-paradigmatic relations. It was shown that the two domain-general principles are at work in the formation of cross-paradigmatic relations and semantic densities in the lexical network.

3.1.11 Aspect and Tense

A fundamental tripartite property of a verb is related to the concept of tense, aspect, and mood (TAM). Recently, Janda and Lyashevskaya (2011) compare the interaction of verbs and TAM system in terms of relative frequency distributions in Russian. In this study, the category of mood is not considered as gerunds and participles were excluded from the sampling frame leaving the categories of tense and aspect. Both of these concepts are related to expressing time. The function of tense is to anchor a linguistic expression in time relative to other points. In contrast, aspect is related to the internal temporal dimension of a linguistic expression. Aspect is another salient property of the Russian verbs in addition to the Reflexive Marker (Бондарко, А. В., 1990; Зализняк, Анна А., Микаэлян & Шмелев, 2010; Храковский, 2005).

Aspect is an inherent property of the verb and it is obligatorily expressed in Russian. Moreover, Russian has invested heavily to mark this distinction through morphology, utilizing affixation, (e.g., ***с****делать*$_{perf}$ ~ *делать*$_{imp}$ 'do,' and *подписать*$_{perf}$ ~ *подпис**ыва**ть*$_{imp}$ 'sign') (Кронгауз, 1993; 1997). The morphological components are given in bold. Similarly to the concept of reflexive and non-reflexive verbs, the aspectual system is analyzed as consisting of pairs, as is illustrated in the previous examples. The crucial criteria in establishing them are the assumption that they have the same lexical meaning, but differ only in terms of their aspectual meaning, (i.e., imperfective and perfective meaning, at the maximally coarse-grained level). Additionally, the imperfective is considered functionally to be the unmarked form (Jakobson, 1989 [1932]:3; Виноградов, 1975; Тихонов, 1998; Шахматов, 1925).

In addition to these two mechanisms to form aspect, four minor groups are present in Russian, displaying deviations from the canonical formation. Suppletive forms constitute a small category, such as *взять* ~ *брать* 'take' and *поймать* ~ *ловить* 'catch.' Importantly, this category merges with the Russian Reflexive Marker, for such verbs as *стать*$_{\mathrm{perf}}$ 'become' ~ *становиться*$_{\mathrm{imp}}$

'become' and *лечь*$_{perf}$ 'lie down' ~ *ложиться*$_{imp}$ 'lie down' (Тихонов, 1998:22-23). Another deviation is demonstrated with two groups of verbs labeled as imperfectiva and perfectiva tantum verbs, a similar label to the reflexiva tantum verbs. These verbs only appear in either one of the aspectual forms. However, Zaliznyak et al. (Зализняк, Анна А. et al., 2010:14) point out that the tendency to form pairs is strong, especially in spoken language and dialects. These verbs display aspectual distinction, such as *поскользнуться* 'slip, slide once' ~ *поскальзываться* 'slip, slide' (cf. Ровнова, 1998; Соболев, 2005).

Another small group of verbs form the category of the so called biaspectual verbs. These verbs do not have morphologically distinct forms to encode aspect, like, such verbs as *реализовать* 'implement, realize' and *организовать* 'organize. The verbs pertaining to this group are typically loan words (Тихонов, 1998). Although the aspect is not morphologically marked with these verbs, Isachenko (Исаченко, 1960) has objected the analysis of these verbs as neutral. In usage, these verbs are used to mark either imperfective or perfective aspect.

Although the concept of the aspectual pairs is the dominant account in studies of aspect, all the realities of the aspectual system do not confine to pairs. A challenge for the pair account, are called aspectual triplets. An example is *съесть ~ съедать / есть* 'eat' (Храковский, 2005). A radical departure from the pair account is discussed by Janda (2007; 2008a). She proposes a clustered model for analyzing Russian aspect where aspectual distinctions form a network. A defense for the pair account is recently discussed in Zaliznyak et al. (Зализняк, Анна А. et al., 2010). For the purposes of the present study, the relation between these different approaches is beyond the scope of this study. Nonetheless, verbs are not collapsed into pairs or clusters.

One motivation for this is that previous studies have shown that aspect might interact, at least, with certain subcategories of the Reflexive Marker. The primary candidate to illustrate this possibility is the formation of the Passive Construction displaying an affinity towards the imperfective aspect. Another category is the decausative reflexive verbs in the diathesis tradition, which is claimed to gravitate towards the perfective forms (Падучева, 2001). At the lexical level of granularity, a certain lexical form can function as a gravitational center. By collapsing verbs into aspectually motivated categories, this possibility would be lost.[61] Moreover, the collapse approach would also impose a strong theoretical construct on lexical forms and introduce aspect as the primary grouping factor. By keeping these factors separate, the model allows to distinguish a possible unique contribution of aspect in the formation of the argument construction types. For comparison, a similar approach is taken in the Russian Grammar. Accordingly, aspectual pairs constitute separate lexical items, which are linked together through motivation, maintaining the aspectual pair distinction (Шведова & другие, 1982:584-585).

Another important facet related to aspect is the semantic classes of the

[61] Another possibility would have been to establish root forms of the reflexive verbs and use these as links between construction types.

imperfective and perfective aspect. However, semantic categories of aspect are not considered in this study. As a grammatical category, morphologically marked aspectual distinctions subsume semantic definitions. The most prominent account on semantic aspect is proposed by Vendler (1957). His classification is based on four categories that consist of activity (*run*), accomplishment (*build a house*), achievement (*find something*) and state (*love something*). (Braginsky & Rothstein, 2008; cf. Падучева, 2004:30-31) This position is clearly illustrated by Zaliznyak et al. (Зализняк, Анна А. et al., 2010:12-13). According to them, Russian verbs can be divided into two semantically broad groups which are state and non-state verbs. The perfective and the imperfective aspect carve this semantic space, allowing positing the more fine-grained analysis of the aspectual properties of a particular verb (cf. Шелякин, 1983). For the purposes of the present study, tense is defined following the definition in Russian Grammar. The periphrastic future is formed by combining the copula *быть* 'be' with the infinitive in the imperfective aspect (Шведова & другие, 1982:626-627).

	Aspect	
Tense	Imperfective	Perfective
Past	####	####
Present	####	*
Periphrastic Future	####	*
Future	*	####

Table 3.1-9 Interaction of aspect and tense in Russian. The hashes are used to indicate presence of the feature and the asterisk is used to indicate absense.

Table 3.1-9 illustrates the interconnectedness between tense and aspect. The hashes are used to indicate presence and the asterisk to indicate absence.[62] A slight complication was introduced with the sampling frame, as infinitives were not excluded in order to obtain periphrastic future forms. Surprisingly, the periphrastic future appears to be fairly infrequent, cf. Table 3.1-10. Thus, the data contains several infinitive forms yielding a complex predicate structure, typically a modal construction as in 3.1-3.

3.1-3 *бывший* *чекист* *мож-ет* *хорош-о*
former.NOM KGB.agent.NOM can-3S.PRS good-ADV
разбира-ть-ся *в* *экономик-е.*
grasp-INF-RM PR economy-PREP
A former KGB agent can easily grasp economy.
[1120, RNC, Беседа в Самаре (2001.08.31)]

[62] One solution would be to collapse the present (imperfective) and future (perfective) as is done by Janda and Lyashevskaya (2011:721 footnote 1). The higher level of granularity is used to tease apart possible support originating from Tense.

All these instances appeared in a personal construction type. Additionally, the subject argument is indexed with the infinitive and the main verb.[63] However, these verbs are not marked for tense. Thus, the label Infinitive is used to avoid missing values the data.

	Aspect		
Tense	Imperfective	Perfective	Sum
Past	334	523	857
Present	871	0	871
Future	0	114	114
Periphrastic future	34	0	34
Infinitive	32	81	113
Sum	1271	718	1989

Table 3.1-10 Distribution of Tense and Aspect of the reflexive verbs in the database. The class Other (*n* = 11) was excluded.

Table 3.1-10 gives the distribution of tense and aspects of the reflexive verbs in the database, excluding the small class Other. The biaspectual reflexive verbs were disambiguated based on the context. Systematic data on the connection between aspect and the Reflexive Marker is lacking. Dankov (Данков, 1981:66) states that it is generally known that the connection towards the perfective aspect is stronger in contemporary Russian compared to Old Russian. In contrast, the encoding used for the neighbor verbs differs from the reflexive verbs. The primary reason is that the neighbor verbs are not used to model derivation but cross-paradigmatic support. Thus, tense is excluded and aspect is not disambiguated.

	Aspect			
	Biaspectual	Imperfective	Perfective	Sum
Neighbor Verb	29	401	287	717

Table 3.1-11 Distribution of Aspect of the unique neighbor verbs (*n* = 717).

Table 3.1-11 gives the distribution of the aspect of the unique neighbor verbs in the database. This section introduced the fundamental morphological categories of the Russian verbs, namely tense and aspect, and their encoding in the database.

3.1.12 Summary: Lexical Networks

This section offers an intermediate summary of the previous discussion and anchors the operationalized variables relative to the gradient and dynamic

[63] The subject argument is not necessarily indexed between the verbs in all construction types. Especially command verbs, when combined with the infinitive, do not typically index the subject argument of the main verb.

structure of the lexicon at various levels of schematicity. Additionally, the last domain-general principle is formulated that is intended be at work in the process of category formation in conjunction with the Hypothesis of Connectivity and Distance, but before that, the operationalized relations are exemplified. Figure 3.1-7 represents the relations investigated in this study.

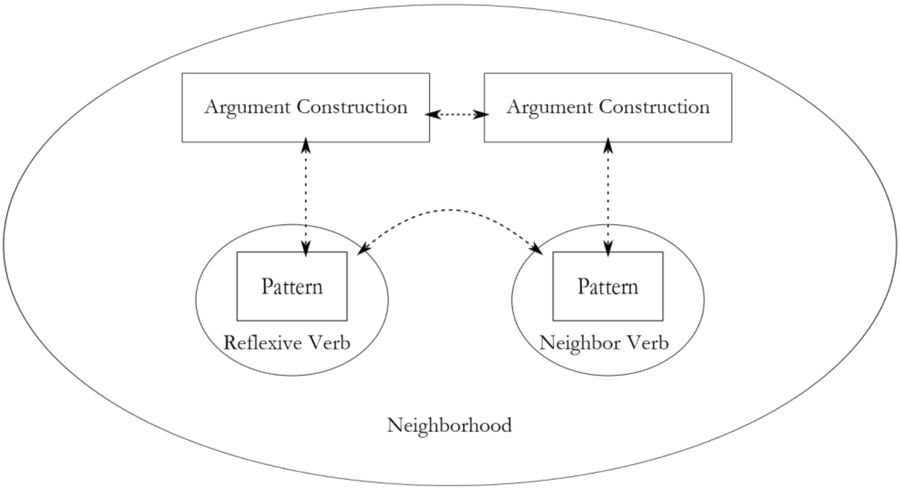


Figure 3.1-7 Idealized relations at different levels of schematicity.

From a usage-based perspective, the lexicon is an integral part in the process of category formation. As this study focuses on the relations of the verb, the lexicon is carved, represented by the outer circle and defined here as the Neighborhood. Additionally, it is argued that the lexicon is structured and not just a repository of items (Bybee, 2010). The Neighborhood was operationalized with the rhyme densities. The Reflexive and the Neighbor Verb further partition the Neighborhood forming paradigms. Consequently, a specific instantiation will partition its own niche within its paradigm represented with the inner circle. Another consequence of the structured view on the lexicon is that the paradigms are intertwined, forming cross-paradigmatic relations represented with the horizontal dashed line connecting the Reflexive and the Neighbor Verb. The arrows are indented to convey the non-directionality of the relations, but it does not deny the possibility that certain relations may be more prone to directionality. However, these are more likely to be locality effects, differences between individual items (Hay, 2001; 2002).

The cross-paradigmatic relation is not static or rule-governed, but shaped by language use and the structural properties of the Neighborhood (Bybee, 1985). The gradient cross-paradigmatic relation is modulated through the Neighborhood Density, the Neighborhood Distance, frequency of use, and the perceived semantic similarity, labeled here as the variable Reflexiva Tantum. Additionally, the cross-paradigmatic relation can also be modulated through semantic components of a particular subtype. The variable Causation represents such a subtype where it is an inherent property of the Neighbor Verb, but due to the cross-paradigmatic relation it may be subjected to conceptualization, leading to a modulation of the cross-paradigmatic relation. Importantly, this relation constitues a form of semantic connectivity between the subtype and is not a derivational rule, (i.e., the Causative → Decausative derivation).

Another component is the abstraction over the usage of a particular verb referred to as the Pattern. A particular usage pattern is connected to other grammatical categories in a language. Thus far, the model includes the variables Tense and Aspect to factor in such impacts. Additionally, repeated usage and perceived similarity between patterns lead to Argument Constructions, a high-order abstraction supported by more specific instantiations. However, the Argument Constructions have a similar status as lexical items, but differ in schematicity (Croft, 2001). Thus, the Argument Construction is readily available and can be used as such without reinventing a particular type of abstraction (Goldberg, 2006). This relation is depicted with the dashed vertical line in Figure 3.1-7. The support originating from a particular verb is subjected to the same gradient structure measured in this study with the variable Constructional Entropy. The final component is the connectivity between Argument Constructions leading a network structure similar to lexical items measured here through the Neighborhood Density. It is a topic discussed in Section 11.3 with the focus on the network structure of the Reflexive Marker.

The previous synopsis anchored the proposed variables to the usage-based view of language. As linguistic categories are assumed to be highly intertwined, the network model is an attempt to capture some aspects of it. A network has internal structure (Bybee, 1985) and it is density-based (Baayen & Moscoso del Prado Martín, 2005; Geeraert & Kyröläinen, in prep.).

The results obtained from the lexical network structure were postulated to be governed, at least, by two domain-general principles, labeled as the Hypothesis of Connectivity and the Hypothesis of Distance. They are domain-general in a sense that they are a consequence of the network structure. They are not a language-specific property, but can be observed in any structure that follows the network model (Arbesman, Strogatz & Vitevitch, 2010; Steyvers & Tenenbaum, 2005). The results thus far have shown that most of the reflexive verbs have a neighbor verb in Russian, forming a cross-paradigmatic relation. This relation can be viewed as the primary motivational pathway in the formation of complex categories.

These cross-paradigmatic verbs appear to have proximately equal neighborhood densities (Section 3.1.2) and be perceived as semantically similar (Section 3.1.4). Their frequency distributions are similar (Section 3.1.7) and this property extends to larger subtypes, such as the verbs connected within their neighborhood though the causative component and transitivity (Section 3.1.10). However, the differences appear to lie in the internal structure of these paradigms, namely the reflexive verbs tend to be more loosely connected based on the estimation obtained with the Neighborhood Distance (Section 3.1.3).

In sum, this section anchored the proposed set of variables to the theoretical basis of usage-based models. The underlying properties of the lexical network are critical, as verbs, verb-specific constructions, and argument constructions differ only in schematicity. Thus, they are assumed to be governed by the same principles. Importantly, we have established the degree of connectivity between items, using the variables of Neighborhood Distance and the strength of connectivity through frequency of use and constructional entropy.

3.2 Primary and Secondary Slots

This section establishes the basic configurations and encodings for the primary and the secondary slots of verbs. The term primary slot refers to the subject argument of the argument construction and the secondary slot to the non-subject. In this vein, the following sections assess the possibility of forming abstractions over observed usage patterns. Another motivation is to include morphology, especially cases, as part of the analysis. Surprisingly, patterns are more or less absent from the whole research paradigm of the Russian Reflexive Marker. An exception to this is Gerritsen's (1990) study.

Reasons for the lack of interest towards case patterns are difficult to pin down. Perhaps one reason is its range. Case patterns subsume both the idiosyncratic properties of individual verbs and more systematic patterns. At the same time, the case patterns offer a degree of cue validity when several usage patterns are contrasted. The following sections establish the encoding of the primary and the secondary slots of the verbs used in this study.

3.2.1 Profiles: Cases and Patterns

The inclusion of morphology, such as case, has become increasingly dominant within the Construction Grammar and usage-based approaches. Kempe and MacWhinney have shown that Russian speakers rely on case-marking in on-line sentence interpretation, using the picture-choice paradigm contrasting Russian and German speakers. Participants heard a transitive sentence while making a choice between two possible agentive referents displayed visually. The results showed, in terms of reaction times, that the Russian speakers relied more on the case-marking and the German speakers on the animacy (Kempe & MacWhinney, 1999). Additionally, the implementation of the variables Case and Pattern follows the usage-based approach. Furthermore, the division between structural and lexical case upheld in formal theories is not utilized (cf. Barðdal, 2011). Instead, the concept of profile is employed.

Goldberg defines the concept of the lexical profile in relation to the participant roles of the verb that are obligatorily expressed. In Goldberg's proposal, the profile is the lexical material that anchors the form and meaning relation (cf. Goldberg, 2006:39-40). The variable Profile receives a broader interpretation in study and it is properly discussed in Section 3.2.3. It follows from this that the profile allows to capture the variation of the verb-specific constructions illustrated in 3.2-1 and 3.2-2.

3.2-1 [*з*]*вер-и* *оказа-л-и-сь* *очень* *хорош-ими* *воспитател-ями.*
beast-NOM.PL appear-PST-PL-RM very good-INS.PL tutor-INS.PL
The beasts appeared to be very good tutors.
[1708, RNC, Сергей Седов. Доброе сердце Робина //
"Мурзилка," №7," 2002]

3.2-2 *Потом* *оказа-л-о-сь,* *что* *эт-о*
Then appear-PST-N-RM that this-NOM
более *сложн-ая* *проблем-а.*
more complex-NOM problem-NOM
Then it turned out that it is a more complicated problem.
[1435, RNC, Лекция (отрывок), Москва // (2005)]

Example 3.2-1 profiles the Property Construction following its canonical form pole, namely the Nominative-Instrumental pattern, whereas the same lexical item *оказаться* ‘seem, appear’ is used to profile the Content Construction, cf. Section 3.2-2, where the primary slot is profiled with the clausal subject, *что* ’that.’ The differences in the profile are indicators of different argument construction types. Hence, generalizations are stated based on the surface structure and its profile.

We are faced with a granularity effect, once again, when utilizing case-marking in Russian. Kopotev (Копотев, 2008) points out that a generally accepted list of the Russian cases does not exist (cf. Corbett, 2011).[64] I refer to Kopotev (Копотев, 2008:140-141) for the discussion on the various case systems. For the purposes of the present study, the traditional system advocated in Russian Grammar is used (Шведова & другие, 1982:474-475) with the following cases: nominative, accusative, genitive, dative, instrumental, and prepositional. The second granularity effect concerns the exact profile of an observed pattern. If we take an exact form-based approach, different combinations of prepositions and cases should be included in the model. Considering the sample size and the number of verbs in the database, the exact form-based approach is infeasible. A slight case of abstraction is utilized by introducing two labels. They are the bare case marking, and the non-bare case marking. These labels are also used in Janda’s (1993b) study on Russian case. Non-bare cases are indicated with the encoding P(X), where the P stands for any preposition and the X is the case. For instance, a verb appearing with the argument k_{dat} is encoded as PD. Table 3.2-1 summarizes the encoding.

[64] A system of 12 cases is used in the Russian National Corpus.

Bare	¬ Bare
Nominative	*
*	Accusative
Genitive	Genitive
Dative	Dative
Instrumental	Instrumental
Prepositional	Prepositional

Table 3.2-1 Encoding of the case in the database.

The nominative case only appears in the bare type because all the verbs attested in the sample are intransitive. Thus, there are no instances of the bare accusative. The cases are encoding for both the primary and secondary slots.

Another important factor arises from the morphological cues, namely patterns. Tomasello (2003:29-31) considers that the ability to find patterns one of the fundamental prerequisites for language acquisition. The role of repeated usage patterns figures prominently in Construction Grammar (Goldberg, 2006) and in corpus-driven approaches, as described by Hunston and Gill (2000), and Sinclair and Mauranen (2006). As instantiations are encoded in terms of slots, their combination leads to observed patterns. Typically, case patterns are encoded as idealizations, as possible linear order is not considered (Barðdal, 2008; Divjak & Janda, 2008; Kyröläinen, submitted). A similar position is already taken by Zolotova (Золотова, Г. А., 2005 [1973]), where the combination of case leads to the core sentence types in Russian. Thus, we arrive at something similar to case constructions as they are labeled by Barðdal (2008), albeit minimal profiles, as the patterns are restricted to the combination of the two slots.

However, the concept of pattern is connected to theoretical stipulations. Faulhaber explicates this with English prepositional verbs, such as *niggle at* and *grate on*. These verbs appear as a highly systematic class in Levin's (1993) classification simply due to the fact that the category of prepositional verb is assumed, (i.e., the prepositions are part of the predicate). In contrast, a high degree of variability is introduced if the preposition is analyzed separately from the verb (Faulhaber, 2011:282). Thus, the concept of pattern is disassociated from grammatical function in this study, illustrated in 3.2-3 and 3.2-4.

3.2-3 *Поэтому* *данн-ая* *стать-я* *явля-ет-ся* *незавершённ-ой*, […].
Therefore this-NOM article-NOM be-3S.PRS-RM incomplete-INS
Therefore, this article is incomplete.
[299, RNC, В.В. Ахияров. Гравитация в Солнечной системе // "Геоинформатика," 2002.03.20]

3.2-4 *Исаак* *Ньютон* *увлека-л-ся* *астрологи-ей.*
NAME.NOM NAME.NOM fascinate-PST.M-RM astrology-INS
Isaac Newton was fascinated by astrology.
[95, RNC, Homo играющий: Было или не было? // "Знание — Сила," 2003]

Example 3.2-3 illustrates the copula verb *являться* 'be,' where the adjective functions as a predicate, yielding a pattern of Nominative-Instrumental. In contrast, if only participants are considered relevant, the pattern is simply the Nominative. Similarly, treating the instrumental case as an oblique argument in 3.2-4, outside the core predication, the pattern is simply the nominative case. Thus, we would seem to have arrived at intransitivity. Still, the case patterns would have to be mapped to the verb in order to yield the sentence pattern. In formal approaches, this might be considered as part of the lexicon and not grammar (cf. Grimshaw, 1990; Pinker, 1989; Pinker & Ullman, 2002).

Certainly, there is a difference between a pattern and a grammatical function (cf. Sapir, 1955 [1921]:59-60). This fact is not denied. Instead they are viewed as separate layers. Example 3.2-5 illustrates the issue.

3.2-5 *Волчиц-а* *оказа-л-а-сь* *бешен-ой.*
She.wolf-NOM appear-PST-F-RM rapid-INS
The she-wolf appeared rabid.
[517, RNC, Темниковские охотники отстреливают лис и волков //"Московский комсомолец" в Саранске," 2004.12.23]

The proposed analysis aligns partly with the dependency-based grammar implemented in the syntactic subcorpus of the Russian National Corpus. Figure 3.2-1 gives the dependency tree structure of Example 3.2-5 parsed with the ETAP 3 parser.[65] The syntactic model of the parser is based on Meaning—Text theory and it is a rule-based model (Apresian, Boguslavsky, Leonid Iomdin, Lazursky, Sannikov, Sizov & Tsinman, 2003). The parser is not publically available but the on-line demo version was used, located at http://proling.iitp.ru/.

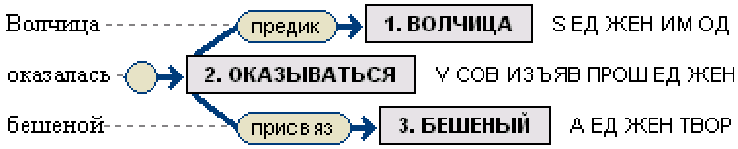


Figure 3.2-1 Dependency tree structure of Example 3.3-1 based on ETAP 3 parser.

The reflexive verb *оказаться* 'seem, appear' has a dependency relation to *бешеной* 'rapid' and the grammatical function of copula (*присвязочное*) is the type of the relation, as indicated in the tree rather than imposing the constituency-based analysis consisting of Noun Phrase and Verb Phrase.

The combination of the layers yields grammatical roles discussed in Section

[65] The basic grammatical information is provided in the parsed tree in the following order: *волчица* - noun, singular, feminine, nominative and animate, *оказаться* - verb, perfective, indicative, past, singular and feminine, and *бешеной* - adjective, singular and instrumental.

3.2.3. The rationale behind the variable Pattern is to account for the possible form-based similarities among the various construction types. A possible type-effect or support originating from certain repeated combinations similar to Goldberg's (2006:105-126) hypothesis is that both forms and function may be subject to generalization. Goldberg demonstrates that constructions may be better predictors for clausal meaning compared to verbs in certain situations as in 3.2-6 and 3.2-7.

3.2-6 *Pat got the ball over the fence.*
get + VOL pattern → "caused motion"
(Goldberg, 2006:106)
3.2-7 *Pat got Bob a cake.*
get + VOO pattern → "transfer"
(Goldberg, 2006:106)

The verb *get* appears in the VOL pattern (Subject Verb Object Oblique) instantiating the Caused Motion Construction, whereas the VOO pattern (Subject Verb Object Object$_2$) conveys the Transfer. Simply knowing the verb does not necessarily entail that the clausal meaning can be predicted. This prediction strength is estimated as cue-validity by Goldberg (2006:105-107).

The formation of patterns rests on the assumption that generalizations are formed over re-occurring strings, such as *Волчица оказалась бешеной* 'The she-wolf appeared rabid,' leading to a generalized pattern of Nominative-Instrumental (cf. Bod, 2009:130-131; Bybee, 2010:25, 34-37). Thus, argument constructions are generalized over these re-occurring patterns.

3.2.2 Referents and Encodings

Animacy figures prominently in linguistic description of argument structures connected both to subjecthood and the description of semantic roles. Additionally, animacy is a central component in diathesis formulated by Geniušienė (1987). Bock et al. demonstrate that animacy is a determining factor in subject/object selection in English. Animate subject arguments occur more often with verbs that allow both animate and inanimate subjects, contrary to object selection. They also show, on the basis of a priming test with the passive construction, that the status of a subject is best understood as a direct mapping between arguments rather than an underlying structure (Bock, Loebell & Morey, 1992:154-159, 162).[66] Crucially, this can be taken as indirect evidence in favor of the syntactic role position adapted in this study.[67]

[66] The rationale behind the priming test is that if the subject slot is indeed a by-product of an underlying structure, it should show a priming effect with the underlying object-arguments, (i.e., a priming effect with the surface objects of the active and the surface subject of the passive). According to the direct-mapping hypothesis, the priming effect should follow the animacy of the primes and not of the underlying structure.

[67] The inclusion of the referent type partly mimics the concept of the subject- and

Several hierarchies have been proposed in the literature. For example, Comrie (1989:184) proposes the following: human > animal > inanimate. Silverstein's (1976) animacy hierarchy contains pronouns also. A commonality between various animacy hierarchies is found in the contrast between human versus other types. Garretson et al. (2004) have developed an animacy coding hierarchy consisting of three tiers with subtypes yielding a set of seven categories: top > human, middle > animal and organization, and bottom > concrete inanimate, non-concrete inanimate, place, and time. Considering the sparseness of the data in terms of different construction types, a simplified encoding of the variable Referent is used in this study: human, animate, inanimate, and abstract. The counts are given in Table 3.2-2. On a coarse level of granularity, the referent types appear to be centered on two poles: Person and Abstract.

	Abstract	Animate	Inanimate	Person	Sum
Instance	779	107	238	865	1989

Table 3.2-2 Distribution of the Referent Layer of the primary slot. The type Other (*n* = 11) is excluded.

The variable preserves the possible importance of the human referent versus others. Additionally, the referent type is considered in context and not in isolation as in 3.2-8.

3.2-8 *Когда подходило время сна*,
[…]
Эле-Фантик на одн-о ух-о ложи-л-ся,
NAME.NOM PR one-ACC ear-ACC lie.down-PST.M-RM
а друг-им укрыва-л-ся.
and other-INS cover-PST.M-RM
When it was time to sleep, Ele-Fantik lied down on one ear and covered himself with the other.
[1573, RNC, Александр Дорофеев. Эле-Фантик // "Мурзилка," №1-5," 2003]

This and similar ones were encoded as instances of Person, as they are characters in stories. Another difficulty is associated with the possible distinction between organizations or groups of people as in 3.2-9.

3.2-9 *Даже после слияния ПВО и ВВС*
[…]

object-oriented verbs advocated in the diathesis tradition. Typically, the subject argument is human or, at least, animate and the object argument is nonhuman.

войс-ка *противовоздушн-ой* *оборон-ы*
force-NOM.PL anti-aircraft-GEN defence-GEN
оста-ют-ся *важн-ой* *составляющ-ей*
remain-3P.PRS-RM important-INS component-INS
вооружен-ых *сил* *Росси-и.*
armed-GEN.PL force.GEN.PL Russia-GEN
Even after the merging of Air Defence and Air Force, air forces remain an important component of Russian military.
[32, RNC, Саид Аминов. История побед и поражений // "Воздушно-космическая оборона," 2001.04.15

These can be viewed either as a collective group of representatives or an abstract entity. All these instances were encoded as abstract. Additionally, if an infinitive appeared in either of the slots, it was tagged as abstract. Finally, the encoding of the Referent is applied to both slots considered in this study.

Another complication associated with the profile is the actual linguistic encoding. Knyazev (Князев, 2007:169-170) considers that Russian displays "love towards zeros."[68] The non-instantiation of the primary and secondary slots comes in varying shades of covert encoding connecting these types to a larger body of discourse structure. The indexing potentiality of arguments, especially its relation to subjecthood, is discussed by Kyröläinen building on different instantiation types and indexing potentiality across conjoined argument construction types in Russian (Kyröläinen, submitted). On the other hand, Goldberg (2006 Chapter 9) connects argument omission to general discourse structure, specifically to pragmatic principles. Similarly, Helasvuo and Kyröläinen have demonstrated that the encoding of the pronominal nominative subject, either as overtly or covertly, is fairly predictable based on discourse structure in conversational data in Finnish (Helasvuo & Kyröläinen, 2010; 2011).

These findings support the view that discourse has a global, predictable structure. If this was not the case, achieving predictive accuracy would not be possible. Thus, the encoding would be arbitrary (Bresnan & Ford, 2010; Genzel & Charniak, 2003). This position is a stark contrast to the analysis proposed in formal approaches. Perlmutter and Moore claim that the nominative pronoun system consists of two series in Russian. The A-series covers nominative pronouns with phonological shape and the B-series contains the silent ones, (i.e., without phonological shape, for example, A-Series: *он* 'he' → nominative third person singular masculine and corresponding B-Series: NULL → nominative third person singular masculine).

Traditionally, the non-instantiation of the subject argument, (i.e., the primary slot), is assumed to be a genre-specific property in Russian. Nonetheless,

[68] The noninstantiation is not limited to arguments, but also covers verbs in Russian. Because the sampling frame is based on the Reflexive Marker, the noninstantiation of the reflexive verbs does not appear in the data.

quantitative studies of the phenomenon are rare (cf. Zdorenko, 2010 and references therein). Seo compares the omissability of the subject arguments in Russian, Polish, Czech, Bulgarian, and Serbo-Croatian. The data are based on five novels covering 2,000 instances, mostly in dialogues. The results suggest that Russian is the most conservative among the studied languages; 22% omitted subjects versus 80-90% in others (Seo, 2001). Table 3.2-3 gives the counts of the encoding type of the primary slot either overtly or covertly in the database.

	Covert	Overt	Sum
Instance	214	1775	1989

Table 3.2-3 Distribution of the Encoding Layer of the primary slot. The type Other (*n* = 11) is excluded.

The data confine to the previous findings that the encoding of the primary slot (subject) leans towards the overt type in Russian in general. Thus, the variable Encoding covers the instantiation of the usage patterns in discourse.

3.2.3 From Syntactic Relations to Syntactic Roles

Traditionally, the definition of subjecthood is established through syntactic tests, which are assumed to target such categories as subject, indirect object, and extending to possible dative and genitive subject (cf. Jakobson, 1989 [1936]:72-73). The tests are also assumed to indicate the universal properties of subjecthood to the point where subject is a universal category, hard-wired into the brain as in Pinker (1989) and Grimshaw (1990). However, the meaningfulness of employing such tests depends on the definition of their targets, the syntactic relations. This leads to circularity, albeit, not necessarily a vicious one, in that a test is devised to establish the category subject. The meaningfulness of applying such a test depends on the category it is supposed to target. This is also reflected on the labels of the tests, like raising to subject and raising to object.

A further complication of the status of grammatical relations is that the different tests typically create different results and are highly dependent on a construction to which they are applied. Additionally, the tests are not applied equally among different possible subject candidates leading to a discrepancy between different theoretical approaches (Barðdal, 2004; 2006; Barðdal & Eythórsson, 2003; Croft, 2001; Eythórsson & Barðdal, 2005). In contrast, Russian linguistic tradition typically operates on a two-tiered system positing semantic subject, *субъект*, and a grammatical subject, *подлежащее*. However, no criteria are offered on how to exactly and ambiguously establish these two categories (Leinonen, 1985; Бондарко, А. В., 2002; Золотова, Г. А., 2000a; Золотова, Г. А., Онипенко & Сидорова, 1998).

Leinonen (1985:15) uses these notions following the Russian linguistic tradition in her contrastive study on Finnish and Russian impersonal sentences.

The term, *подлежащее*, is defined as a syntactic unit forming the nucleus of the sentence. In contrast, the *субъект* is a semantic notion, the carrier of the quality (Золотова, Г. А. , 1981). In principle, a two-tiered system allows to maintain the traditional criteria of subject and also to include non-canonical subjects.

Generally stated, the status of the oblique, or quirky subject is a controversial issue and its definition is theory-dependent. There are a number of reasons leading to this discrepancy between different approaches. In formal approaches to Russian, the status of the dative argument has been a debated topic. Interestingly, agreement on the status of dative subject is not unified even among the formal approaches (Greenberg & Franks, 1991; Moore & Perlmutter, 2000; Perlmutter, David & Moore, 2002; Perlmutter, David M., 1983; Zimmerling, 2009). Kyröläinen demonstrates, using ten different construction types in Russian with 20 commonly posited features, that the mismatch between the features and the selected construction types is highly construction-specific. For example, if a global category of indirect object exists in a language, it should display identical behavior across different construction types (Kyröläinen, submitted). Example 3.2-10 illustrates the Ditransitive Construction and Example 3.2-11 the Subjective Experiencer Construction with an idiomatic reflexive pronoun construction.[69] Both construction types are conjoined with the reflexive pronoun *себя* ‘oneself.’

3.2-10 *Я*$_i$ *написа-л* *Борис-у*$_j$ *длинн-ое*
I$_i$.NOM write-PST.M NAME$_j$-DAT long-ACC
письм-о *о* *себ-е*$_i$/$_{*j}$.
letter-ACC PR self-PREP$_i$/$_{*j}$
I wrote a long letter to Boris about myself/*himself
[(Moore & Perlmutter, 2000:379)

3.2-11 *Борис-у* *не* *работа-ет-ся* *у* *себ-я* *дома*.
NAME-DAT NEG work-3S.PRS-RM PR self-GEN at.home
Boris can’t seem to work at his own place (at home).
[(Moore & Perlmutter, 2000:378)

To rectify these mismatches, Moore and Perlmutter (2000:375) posit that the dative in 3.2-11 is Inversion nominal, (i.e., the demoted subject becomes an indirect object in the surface structure). In contrast, the recognition of argument constructions, at least partly, solves the issues of the rampant mismatches. In 3.2-10, the nominative argument takes precedence over the dative, whereas the dative is indexed in 3.2-11.

The datives do not constitute a grammatical relation of indirect object in Russian. Instead, they are construction specific roles. The differences follow from the argument constructions and not from some underlying structure. Because of these mismatches or contradictory results, Croft takes a radical stance by denying the status of grammatical relations. Thus, he posits that constructions contain grammatical roles and these are construction-specific

[69] The glossing was added to the examples by the author.

properties. In this manner, the syntactic categories and traditional labels, such as subject, object, and indirect object are defined in relation to the construction as a whole and not to the predicate which would constitute a part-part relationship. This position is upheld within constructionist approaches by Barðdal (2006), Divjak and Janda (Divjak & Janda, 2008), and Kyröläinen (Kyröläinen, submitted). For example, Kyröläinen proposes a clustered model for establishing subjecthood in Russian (cf. Dowty, 1991; Keenan, 1976).[70] When typical tests for subjecthood are systematically contrasted across argument construction types, the traditional properties of the canonical subject emerge from the patterns in Russian, namely agreement, nominative case and personal versus impersonal construction type. (Kyröläinen, submitted).

Section 3.2.4 introduces the canonical subject construction for the Russian reflexive construction types, which allows establishing the deviations that yield the non-canonical subjects. Section 3.2.5 discusses the non-canonical subject roles and offers a possible model to differentiate them based on distributional properties. The extension from the canonical subject roles to non-canonical are established through the account proposed by Divjak and Janda (2008). Thus, these sections are used to establish the properties of two schematic construction types in Russian, namely the personal and the impersonal construction.

3.2.4 The Canonical Subject Role

The intersection between a canonical and a non-canonical subject is established through positing a set of properties, which can be used to separate subjects from non-subjects. In nominative-accusative languages, the case marking is taken as one of the crucial factors in defining subjecthood in functional and cognitive approaches. A second crucial property is agreement between the subject candidate and the predicate. In Russian, this canonical pattern is attested in person, number, and gender (for past tense). This yields the traditional definition of the canonical subject and forms a schematic category of personal construction type.[71] Kyröläinen demonstrated that the proposed properties of subjects (*m*=20) are systematically contrasted across construction types (N = 10). These three properties form a distinctive cluster separating them from all other properties, labeled as the Primary Cluster (Kyröläinen, submitted). This follows the traditional definition, as in the Russian Grammar (Шведова & другие, 1982:480-481). This position is also commonly upheld in studies of the Russian Reflexive Marker (cf. Gerritsen, 1990; Israeli, 1997). In Radical Construction Grammar, the status of subjecthood is defined in relation to the construction as a whole. This has implications especially for the non-canonical subject roles and these are discussed in the following section.

For the purposes of this study, the canonical subject slot can be defined in

[70] More recently, Ackerman and Moore (2009) propose a model which also builds on the notion of clustered properties.

[71] There are obvious deviations even within the canonical personal types. These deviations are discussed by Kyröläinen (submitted and references therein).

terms of a schematic personal reflexive construction type, where Intransitive Construction follows the pattern Nominative Subject and Reflexive Verb. The instantiations of the reflexive constructions are stipulated against this background. Table 3.2-4 gives the distribution of this schematic patterning in the database. As all the attested reflexive infinitive forms were personal, they are encoded based on the main verb. The distribution follows the expected tendency that the majority of the Reflexive Construction types pertain to the schematic personal type.

	Personal	Impersonal	Sum
Instance	1865	124	1989

Table 3.2-4 Distribution of the personal and impersonal types in the database. The type Other (*n* = 11) is excluded.

Additionally, the concept of the personal reflexive construction type is tied to the grammatical category of person. Siewierska (2004) considers that the category of person is used to characterize the roles of the discourse participants and not just simply participants moving away from the traditional tripartite structure: speaker as the first person, the addressee as the second person, and the referents discussed about as third person (cf. Jakobson, 1989 [1932]). However, the possible influence of discourse structure or contextual factors are not considered in this study, the traditional definition of person as delimited to the inflectional category of person marking in verbs is sufficient.

The centre of the category person in relation to personal pronouns is considered to constitute a hierarchical structure. The first and the second person constitute the core of the category of person whereas the third person is related to the periphery along with the various impersonal types (Бондарко, А. В., 1991:19-20).[72] Typically impersonal forms are considered to be part of the category person in the Russian linguistic tradition. For example, Jakobson (Якобсон, 1985a:215) states that the impersonal patterns belong to the so called third person category from a grammatical perspective. Similarly, the impersonal forms can be analyzed as deviations from the personal type and only contain the default agreement pattern third person singular/neuter.

In terms of person marking, a further complication arises due to the fact that the past tense forms do not have a morphological person marker. Instead the gender is used. The inclusion of this category as part of the person paradigm creates a unique combination for the past forms in terms of modeling purposes. Considering the small sample size (*N* = 2,000) and the skewed distribution of

[72] Bondarko states that the personal form of the verb and the personal pronouns establish the centre of the category of person and not separate components. Thus, a division between grammatical and pragmatic functions are established. In contrast, Siewierska (2004:14) considers that a person marker appears in both nominal and verbal domains, and additionally, at various levels whether a phrase, a clause, a sentence, or a text.

tense across different construction type, the data would be highly partitioned. When the type "Other" is excluded, the past tense covers 43% (*n* = 857) of the data points. Additionally, the inclusion of gender as a variable in person marking is related to the question of granularity (Janda & Lyashevskaya, 2011). The main motivation behind the encoding schema of person marking is to tease apart the possible connection between the encodings of subject and verb in this study. Thus, gender is not considered and the past tense forms were manually tagged for Person. Table 3.2-5 gives the distribution of the variable Person in the database.

	Person						
	1S	2S	3S	1PL	2PL	3PL	Sum
Instance	85	39	1327	57	30	451	1989

Table 3.2-5 Distribution of the person marking in the database. The type Other (*n* = 11) is excluded.

The distribution is highly skewed towards the third person singular (*n* = 1337) and plural (*n* = 452) forms. In terms of person marking, the Reflexive Marker appears to be centered on the third person. This property seems to be fairly stable, as Dankov's (Данков, 1981:66) diachronic study has shown that the Reflexive Marker is primarily used in the third person. Certainly not an unexpected distribution considering the categories typically attributed to the Russian Reflexive Marker, such as the Passive (Section 5.1) and the Spontaneous Event (Section 5.3), are typically tied to the third person. Moreover, this distribution can be assumed to follow the underlying population of the reflexive verbs because the Reflexive Marker was used to form the database to minimize any preselection of the data. Considering the overall distribution, the data points are, however, collapsed into four categories in order to make the possible distinction between the third person versus first and second person. The levels are: combined 1S and 2S (*n* = 124), and combined 1PL and 2PL (*n* = 87), leaving the third person singular (*n* = 1,327) and plural (*n* = 451).

3.2.5 Non-Canonical Subject Roles

The status of non-canonical subjects is a controversial matter in Russian and status of these roles is highly theory dependent. This section illustrates one possible solution of establishing the non-canonical subject roles in Russian based on distributional properties of the verb-specific constructions and their interaction with argument constructions. Specifically, this section is devoted to the relationship between dative or oblique subjects in contrast to possible infinitive or subordinate clause subject roles, namely *что*-clauses, that-clause. Additionally, this discussion is limited to the verbs which are attested with these patterns in the data base.

Shahmatov (Шахматов, 1925) gives three groups of impersonal types stating that there might be more. A more comprehensive list is offered by Vinogradov

(Виноградов, 1972:370-371) consisting of eight different types:

1) Verbs denoting existence and state.
2) Verbs denoting phenomenon in nature.
3) Verbs denoting mystical process.
4) Verbs denoting fate.
5) Expression of inner physical state or inner physical change.
6) Expressions of depicting conceptualization of surrounding phenomena.
7) Verbs denoting undergoing a physical sensation.
8) Verbs denoting someone's ability towards activity or activity towards subject.

Vinogradov explicitly connects the last type (8) to the Reflexive Marker. This construction type is labeled as the Subjective Perspectivization in this study, cf. Section 9.4.

The early Russian tradition focused on establishing the semantic basis of the impersonal types. A study by Galkina-Fedoruk is still perhaps the most comprehensive study of Russian impersonal types, covering both non-reflexive and reflexive verbs. Additionally, nominal predicates are also included. The following definition is given for the impersonal type. Impersonal is a construction without subject and with one primary member, the predicate. The predicate is in a form, which does not express person and there is no person in the current context. Thus, this definition hinges on the assumption that there is a stark contrast between a third person personal form and third person impersonal. At the same time, a more dynamic view is also presented by Galkina-Fedoruk (Галкина-Федорук, 1958:126). When interpreted from a usage-based perspective, it could be perhaps stated in the following manner. Certain verbs can be used either in personal or impersonal construction and frequent usage over time in an impersonal type may lead to a stronger association with it and, ultimately, to reduced personal paradigm. This characterization avoids positing a homonymy analysis for these verbs. This is a basic tenet in constructionists approaches (Goldberg, 1995:10-11). Interestingly, a similar argumentation is also presented by Galkina-Fedoruk (Галкина-Федорук, 1958:126-127).

In terms of subjecthood, Bondarko considers that the dative includes both the properties of a subject and an indirect object, yielding the traditional account of two-tiered subjects. Nominative subject constitutes the first-order subject while the dative is the second-order (Бондарко, А. В., 2002:639). However, the question remains whether all datives are equal in terms of being the second-order subject. Table 3.2-6 gives the reflexive verbs that appeared in the impersonal construction types in the database.

		Pattern			
Reflexive Verb	translation	nominative	dative	infinitive	that
вериться	believe		###		###
выразиться	express	###	###		###
выясниться	emerge		###		###
говориться	say		###		###
довестись	happen, manage		###	###	
доводиться	happen, manage	###	###	###	
житься	live		###		
захотеться	start to want		###	###	
иметься (в виду)	have, mean				###
казаться	seem, appear	###	###		###
мечтаться	dream		###	###	###
нравиться	please	###	###	###	###
обнаружиться	come out	###			###
оказаться	turn out, appear	###	###		###
оказываться	turn out, appear	###	###		###
оставаться	remain	###	###	###	
остаться	remain	###	###	###	
подразумеваться	imply	###			###
показаться	turn up, appear	###	###		###
полагаться	suppose		###	###	###
понадобиться	necessary	###	###	###	
предлагаться	suggest	###	###	###	###
предписываться	order, prescribe		###	###	###
предполагаться	intent	###	###	###	###
представляться	appear, arise	###	###	###	###
прийтись	happen, kinship	###	###	###	
приходиться	happen, kinship	###	###	###	
разуметься	mean	###	###	###	###
случаться	happen	###	###	###	###
спрашиваться	ask	###			###
считаться	consider, kinship	###	###	###	###
удаваться	succeed	###	###	###	
удаться	succeed	###	###	###	
указываться	point out	###	###		###
хотеться	want		###	###	
чувствоваться	feel	###			###
чудиться	seem	###	###		###

Table 3.2-6 Patterns available for the reflexive verbs conjoinable with impersonal construction types. The hashes indicate presence of the pattern.

Certain reflexive verbs, such as *мерещиться* 'seem, appear,' can be used either in a personal or an impersonal construction type, but only the personal one is attested in the database. Additionally, Table 3.2-6 includes the potential case patterns of the verbs indicated with the hashes bringing forth the divergent nature of these verbs. Given the small number of instances, the additional patterns available for a given Reflexive verb were compiled from the following dictionaries: Kuznecov (Кузнецов, 2009 [1998]), Denisov and Morkovkin (Денисов & Морковкин, 2002), and Daum and Schenk (1968).

The given patterns portray the situation in a simplified manner. Other possible oblique arguments are not considered, such as, instrumental case with the verb *оказаться* 'seem, appear,' *у*$_{\text{gen}}$ prepositional phrase with the verb *случаться* 'happen,' and, finally, *чтобы* 'that' for purpose or subjunctive subordinate clause. The latter option is possible with the verbs *захотеться* 'start to want' and *хотеться* 'want.' Even with this small set of verbs, the patterns available for a particular verb are highly divergent. From a lexical perspective, the classification of these verbs will depend on the exact range of the patterns recognized in a study leading to different classifications.

The whole range of patterns available for a verb cannot be utilized simultaneously. For example, the verb *приходиться* 'happen, kinship' cannot display all four slots in one expression. From a lexical perspective, some of these verbs are classified as polysemous and others as homonymous depending on whether a non-reflexive verb can be found with similar meaning. Thus, it is sometimes argued that the presence of the Reflexive Marker is accidental or irrelevant for the formation of the impersonal construction type if the non-reflexive and the reflexive verb differ in meaning. This position is clearly stated, for example, by Israeli (1997:130).

The verb *приходиться* can be used to depict kinship, profiled with the nominative, dative, and instrumental pattern. In contrast, the non-reflexive verb profiles the motion domain, (i.e., 'come, arrive'). However, this mode of analysis looses the sight on the overall architecture of the Reflexive Marker as a category and posits a strict correspondence between the reflexive and non-reflexive verb. If we take the position upheld in the Moscow linguistics school that two instances are enough to posit regular polysemy (Апресян, Ю. Д., 1974:189; Падучева, 2004:28), the construction type depicting kinship is a matter of multiple argument constructions because the same type is available for the verb *считаться*. By taking an even broader perspective, the semantics of kinships can also be profiled with non-reflexive verbs, leading to increased overlaps in patterning, for example, *Он мне отец* 'he is my father.' A strict binary perspective may posit an accidental formation while constructionist perspective posits, albeit marginal it may be, regularity. In addition to this support by patterns, the divergent periphery appears to be supported by relative frequent verbs.

Turning to a constructionist perspective, Divjak and Janda argue that the Nominative-Accusative Construction occupies a central position in the network of the Russian construction types and the various impersonal types have a peripheral one. They consider that one contributing factor to this status is the overall number of verbs capable of displaying this behavior. The central position is attributed to the transitive construction displaying the following pattern: nominative case, finite verb, and accusative case (Divjak & Janda, 2008). In order to establish different impersonal construction types, specifically the interaction between the dative oblique and the infinitive, Divjak and Janda consider that the dative depicts the semantic role Experiencer and establish links between argument constructions based on the deviation from agentivity to Experiencer.

Divjak and Janda analyze the relationship between a dative oblique and an infinitive pattern and its association to non-canonical subjecthood, of which these elements occupy the subject position. According to them, there are 81 verbs displaying this behavior in Russian. Generally stated, the function of these elements is a debated topic in Russian linguistic tradition. Divjak (2004:19-33) proposes a pronominalization test based on small design elicitation test with 15 native speakers. The test was designed as a rating task on a three-point scale. The test was used to probe the possible difference between the dative and the infinitive. Examples 3.2-12–3.2-14 illustrate the rationale behind the proposed test, that both the noun phrase and the infinitival clause are acceptable answers. (Divjak & Janda, 2008:166).

3.2-12 *Что* *он* *планиру-ет?*
What.ACC he-NOM plan-3S.PRS
What he plans to do?
3.2-13 *Поездк-у* *в* *Москв-у.*
Trip-ACC PR Moscow-ACC
A trip to Moscow.
3.2-14 *Поеха-ть* *в* *Москв-у.*
Travel-INF PR Moscow-ACC
To travel to Moscow.

Pronominalization as a test is part of Keenan's (1976:315) concept of clustered properties of subjecthood. This test is also briefly discussed by Leinonen (1985) in her study of impersonal sentences in Russian. Moreover, this is a standard test in formal approaches to establish constituents. For example, a clausal complement is taken to possess the same status as verbal arguments. In contrast, this test is typically criticized for yielding contradictory results in cognitive and functional approaches. Verhagen (2008:136) explicitly states that reducing complex clauses (complement clauses) into simple clauses in the form of X verb Y does not state anything about the mismatches between these two types. Although Verhagen is only considering the *that*-clause types, the parallelism to the Dative Verb Infinitive pattern is there. The mismatches are apparent in the data provided by Divjak and Janda, cf. Table 3.2-7.

Reflexive Verb	translation	Obligatory dative	Optional dative	Infinite subject	used in a morp. def. Sense	exists in a morph. def. sense
довестись	happen, manage	###				###
доводиться	happen, manage	###				###
захотеться	start to want	###			###	
нравиться	please	###		###		
оставаться	remain		###		###	
остаться	remain		###		###	
полагаться	suppose		###		###	
понадобиться	necessary	###			###	
предписываться	order, prescribe	###		###		
предполагаться	intend	###		###		
прийтись	happen		###		###	
приходиться	happen		###		###	
случаться	happen	(###)			###	
удаваться	succeed	###			###	
удаться	succeed	###			###	
хотеться	want	###			###	

Table 3.2-7 Aggregated list of patterns available for Russian impersonal reflexive verbs based on Divjak and Janda (2008). The hashes are used to mark the presence of the property.73

Table 3.2-7 gives an aggregated binary list based on Divjak and Janda (2008). The list only includes those verbs attested in the database. The status of the oblique dative is divided into two parts: oblique and optional. The oblique label states that the dative is an obligatory element of the construction. Although it is given as a binary category, Divjak and Janda (2008:141 footnote 3) posit that it is a simplification and can be interpreted in probabilistic terms rather than a categorical binary property. The focus in this study is the relationship within this pattern. Thus, a verb-specific property of obligatoriness is beyond the scope of this study (cf. Section 3.2.7 for further discussion). Nonetheless, the data according to Divjak and Janda suggest a high level of specificity. For example, the dative is labeled as optional with the verb *приходиться* and the infinitive does not occupy the subject position. In this vein, it seems plausible that there is a verb-specific continuation from obligatorily profiled participants to optionally profiled ones and the verb-specific constructions carve this continuation based on their distinctive usage patterns.

The Infinitive subject -column gives the status of the infinitive, whether it can occupy the subject position or not. Furthermore, the status of the whole pattern is divided into two sets. These sets are: 1) used in a morphologically defective sense, and 2) exists in a morphologically defective sense. The former label is used to capture multiple argument constructions. The verbs which can have a nominative slot, belong to the latter type while verbs which cannot have a nominative slot pertain to the former (Divjak & Janda, 2008:168-169). As an example of this behavior, the verb *нравиться* 'please, like' can appear in both the Dative-Nominative, and Dative-Infinitive patterns. Following the classification

[73] The translations were taken from Divjak and Janda.

proposed by Divjak and Janda, the syntactic role slot can be filled with two types of non-canonical subjects: the Dative subject or the Infinitive subject. At the same time, mismatches emerge in specific configurations, as illustrated in 3.2-15. The complement, *пассивной* 'passive,' of the copula *быть* 'be' agrees in gender and number with the dative argument.

3.2-15 *[е]й нрав-ит-ся бы-ть пассивн-ой.*
She.DAT like-3S.PRS-RM be-INF passive-INS
She likes to be passive.
[1295, RNC, Программа "Культурная революция" на телеканале "Культура" (2006.04)]

When weak subject candidates are merged together into a complex construction, mismatches are to be expected. From a constructionist perspective, Example 3.2-15 consists of the *нравиться* -type and the so called Dative-Infinitive Construction. The latter type is generally considered to be the strongest Dative subject candidate in Russian, as it can display reduced indexing properties similar to the structure in 3.2-15 (Kyröläinen, submitted; Moore & Perlmutter, 2000; Zimmerling, 2009).[74] In order to preserve the classification proposed in Divjak and Janda (2008), the verb *быть* 'be' is analyzed as occupying the subject slot with the *нравиться* -type and the indexing properties are a residue of the operation of conjoining with the Dative Infinitive Construction.

3.2.6 Non-Canonical Subject Roles and Other Deviations

The final question is the status of the *that*-clause within the network of the Russian Reflexive Marker. Crucially, this question is related to the level of granularity imposed on the category of the Reflexive Marker. Gerritsen's study is the most comprehensive in this regard and she offers highly nuanced analysis between the reflexive and the non-reflexive verb. An outline of her taxonomy is given in Table 3.2-8 with definitions and examples. First, Gerritsen divides the reflexive verbs into two broad groups called impersonal improper and impersonal proper. The two are then further divided into several subgroups. The difference between the subtypes of the impersonal improper is the presence of a possible substitute for the subject (Gerritsen, 1990:125).

[74] Sigurðsson (2002:704, 713) denies the subject status of the dative argument in the Dative Infinitive Construction and posits a silent, (i.e., zero) copula as part of the structure to handle the indexing properties.

	Impersonal improper
1)	IR in which a constituent is present and may be said to occupy the position of a subject, being its substitute, for instance, an infinitive, a subordinate clause, a prep.O. Such IR correspond to NR in which the same constitutients occupy the place of acc.O.
2)	IR in which no SubS is present, but in which there are indications for the presence of an implied subject
	Impersonal proper
1)	IR which do not have a corresponding PR
2)	IR which do have a corresponding PR, but with which the PR has an interpretation which differs more or less from that of the IR

Table 3.2-8 Definitions of the impersonal types according to Gerritsen.
IR = impersonal reflexive verb, PR=personal reflexive verb, NR = non-reflexive verb, prep. O. = prepositional object, and acc. O = accusative object.

Examples 1-4 are taken from Gerritsen to illustrate the classification.[75]

1) Impersonal improper: NR) *Он говорил, что* 'he said that' → PR) *Это говорилось* 'it was said' → IR) *говорилось, что* 'it was said that'
2) Impersonal improper: *рассказывалось о Пушкине* 'About Pushkin was spoken' → implied subject: *рассказ* 'narrative' etc.
   Impersonal proper 1) *мне хорошо работается* 'the work is going well for me'
3) Impersonal proper: PR) *Он делался веселым* 'he was happy,' IR) *ему делалось весело* 'he became happy'

The basic tenet of this classification is the existence of a correspondence between a reflexive and non-reflexive verb. However, Gerritsen's very first definition of IR improper states: "IR improper usually have a corresponding PR" (Gerritsen, 1990:125). This definition apparently covers verbs which may not have a corresponding reflexive verb. Additionally, Gerritsen's subtypes, proper and improper, are, in a sense, akin to the labels of the two defective senses proposed by Divjak and Janda. Nonetheless, Gerritsen's proposed set of types makes an interconnection between different reflexive constructions, although this possibility is not explored in her study. Instead of trying to cycle reflexive verbs through constructions available for the non-reflexive verbs, the impersonal *that*-clause type is taken to constitute a construction of its own right. Similar argumentation is made by Verhagen for a general *that*-clause construction. He argues that a bottom-up approach, in which patterns are not reduced into more abstract representation of constituency, can be used to handle the apparent mismatches between different construction types. Additionally, Verhagen supports this view in terms of type and token frequency attested with the *that*-clause construction type in Dutch (Verhagen, 2005:82-83, 102-103).

[75] Translations are not provided in Gerritsen's study. They were added by the author.

The argumentation in this study builds on the same principle that the *that*-clause construction is a subtype in the network of the Russian Reflexive Marker. It can be considered as a product-oriented schema in Bybee's terminology as was outlined in Section 2.2.2. It follows from this that the construction type may function as a gravitational center. Table 3.2-9 gives the attested reflexive verbs appearing in an impersonal construction type, which potentially can be conjoined with the *that*-clause construction. It is apparent in the table that there are multiple pathways to the *that*-clause construction type, even if we are only considering the variation with the reflexive verbs. The semantic content of these instances is the shared commonality between them. They are used to profile a content, which can be either a content of a speech or a content of a perception.

	Pattern			
Reflexive Verb	Translation	nominative	dative	that
вериться	believe		###	###
выразиться	express	###	###	###
выясниться	emerge		###	###
говориться	speak		###	###
иметься (в виду)	have, mean			###
казаться	seem, appear	###	###	###
мечтаться	dream		###	###
нравиться	please	###	###	###
обнаружиться	come out	###		###
оказаться	turn out, appear	###	###	###
оказываться	turn out, appear	###	###	###
подразумеваться	imply	###		###
показаться	turn up, appear	###	###	###
полагаться	suppose		###	###
предлагаться	suggest	###	###	###
предписываться	order, prescribe		###	###
предполагаться	intent	###	###	###
представляться	appear, arise	###	###	###
разуметься	mean	###	###	###
случаться	happen	###	###	###
спрашиваться	ask	###		###
считаться	consider, kinship	###	###	###
указываться	point out	###	###	###
чувствоваться	feel	###		###
чудиться	seem	###	###	###

Table 3.2-9 Attested impersonal reflexive verbs with the potentiality to combine with the that-clause construction type in the database. The hashes mark the presence of the pattern.

By breaking the table into patterns, a structure emerges displaying the pathways to the *that*-clause type. Assuming that a high frequency verb can function as a pathway for extensions, the verb *оказаться* 'seem, appear' (482,7 freq) is a strong candidate. The personal pattern consisting of the Nominative Verb Instrumental -pattern is given in 3.2-16. This construction type and similar instances are discussed in Chapter 8. Additionally, this is a reflexiva tantum verb in contemporary Russian although it has the neighbor verb *оказать* 'render.'

3.2-16 [*ч*]то *воронк-а* *от взрыв-а* *оказа-л-а-сь*
that crater-NOM PR explosion-GEN appear-PST-F-RM
настолько *больш-ой* […].
so great-INS
That the crater from the explosion appeared to be so great.
[552, RNC, Сейсмологи утверждают, что в КНДР произошло 2 взрыва // "РБК," 2004.09.12]

Instead of positing derivations (cf. Moore & Perlmutter, 2000:400-401), the patterns in Examples 3.2-16 and 3.2-17 constitute separate argument construction types supported by high frequency verb.

3.2-17 [*и*] *вдруг* *оказа-л-о-сь* *что*
and suddenly seem-PST-N-RM that
это огромный кусок жизни моей.
[…]
And suddenly it turned out that this is a huge proportion of my life.
[1301, RNC, Радиоинтервью Юрием Маликовым (2006.04)]

Another pattern functioning as the pathway is exemplified with the high frequency verb *нравиться* 'please, like' given in Examples 3.2-18–3.2-20.[76] The patterns with this verb offer a wide range of potential combinatory possibilities: Nominative Reflexive Verb Dative, Infinitive Reflexive Verb Dative, and That Reflexive Verb.

3.2-18 [*м*]*не* *она* *тоже* *нрав-ит-ся.*
I.DAT she.NOM also like-3S.PRS-RM
I also like her.
[1054, RNC, Разговор при выходе из дома, Москва (2005.04)]

3.2-19 [*е*]*й* *нрав-ит-ся* *бы-ть* *пассивн-ой.*
she.DAT please-3S.PRS-RM be-INF passive-INS
She likes to be passive.
[1295, RNC, Программа "Культурная революция" на телеканале "Культура" (2006.04)]

[76] The last example is specifically extracted from RNC because the *that*-clause construction type is not attested in the sample.

3.2-20 *Нрави-л-о-сь,* *что* *вс-е* *сме-ют-ся.*
Please-PST-N-RM that everyone-NOM.PL laugh-3P.PRS-RM
It is pleasing that everyone laughs.
[RNC, Андрей Геласимов. Фокс Малдер похож на свинью (2001)]

Yet another pattern is evident with the verb *вериться* 'believe' and *мечтаться* 'dream' consisting of dative and *that*-clause, exemplified in 3.2-21. Although this pattern is more restricted, these verbs always appear in the impersonal construction, but their semantic range overlaps with the previously mentioned verbs due to the inclusion of the dative.

3.2-21 *Мне* *прост-о* *не* *вери-л-о-сь,* *что* […].
I.DAT simple-ADV NEG believe-PST-N-RM that
I simply did not believe that.
[616, RNC, Джим Кэрри - изнутри и снаружи // "Экран и сцена," 2004.05.06]

Finally, there is a small group of verbs which appear both in the personal and impersonal type, but display a divergent, verb-specific affinity towards the inclusion/exclusion of the dative argument, which include such verbs as *подразумеваться* 'imply' and *предлагаться* 'suggest' in 3.2-22 and 3.2-23.

3.2-22 [*п*]*одразумева-ет-ся,* *что* *следстви-е* *появи-л-о-сь*
imply-3S.PRS-RM that consequence-NOM appear.PST-N-RM
после *причин-ы.*
PR cause-GEN
It implies, that the consequence appeared after the cause.
[487, RNC В.Н. Комаров. Тайны пространства и времени (19952000)]

3.2-23 *В* *стать-е* *предлага-ет-ся* *подход* *к*
PR article-PREP suggest-3S.PRS-RM approach.NOM PR
решени-ю *задач* синтез-а топологи-и […].
solution-DAT problem.GEN.PL synthesis-GEN topology-GEN
In this paper, an approach to solving problems of synthesis of the topology is proposed.
[17, RNC, Задачи синтеза сетей синхронной иерархии // "Информационные технологии," 2003]

It seems reasonable to assume that the *that*-clause construction is a schematic type with lower level subtypes supporting it, for example, highly verb-dependent configurations in terms of the inclusion/exclusion of the dative argument. Describing the impersonal construction types in this manner offers a motivational pathway of the divergent nature of these patterns. Figure 3.2-2 illustrates one possible network for the *that*-clause construction type in Russian. The *that*-clause construction type functions as a gravitational center, a product-oriented schema. Moreover, Figure 3.2-2 displays the configuration only in terms of patterns supported by the Reflexive Constructions. From a lexical perspective, the proposed preliminary network could easily be extended to cover non-reflexive construction types, as well.

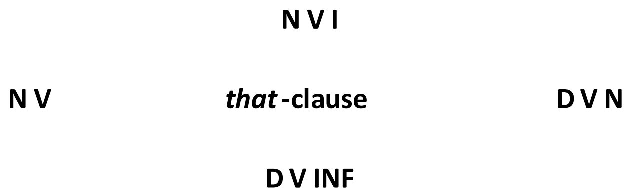


Figure 3.2-2 Network for the *that*-clause construction in Russian based on the sample. N = nominative, V = verb, D = dative, I = instrumental, and INF = infinitive.

In sum, the model on the canonical and non-canonical subject roles is divided into four roles. The distribution is given in Table 3.2-10.

	Subject Role				
Agreement	CL.Sbj	D.Sbj	INF.Sbj	Sbj	Sum
Impersonal	56	59	9	0	124
Personal	0	0	0	1865	1865
Sum	56	59	9	1865	1989

Table 3.2-10 Distribution of the subject roles in the database: CL.sbj (*that*-clause), D.sbj (Dative subject), INF.sbj (Infinitive subject) and Sbj (Nominative subject). The type Other (*n* = 11) is excluded.

The canonical subject role is supported by nominative case and agreement in person and number. The non-canonical subject roles are established through two properties: non-agreement and pattern. The latter is further divided into three types: dative, infinitive, and *that*-clause. On one hand, the status of these roles is dependent on the construction they appear in. On the other hand, they arise from verb-specific constructions.

3.2.7 Non-Subject Roles

Thus far, the concept of argument has been applied to the primary slot and its layered structure, (i.e., the subject role). The secondary slot covers a problematic area in any linguistic theory. In terms of syntactic roles, the status of the secondary layer is crucial in establishing usage patterns. Thus, the questions pertain to the status of the entities that can be used to profile the secondary syntactic role. Additionally, the question is framed against obligatoriness or optionality. The most schematic claim would state that only the subject role is an obligatory element of the Intransitive Construction following the position, as described by Dowty (1991). Very little semantic information, however, would be included in a model which would operate on a single role. At the same time, this has to be understood in relation to the goals set for the model. If only the mapping from verbs to roles is the ultimate goal, fewer distinctions need to be stated. On the other hand, if usage patterns are the designated target, as is the case in this study, a richer representation has to be considered. Examples 3.2-24–3.2-26 serve to illustrate the situation.

3.2-24 *Также в студи-и наход-ят-ся музыкант*
Also PR studio-PREP locate-3P.PRS-RM musician.NOM
Серге-й Галанин и наш-и гост-и.
NAME-NOM NAME.NOM and our-NOM.PL guest.NOM.PL
A Musician Sergey Galanin and our guests are also in the studio.
[1048, RNC, Беседа с рок-музыкантами о проблемах наркотиков, НТВ "Кома" (2002.06.03)]

3.2-25 *Вс-е бы-л-и безумн-о рад-ы и*
All-NOM-PL be-PST-PL insane-ADV happy-NOM.PL and
готови-л-и-сь к больш-ому весель-ю.
prepare-PST-PL-RM PR grand-DAT festivity-DAT
Everybody were insanely happy and prepared to the grand festivity.
[1727, RNC, Сергей Седов. Доброе сердце Робина // "Мурзилка," №7," 2002]

3.2-26 *Мы не мож-ем,*
We.NOM NEG can-1P.PRS
по крайней мере, на существующем уровне науки и техники,
[…]
возврати-ть-ся в прошл-ое […].
return-INF-RM PR past-ACC
At the current level of science and technology, we cannot return to the past.
[483, RNC, В.Н. Комаров. Тайны пространства и времени (1995 2000)]

Typically, linguistic expressions are framed against certain location or time as in 3.2-24 and 3.2-26. Traditionally, these expressions are considered to be optional in terms of the structure of the verb, in other words, adjuncts (cf. Якобсон, 1985a:66-67). However, the location is an obligatory argument with the verb *находиться* 'be located' (cf. Золотова, Г. А., 2005 [1973]:131). A similar situation is present in Example 3.2-25 with the verb *готовиться* 'prepare.' The secondary slot profiled with *к*$_{\mathrm{dat}}$ is typically optional, but in a specific context it may become closer to obligatory, as is the case in Example 3.2-25. Thus, the relation of obligatoriness and optionality appears to be gradient rather than binary (cf. Croft, 2001:179-180; Langacker, 1988a).

In terms of the derivational approach, the mapping between the base and the derived form does not entail that the same semantic structure is available with both verbs. Geniušienė proposes that both primary and secondary functions have to be separate. The primary semantic function is used to describe the base form and the secondary pertains to the derived structure (Geniušienė, 1987:50-51). This distinction hinges on the assumption that the base form is identified. Additionally, the account must find pairs or the distinction between the roles is lost. In contrast, syntactic roles are defined relative to argument constructions in most versions of Construction Grammar.

The terms actant and circumstantial are commonly employed in the diathesis tradition and in the Meaning–Text theory analogous to Tesnière's (1959) concepts (cf. Mel'čuk, 2004 for an overview). The actant corresponds to the traditional obligatory argument and the circumstantial to adjunct. Although the theoretical underpinnings between constructionist and diathesis approaches are different, certain similarities are, nonetheless, present. For example, Mel'čuk (2004:13) proposes the Obligatory Participant Inheritance Principle. When the logical structure is stripped away, the principle might be paraphrased in the following manner: a certain situation type inherits all the obligatory participants of a certain lexical item.[77] Paraphrased in this manner, this principle becomes in close proximity with Goldberg's Construction Grammar. The Correspondence Principle links the participant roles of the verbs to the argument roles of the argument constructions, cf. Section 2.2.3 (Goldberg, 1995).

Paducheva (2004:72-73) notes that the criteria to establish distinctions between them is lacking (cf. Паневова, 1978). One motivation for the situation is that the status of the circumstantial or adjunct is considered to be located at the periphery of grammar. Perhaps the most prominent feature is syntactic obligatoriness. The omission of the obligatory argument leads to ungrammaticality. The potential omissability comes under various shades, cf. Section 3.2.2, and crucially depends on what kind of data is considered, as evidence ranging from hand-crafted minimal pairs to spoken discourse will certainly lead to different outcomes.

In contrast, the distinction between the traditional concept of obligatory and optional argument is not maintained in FrameNet. Instead, elements of a frame are established either as core or periphery. Crucially, the distinction between core and periphery is defined relative to a frame and not to a lexical item. From a syntactic point of view, typical oblique elements may be considered as core depending on the evoked frame semantics (Fillmore, 2007:133). In this vein, even traditional locative elements may be considered obligatory if they are used to profile the Locative Relation Frame. This position allows us to maintain the importance of the patterns with verb-specific constructions. Changes in the pattern are naturally incorporated as part of the description. For example, the preposition *без* 'without' is another typical candidate for adjunct. However, certain verbs, such as *остаться* 'stay, remain' form an idiosyncratic pattern with it, given in 3.2-27 (Кузнецов, 2009 [1998]).

[77] This is a simplification, as Meaning–Text theory operates with semantic, deep-syntactic, and surface-syntactic actants.

3.2-27 *Кстати, и Эдвард Дженнер не*
By.the.way also NAME.NOM NAME.NOM NEG
оста-л-ся без наград-ы.
remain-PST.M-RM PR reward-GEN
By the way, Edward Jenner did not go unrewarded either.
[26, RNC; Как родилась иммунология // "Знание — сила", №7", 2003]

For purposes of the present study, obligatoriness is not considered because building a pattern dictionary for the unique reflexive verbs is not the ultimate goal.[78] The profile of the secondary slot was primarily determined based on Kuznecov (Кузнецов, 2009 [1998]), and Denisov and Morkovkin (Денисов & Морковкин, 2002). Both dictionaries contain a fairly detailed description of basic patterns in Russian. Thus, the profile of the secondary slot is not defined in terms of obligatoriness.

As the exact profile of the secondary slot depends on the profiled argument construction type, the labels associated with them are defined relative to them. Three labels are used in this study. First, the label Complement is assigned to the secondary slot for types that are traditional analyzed as (semi-)copulas, primarily used with the instrumental or nominative case (cf. Krasovitsky, Long, Baerman, Brown & Corbett, 2008), cf. Chapter 8. Second, the same label is used with *that*-clauses and infinitives (cf. Падучева, 2004:249, 268, 314). Third, all other types are labeled as Oblique. Finally, a dummy “None” is used in cases where the secondary slot is not profiled. The distribution is given in Table 3.2-11.

	Non-Subject Role			
	Complement	None	Oblique	Sum
Instance	322	531	1136	1989

Table 3.2-11 Distribution of the non-subject roles in the database. The type Other (n = 11) is excluded.

In sum, this section introduced the encoding used for argument constructions, generalizations over verb-specific constructions, but the inherent variability of reflexive verbs is incorporated in the model through the variable Pattern. Regardless of the theoretical inclinations of modeling syntactic and semantic configurations, any model has to arrive at the profile. The profile is the minimal structure used to convey the meaning of the sentence. Examples 3.2-28–3.2-30 illustrate the variability with the verbs *оставаться*$_{\text{imp}}$ ‘stay, remain’ and *остаться*$_{\text{perf}}$ ‘stay, remain.’

[78] Certainly corpus-based methods, such as constructing statistical profiles for lexical items, is one possibility as is proposed by Azarova et al. (2004) for Russian verbs or structured native speaker interviews for determining the degree of obligatoriness as in Faulhaber (2011) for English.

3.2-28 [*у*] *стенок* *цилиндр-а* *оста-ёт-ся*
PR wall.GEN.PL cylinder-GEN remain-3S.PRS-RM
почти чист-ый *воздух.*
almostpure-NOM air.NOM
Almost pure air remains at the walls of the cylinder.
[342, RNC, Ликбез: Что такое непосредственный впрыск бензина // "Автопилот," 2002.02.15]

3.2-29 [*к*]*лиент-ами* *Moet* *остава-л-а-сь* *состоятельн-ая*
Client-INS.PL NAME.GEN remain-PST-F-RM wealthy-NOM
и *не* *очень* *юн-ая* *публик-а.*
and NEG very young-NOM audience-NOM
Wealthy and not so young audience stayed as clients of Moet.
[893, RNC, Владимир Ляпоров. Молодая гвардия. Искусство быстрого завоевания новых рынков сбыта // "Бизнес-журнал," 2003.10.23]

3.2-30 *Поэтому* *автор-у* *настоящ-ей* *стать-и* *оста-ёт-ся*
Therefore, author-DAT this-GEN article-GEN remain-3S.PRS-RM
упова-ть *на то*
hope-INF PR that.ACC
что он не является “типичным западным учёным” […].
[…]
Therefore, it remains for the author of this article to hope that he is not a typical Western scholar.
[377, RNC, Владимир Успенский. Виттенштейн и основания математики (2002)]

Any adequate theory of the (Russian) Reflexive Marker must be able to model the patterns in 3.2-27.–3.2-30. The inclusion of the variable Profile incorporates the verb-specific connections across generalizations and, at the same time, maintains the verb-specific construction

3.3 Summary: Layered Structure

The previous sections have outlined the basic variables used in the encoding of the verb-specific constructions in terms of the Layered Model.[79] The model offers an augmented version of the basic slot types presented in Construction Grammar. The Layered Model makes it possible to globally and systematically compare different relations and patterns in the data across different abstractions. A key component in the proposed model is the ability to take different levels of granularity into consideration, assuming that a variation in the level of abstraction can bear consequences to the overall descriptive aspect of the model. Figure 3.3-1 demonstrates the analysis of Example 3.3-1 as layered

[79] The term layer already appears in Cognitive Grammar. Langacker (2002) has used concept of layer to analyze morphological structures and later clausal patterns (Langacker, 2009).

structure bringing together the variables discussed in the previous sections. The same example was used in Section 3.2.1 to illustrate the dependency-based model.

3.3-1 *Волчиц-а* *оказа-л-а-сь* *бешен-ой.*
She.wolf-NOM appear-PST-F-RM rapid-INS
The she-wolf appeared rabid.
[517, RNC, Темниковские охотники отстреливают лис и волков //"Московский комсомолец" в Саранске," 2004.12.23]

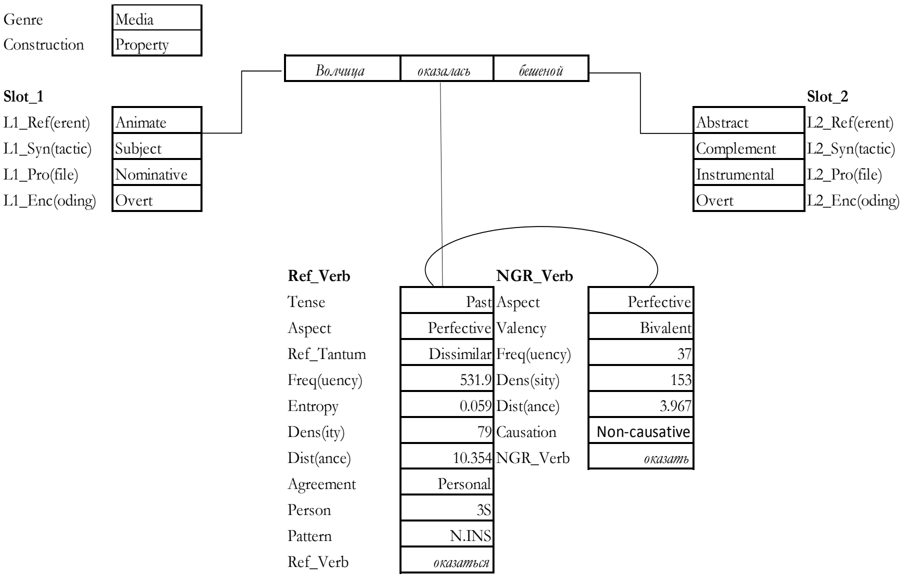


Figure 3.3-1 Layered structure of the Argument Construction.

Formalizing constructions in the above manner resembles unification-based Construction Grammar. This formalization is defined as Attribute-Value Matrix (AVM). The commonality is that an attribute may consist of multiple values, but only one can be present at a given time. Additionally, the variables are not considered to be universal (cf. Fried & Östman, 2004:29-30; Kay & Fillmore, 1999 for elaborated discussion on AVM). Although Goldberg (2006:216-237) criticizes using values, as they tend to overemphasize recurrent elements and subtle semantic differences are not easily captured with them. However, any model requires operationalization, defining the subject matter at hand in quantifiable terms (cf. Bod, 2009). Due to adapted formalization, argument constructions are not defined as the pairing of form and meaning in the general sense. Instead, they constitute the combination of the layers.

At the same time, the notation given in Figure 3.3-1 is just one possible mode of formalizing the proposed set of the variables, but it confines to the non-derivational and monostratal model of constructionist approaches. In this vein, the analysis does not hinge on postulating basic forms and then somehow deriving the surface structure from these. Importantly, all information is located on the same level without separate modules. A citation from Kay (2002a:18) conveniently captures this: "To be sure, an architecture imposes constraints on the theories that can be expressed within it, but we should not expect a theory of grammar to follow from the notation in which it is expressed any more than we expect a theory of planetary motions to follow from the notation of differential equations."

Returning to Construction Grammar, there are actually very few studies on certain fundamental aspects related to argument structures, namely how argument constructions identified and separated. Croft (2001:51-53) simply states that the separation of construction types is a matter of categorization. Goldberg (1995:9) provided evidence for positing a construction in terms of coercion, (i.e., the famous *he sneezed the napkin off the table*). The intransitive usage of the verb *sneeze* is conceptualized in terms of the Caused Motion Construction providing evidence for the abstraction. However, when compositional structures are also viewed as constructions, coercion cannot be used as a diagnostic tool (cf. Michaelis, 2004). Some attempts have been made in computational linguistics, for example, by Bod (2006; 2009), and Lagus et al. (2009).

Lagus et al. investigate the possibility to derive constructions from plain text. As Lagus et al. note that the constructions formed by the statistical model, such as the Finnish *olipa kerran* [*x*] 'once upon a time [x],' are far from the typical types of constructions proposed in the literature. An edited volume by Sahlgren and Knutsson (2010) also presents several papers on this topic. Thus, a great deal of work is still required if a more data-driven approach to forming constructions is to be achieved. Perhaps this might be outside the scope of (theoretical) linguistics and only by a joined effort with computational linguistics and psycholinguistics might there be a way to tackle this issue, (i.e. an unsupervised model to form argument constructions from data).

Another issue related to formulating construction types is the reliability of the encoding schema. As Artstein and Poesio (2008) point out, disambiguating tasks of verb senses tend to yield fairly low inter-rater agreement. Generally, Kirppendorff has proposed the most stringent criteria for data that may be analyzed for reliability. The criteria consist of: 1) exhaustive formulation of the encoding schema fixed in advance, 2) stating the number of coders used the minimum being three, and 3) having the coders work independently from each other and having a divergent background. The last criterion is intended to minimize the possible influence of theoretical background. For example, experts may simply agree solely based on their shared knowledge of the subject matter and not based on the encoding schema. Having access to, at least, three independent raters after the encoding schema was finalized would require a research group. Thus, inter-rater agreement on the encoding schema was not carried out. However, the encoding of the verbs is made publicly available in machine readable format, ensuring that, a comparison to future studies can be made similar to studies as Barðdal (2008), and Schulte im Walde (2006) that contain a fairly extensive number of verbs.[80]

At the same time, it is possible to distill information from the existing body of literature to form two guiding principles. First, Goldberg's (2006 Chapter 6) model on multiple patterns available for a particular verb can be applied in situations where coercion is not applicable. If a particular verb appears in multiple argument constructions and these types are able to attract other verbs or are supported by other verbs, generalizations over them may be considered warranted. A similar argumentation is evoked by Paducheva (Падучева, 2004:149). Second, argument constructions tend to have a few frequent members that align with the contribution of the verb-specific construction with the argument construction. For example the Ditransitive Construction aligns with the verb *give* in English. Goldberg (2006:88-89, 92) has proposed that high frequency items may function as anchors where a form of standard of comparison that may facilitate both learning and assimilation of members into a category. Thus, these re-occurring frequent verbs may be considered to be the central members of a particular abstraction (cf. Bybee, 2010:25-31).

Chapters 5–9 are used to formulate the proposed argument constructions type. Specifically, the following properties are discussed:

1) Properties of the argument construction type.
2) The function and form of the proposed argument construction.
3) Possible central members.
4) The nexus between the verb-specific constructions and argument constructions.
5) The relation of the proposed argument construction to previous established taxonomies.
6) The Classification accuracy of the model.

Detailed discussion about the formation of the argument constructions is offered for two reasons. First, the lack of a constructionist model established through surface-structure generalizations has not been proposed for the (Russian) Reflexive Marker, at least to my knowledge. Second, because inter-rater agreement was not performed on the data, the solutions to formulate the generalizations need to be made explicit. Thus, Chapters 5–9 follow the same format which include a discussion on the formation of the argument construction and an evaluation of the performance of the model based on the input. The proposed set of argument constructions are grouped under four semantic generalizations following Goldberg (1995:39-43), Barðdal (2008:46-47, 66), and Langacker (1987:148) in Chapters 5–8. Chapter 9 contains the argument constructions, which have a low type frequency in the sample. These types are labeled as minor argument constructions. Before turning to the

[80] If we follow Krippendorff's criteria, the encoding of the variable Causation extracted from the Russian National Corpus is also unreliable as the number of encoders, agreement rate, and the exact encoding schema are not publicly available.

argument constructions, the properties of the chosen machine learning algorithm in relation to the conceptual basis of usage-based models and the fitting process of the data are discussed in Chapter 4.

4 Multivariate Methods: Random Forests

This chapter introduces random forests machine learning algorithm. The method is an extension of the classification and regression trees introduced in Section 1.4.2. The potential conceptual connection between random forests and usage-based grammars is discussed in Section 4.2. This connection is critical for linguistic models that assume cognitive plausibility. If some aspects of the linguistic model are assumed to be connected to the mental lexicon, the same principle applies to methods. At least partly, the chosen method should align with what is known about processing. Finally, random forests contain parameters that influence their ability to learn similar to most machine learning algorithms. Although, the tuning of these parameters is specific to a particular data set, it is demonstrated that random forests are typically fairly insensitive to changes in the parameters in Section 4.4. This section also covers the fitting process of the data and validation measures of random forests model are outlined in Section 4.5. Before taking on the fitting process of the RF model, the data contains infrequent construction types. The selection of the data for the modeling is discussed in the following section.

4.1 Data Selection and Summary of the Variables

The data contain infrequent argument construction types that are problematic for the purposes of modeling. All things being equal, there might not be enough data points to learn these types from the input. Table 4.1-1 gives the distribution of the argument construction types in the database.

	Argument Construction				
	Content	Existence	Experiencer	Experiencer Extension	**Experiencer Perspective**
Instance	113	85	137	58	1
	Inclusion	Location	Motion	**Other**	Passive
	38	44	212	11	208
	Permissive	Phase	Property	Reciprocal	Reflexive Engagement
Instance	9	145	171	103	215
	Relation	Spontaneous Event	**Semantic Reflexive**	Stimulus	**Stimulus Extension**
Instance	26	330	25	65	4

Table 4.1-1 Frequency of the argument construction types in the database. Types in bold are exluced from the RF model.

The minimum threshold value was set to the frequency of 44, the Location Construction. The excluded construction types are given in bold. The minimum threshold value is not completely arbitrary. Random forests model is trained on random samples of the data. Highly imbalanced data leads to a situation where the smallest type might not be included in the random samples (cf. Westerhout, 2009). Nonetheless, the difference in frequency between the Location and the Inclusion is small. The removal of the infrequent types amounts to a loss of 114 data points. Another potentially problematic variable is the Pattern, even after removing the infrequent argument construction types. Table 4.1-2 gives the frequency of the Pattern after the infrequent types had been removed.

Pattern	Instance
CL	17
D.CL	27
D.G	**4**
D.INF	60
I.INF	**1**
INF	**2**
N	478
N.ADV	41
N.CL	15
N.D	80
N.G	38
N.I	416
N.INF	82
N.N	10
N.P	201
N.PA	136
N.PD	49
N.PG	126
N.PI	92
P.CL	10
PG.CL	**1**
Sum	1886

Table 4.1-2 Frequency of the variable Pattern after the minor types had been removed. Infrequent patterns are given in bold.

Certain patterns are still infrequent in the dataset. These values are given in bold (*n* = 8) and removed from the dataset, leaving a total of 1,878 data points for the purposes of modeling. In theory, the remaining frequency counts might be sufficiently high to apply some univariate method. However, once they are conditioned by the argument construction types, they become sparse. Certain types only appear in a specific configuration leaving empty cell values. This issue would remain even in a large scale study simply because it is a property of verbs (cf. Surdeanu, Harabagiu, Williams & Aarseth, 2003). Generally, data sparseness is an issue in language models when random samples are used instead of selected items (Xu, Peng & Jelinek, 2007). Thus, the final dataset contains 1878 data points, a loss of 6.1% of the data compared to the original sample size of 2000. Another aspect regarding data loss is the number of the unique reflexive verbs. The final dataset contains 770 unique reflexive verbs compared to the full sample of 819 unique reflexive verbs. Thus, a large number of reflexive verbs are still maintained in the final dataset. Table 4.1-3 gives the summary of the variables.

Variable	Level	Range	Description
Genre	4	NA	Genre based on the Russian National Corpus
L1_Ref	4	NA	Referent type of the subject slot
L1_Syn	4	NA	Syntactic role of the subject slot
L1_Pro	4	NA	Profile of the subject slot
L1_Enc	2	NA	Encoding of the subject slot
L2_Ref	5	NA	Referent type of the non-subject slot
L2_Syn	3	NA	Syntactic role of the non-subject slot
L2_Pro	13	NA	Profile of the non-subject slot
L2_Enc	3	NA	Encoding of the non-subject slot
Tense	5	NA	Tense of the reflexive verb
Ref_Aspect	2	NA	Aspect of the reflexive verb
Construction	13	NA	Argument construction type
Pattern	17	NA	Pattern of the reflexive verb
Agreement	2	NA	Indexing type of the reflexive verb
Person	4	NA	Person marking of the reflexive verb

Entropy	NA	min. = 0.001,
		mean = 0.044,
		max. = 0.195
			Contribution of the reflexive verb to the information content of the argument construction type

Ref_Density	NA	min. = 0,
		mean = 141,
		max. = 865
			Neighborhood density of the reflexive verb based on its rhyme neighborhood

Ref_Distance	NA	min. = 0,
		mean = 4267,
		max. = 8.737
			Average pair-wise distance of the reflexive verb to its rhyme neighbor verbs

Causation	2	NA	Causative type of the neighbor verb
NGR_Aspect	4	NA	Aspect of the neighbor verb

NGR_Density	NA	min. = 0,
		mean = 239.6,
		max. = 1445
			Neighborhood density of the neighbor verb based on its rhyme neighborhood

NGR_Distance	NA	min. = 0,
		mean = 3.449,
		max. = 8.778
			Average pair-wise distance of the neighbor verb to its rhyme neighbor verbs

Ref_Tantum	4	NA	Degree of semantic connectivity between the reflexive and the neighbor verb
NGR_Valency	4	NA	Valency pattern of the neighbor verb

Ref_Freq_L	NA	min. = 0,
		mean = 3.738,
		max. = 6.278
			Normalized log frequency of the reflexive verb (+1)

NGR_Freq_L	NA	min. = 0,
		mean = 2.626,
		max. = 7.782
			Normalized log frequency of the neighbor verb (+1)

Table 4.1-3 gives the summary of the variables available for modeling with a descriptive label. Additionally, the levels of the categorical variables and the range of the numeric variables are also provided.

Dummy variable "None" was used to avoid missing values in two instances. First, a dummy neighbor verb "None" was used to avoid a missing value for the reflexive verbs that do not form a cross-paradigmatic relation, as was outlined previously. Second, a dummy coding "None" was used for certain levels. For example, the aspect of the neighbor verbs (NGR_Aspect) has four levels: Imperfective, Perfective, Biaspectual, and None. This issue was discussed earlier in Chapter 3.

It is common practice to dichotomize categorical variables before modeling with machine learning algorithms and certain algorithms may even require it (Arppe, 2008:118-119; Nielsen & Pradhan, 2004).[81] Instead of representing the referent type of the subject (L1_Ref) with four levels, they would be dichotomized into four new binary variables: L1_Ref_Person, L1_Ref_Animate, L1_Ref_Inanimate, and L1_Ref_Abstract with values present or absent. In Cognitive Linguistics, Gries and Divjak (2009) have popularized this kind of binary encoding of variables as ID-tags in their behavioral profile analysis. It enables one to directly compute frequencies of co-occurrence. In this case, Dichotomization would create 86 categorical variables. The seven numeric variables would also have to be discretized, creating an even larger set of variables. Additionally, lexical verbs are not used as variables. Instead, the Entropy and the Frequency are used to modulate their potential influence to the argument constructions. The reflexive verbs (*n* = 770) could be included as dichotomized variables. Effectively, 480 new variables would be included, as that is the number of hapax legomena, verbs appearing only once, in the final data set (*N* = 1,878).

From this perspective, the classical regression model is not applicable to the data at hand. As a rule of thumb, the minimum frequency of the least frequent level of the response variable would have to be at least 10 or 20 times the number of predictors in the model, setting aside all potential other issues such as correlation between the predictors and empty cell values (Arppe, 2008:116). Thus, the minimum frequency of a certain construction type would have to be at least 930.[82] Considering that the lowest frequency, 44, is attested with the Location Construction, this study is situated in the realm of small *n* and large *p*. This should not be understood as promoting an analysis on small data sets. In contrast, a number of advanced techniques could be implemented to reduce the dimensionality of the data, (i.e., a reduction in the number of the predictors). Some of these will be discussed in conjunction to the variable importance measures with random forests.

Nonetheless, dichotomization is not carried out in this study. The variable importance measures will reflect the importance of a given layer as a whole and not its individual levels. For example, the importance is estimated for the referent type of the subject (L1_Ref) rather than for its individual four levels. This coarse-grained level offers, at least in my view, a more comparable ranking

[81] The maximum number of levels for categorical variables is 32 with RF (Liaw & Wiener, 2002). This limitation is bypassed with dichotomized variables.

[82] The default option in R for encoding multilevel categorical variables is dummy-coding where the levels become *k* -1 lowering the required number of parameters in regression, for example.

of the predictors for future studies. Encoding schemata may vary across studies and certainly not all languages have exactly the same range of levels, such as the number of cases or the number of different types of referents used in a study. Thus, the relative importance of the variables based on the coarse-grain level of granularity is more prone to be either verified or falsified in future studies.

4.2 Random Forests and Usage-Based Models

In recent years, a number of algorithms have been developed that employ classification and regression trees as base-learners connected to the concept of ensemble methods. Instead of operating on a single tree, ensemble methods combine multiple tree models constructed with the base-learner. The final model is the result of averaging over the ensemble. From a linguistic perspective, the concept of ensemble methods is interesting, effectively going against the basic principles of invariance. In theoretical linguistics, a single and most parsimonious representation is typically sought after, aimed to discover a single rule in the generative paradigm or some single function in functional or cognitive linguistics. In ensemble methods, the base-learners are also referred to as weak learners because as individuals they are only slightly better than random guessing. Schapire (1990) proved that weak learners can be boosted to stronger learners leading to the concept of boosting, where weak learners are sequentially added to the model. Interestingly, Goldberg uses this same argument to motivate the existence of multiple cues present in argument constructions. Even knowing the verb can constitute a weak cue simply because there are typically multiple argument realizations available for a particular verb. In order to arrive at the interpretation that the Ditransitive Construction is profiled with the verb *give*, multiple cues are combined. Goldberg goes further by linking the multiple cues to the possibility of modeling them with Adaboost, a well-known machine learning algorithm (Goldberg, 2006:101-102).

Another related method to boosting is to add base learners in parallel. This model is exploited with random forests. Breiman (1996a) introduced the concept of bagging (bootstrap aggregating) where random samples with replacement are drawn from the data and the base learner is trained on them and then averaged over the ensemble. A data point is allowed to appear multiple times in bootstrap sampling increasing the divergence of the sample. Additionally, individual trees are not pruned or stopping criterion is not implemented. Bühlmann and Yu (2002) have shown that bagging typically outperforms classification and regression trees in prediction accuracy. The source of the increase in performance is smoothing. The hard decision boundaries of CART are smoother in the ensemble. This will be illustrated in Section 4.3. From a statistical perspective, the random sampling lowers the variance.

Later Breiman formally introduced the principles of random forests by extending the principle of bagging. An additional layer of randomness is injected with random forests by including random selection of predictors at each split in the tree. This increases the divergence of the trees in the ensemble even further.

Thus, random samples are drawn from the data that are independent of the previous samples and the base learner is trained on them with a random selection of predictors at each split. Another benefit of including the random selection predictors in splitting is that it allows the possible weaker predictors to be included in the tree because a strong predictor might not be available in the split due to the random selection of predictors. These two sources of randomness enables one to by-pass some of the weakness of the individual tree model (Breiman, 2001a). There are a number of extensions of the principles of the random forests but the discussion is limited to two algorithms available in R, package randomForest (Liaw & Wiener, 2002), and conditional random forests in party package (Hothorn, Bühlmann, Dudoit, Molinaro & Van Der Laan, 2006a; Strobl, Boulesteix, Kneib, Augustin & Zeileis, 2008; Strobl et al., 2007). The former is based on the original version by Breiman, whereas the latter builds on the conditional inference trees described in Section 1.4.2.[83]

The prediction accuracy of random forests is competitive with other state-of-the-art machine learning algorithms demonstrated in a number of benchmark studies (Hastie et al., 2009:412-414) and in a wide range of applications in different scientific fields (Arppe & Baayen, 2012; Díaz-Uriarte & Alvarez de Andrés, 2006; Lehmann, Koenig, Jelic, John, Lars-Olof, Dogde & Dierks, 2007; Wu, Abbot, Fishman, McMurray, Mor, Stone, Ward, Williams & Hongyu, 2003).

The performance of random forests is good considering that it is a pure machine learning algorithm. The model is driven inductively from the data. Nonetheless, the performance of the random forests appears to be related to data structure. If data contain a highly complex structure of interactions among the predictors, the performance of the random forests is typically increased. Furthermore, random forests have a number of desirable properties in terms of application. Random forests are essentially an "off-the-shelf" method requiring minimal tuning; whether such methods are preferred or not is another question. The following are the central properties of random forests promoted in the literature (Breiman & Cutler, 2006; Díaz-Uriarte & Alvarez de Andrés, 2006):

1) Can be used for classification with polytomous categorical response variable and for regression.
2) Can handle mixed types of predictors without scaling similar to CART.
3) Can be used in "small *n* large *p*" situations where the number of predictors exceeds the number of observations.
3) Interactions are modeled based on the data similar to CART.
4) Produces a variable importance measure for the predictors.
5) Includes a proximity measure.

The first three properties are critical for the purposes of the present study. First, the chosen algorithm has to be able to handle polytomous response variable.

[83] Another version of the random forests developed further by Breiman and Cutler (2006) is available at www.stat.berkeley.edu/~breiman/RandomForestss/.

Second, the predictors are of mixed-type. Thus, the numeric variables can be included without discretizing them, for instance. It is worth pointing out that the log transformation of the frequency variables is not required, but they are kept as such in the model for consistency.[84] A clear advantage of random forests over the classical regression analysis is that it can be applied to data where the number of predictors exceeds the number of data points. This is certainly one of the contributing factors for the increased use of random forests in genetics and bioinformatics to identify a set of important variables among hundreds or thousands of predictors (Boulesteix, Bender, Lorenzo Bermejo & Strobl, 2012; Díaz-Uriarte & Alvarez de Andrés, 2006; Lunetta, Hayward, Segal & Van Eederwegh, 2004).

In contrast, the classical regression model would require an implementation of some form of variable selection procedure, for example, using univariate methods to find the predictors with the strongest association with the response variable before modeling (cf. Arppe, 2008:116-117). Nonetheless, if the chosen univariate method is conceptually related to the chosen multivariate method, the procedure may lead to bias (Guyon & Elisseeff, 2003). Additionally, potentially higher-order interactions are missed with univariate methods (Lunetta et al., 2004). The inclusion of complex interactions increases the number of required data points in regression because the number of estimated parameters is increased. Finally, complex interactions may lead to a perfect partitioning of the response variable in certain combinations, creating another problematic data issue.

Related to the issue of the variable selection is the potential correlation between them (Cohen, Cohen, West & Aiken, 2003:419-430; Gelman & Hill, 2007:68). In linguistics, variables tend to be correlated, especially in studies of sentential structure, where case marking and syntactic roles are interconnected (cf. Arppe, 2008:116-117), or when a large number of different frequency-based variables are included in the model (cf. Baayen et al., 2006). Baayen considerers that the correlation between variables may be indicative of language processing; when multiple variables tap into the same functional region, the individual burden of a single variable is smaller (Baayen, 2011:209-310). This position does not only relate to linguistic structures but also to common communication situations. For example, missing one or two cues, variables, due to some noise factor do not cause a communication to fall apart because other cues can compensate due to interconnectedness, (i.e., the correlation between the different cues).

From a modeling perspective, the potential correlation between the variables is an undesirable property, but a from linguistic perspective, it can be viewed as an inherent property of language. Nonetheless, these issues are alleviated, at least partly, with random forests because the predictors are modeled one at a

[84] The scaling also concerns the unit scale of measurements, for example, the Entropy and the frequencies differ considerable in their unit scale. In classical regression, these might be required to be brought closer together.

time. The trade-off is that the exact estimations obtained with the classical regression model are lost (cf. Cutler, Edwards, Beard, Cutler, Hess, Gibson & Lawler, 2007:2791-2792). In this vein, random forests can be used to identify potentially important variables which can be later analyzed with other methods, (i.e., used as a data-driven approach to variable selection).

Another important property is the proximity measure. Random forests estimate a proximity measure for the data points in the terminal nodes. It will be used later to construct linking constructions, abstractions over argument constructions, in a data-driven manner. The final property concerns overfitting. Strobl et al. (2009b:33) point out, that random forests being a fairly new approach, some misconceptions are present regarding overfitting when a high number of trees is grown, typically maximally deep. There is very little indication that random forests overfit due on the number of grown trees in the forest (Breiman, 2001a; Strobl et al., 2009b), although the depth of the trees, at least in regression problems, may cause overfitting (cf. Segal, 2004).

In recent years, a number of simulation studies have been conducted to further aid the interpretation and to understand the mechanism behind the performance of random forests (Biau, 2012; Boulesteix et al., 2012; Díaz-Uriarte & Alvarez de Andrés, 2006; Lin & Jeon, 2006; Nicodemus, Malley, Strobl & Ziegler, 2010; Strobl et al., 2007). Random forests are conceptually connected to the principles of weak learnability aligning with Construction Grammar, at least with its cognitively oriented branch.

Related to cognitive plausibility, Baayen (2011) argues that most machine learning algorithms can be viewed as a description of human knowledge, contrasting such methods as regression, nearest the neighbor algorithms, and random forests. This kind of alignment between probabilistic models and human knowledge is demonstrated by Bresnan and Ford with the dative alternation in varieties of English using rating, lexical decision, and sentence completion tasks. The results of the study displayed a concordance with the corpus-based probabilities and human knowledge going beyond a task-specific skill to more implicit sensitivity to probabilistic knowledge. (Bresnan & Ford, 2010:184-185, 191-192, 200-201, 205-206).

Nonetheless, Baayen considers that out of the previously mentioned algorithms only the nearest neighbor, being a memory-based method, may be considered to reflect how language is acquired and used by speakers (Baayen, 2011).[85] In the simplest form, the nearest neighbor algorithm tries to identify prototypes in the data and classify instances based on the distances from the prototypes (Hastie et al., 2009). Lin and Jeon have shown that the random forests can be understood as a weighted version of the nearest neighbor algorithm where the terminal nodes represent the size of the neighborhood. The

[85] In Baayen, a new algorithm, Naive Discriminative Learning classifier (NDL), is proposed that should in theory reflect more closely human learning. Any comparison between NDL and random forests is not made here. I refer to Baayen (2011), and Arppe and Baayen (2012) for discussion on this matter.

random forests assign weight to the data points adaptively based on the input (predictors) available (Lin & Jeon, 2006:579-581). A similar connection is made by Hastie et al. They point out that the individual tree can be viewed as a search for the optimal path to a given target data point and the weight assigned to other data points depends on the proximity to the target data point. The optimal path is not the most optimal. The input and training data are randomized for the base-learners in random forests. Furthermore, Hastie et al. demonstrate that the decision boundaries between the random forests and the nearest neighbor algorithm are similar (Hastie et al., 2009:601-602).

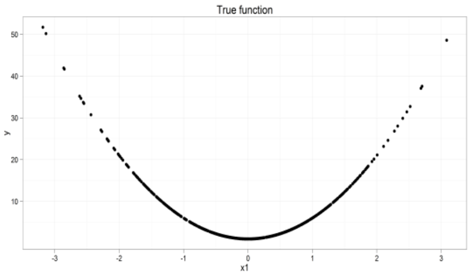
These findings connect random forests to memory-based learning. Baayen, nonetheless, points out that the input made available for a machine learning algorithm is not necessarily available for humans in a similar manner. An example is the range of co-occurring items (Baayen, 2011:296-297). Another factor concerns how the estimations for the predictors are derived from the model. Most of the state-of-the-art machine learning algorithms offer a fairly similar classification accuracy on average (cf. Arppe & Baayen, 2012). However, the relative importance of the predictors estimated by different algorithms may vary considerably (cf. Baayen, 2011:317-318). Keeping these caveats in mind, random forests as a machine learning algorithm may be considered to constitute a form of memory-based learning aligning with the conceptual basis of usage-based grammars (Bybee, 2010; Goldberg, 2006; Tomasello, 2003).

4.3 Extending Trees to Forests

This section focuses on the rationale behind the random forests, contrasting the original algorithm referred to as RF and the conditional as cRF from now on. The advantage of ensemble methods is their ability model smooth surfaces in the data contrary to CART. In Section 1.4.2, the hard decision boundaries of CART were visually compared with the regression analysis. The same example is replicated here for convenience and the same simulated variables *y* and *x1* are used, adapted after Berk (2008). The following models were fitted to the data:

lm(*y* ~ *x1*) (linear regression)
ctree(*y* ~ *x1*) (default settings) (conditional inference tree)
randomForest(*y* ~ *x1*) (default settings) (randomForest algorithm)

The target function and the fitted values extracted from the models are given in Figure 4.3-1.



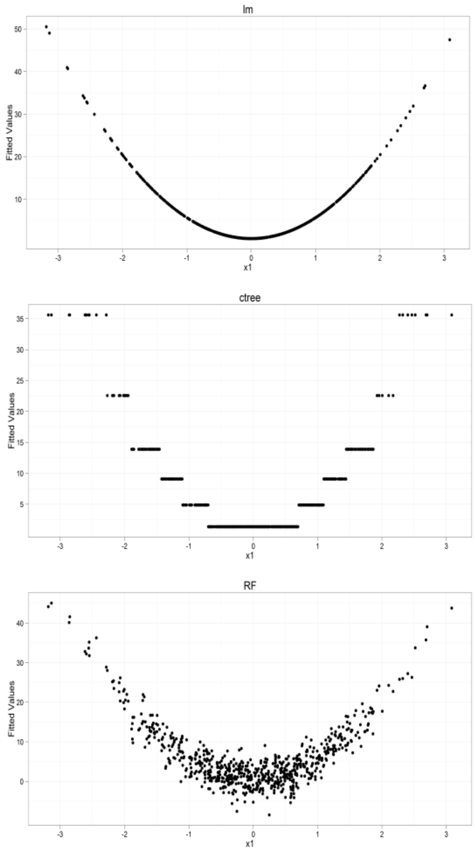


Figure 4.3-1 Scatterplots for estimating target function with three models: Target function (top panel), fitted values from linear regression (second panel), fitted values from conditional inference tree (third panel) and fitted values from RF (bottom panel).

Figure 4.3-1 illustrates the smoothing achieved with RF (bottom panel) compared to the CART model (third panel). Similar to CART, RF is able to inductively learn the function from the data, although the linear regression model still offers a superior fit to the data. In RF, the smoothness is achieved by introducing randomness to the model. Breiman formulated the foundation of random forests by showing that the performance of RF is connected to strength and correlation. The strength is a measure of accuracy of the individual base learners and the correlation is the dependency between them. In order to improve accuracy, the correlation between the based learners has to be minimized and the strength has to be maintained (Breiman, 2001a:7-9). Correlated trees tend to yield similar results, increasing the chance that the mistakes in the classification are repeated across the ensemble. Reduction in correlation increases the divergence of the trees but lowers the chance of repeated classification errors through the ensemble. The first source of randomness is introduced by training the base learners on random samples of the data. The sampling process is illustrated in Figure 4.3-2.

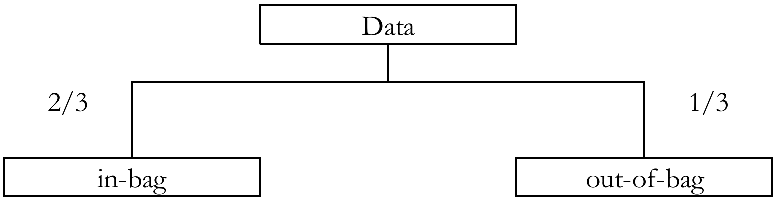


Figure 4.3-2 Sampling process in constructing RF.

The data are randomly split before training with sampling to two sets referred to as in-bag which contains proximately 2/3 of the original data, and out-of-bag (OOB) which contains proximately 1/3. Two standard sampling methods are used. They are bootstrap sampling, where a given data point can appear multiple times, or subsampling without replacement.[86] Subsampling is used in this study and the reasons are explicated when the data are modeled in Section 4.4. The sampling process is repeated for every tree in the ensemble increasing divergence (Hastie et al., 2009:598). Thus, the base learners are only trained on the in-bag data. This procedure also has practical importance. Because sampling is used, a form of cross-validation is built-in to the model. An individual base learner has never seen the OOB data. Thus, it functions as test data, an estimation of the prediction accuracy of the model for unseen data. In classification, the prediction of the RF model is simply a majority vote.[87] For each data point, an estimated class label is assigned by a tree in the forest. The class label receiving the majority of the votes is the estimated class of the RF algorithm. The actual voting process cannot be inspected (Breiman, 2001a:29; Liaw & Wiener, 2002:18).

Breiman (1996b; 2001a:11) has shown that the OOB error is a good, unbiased estimation of the prediction accuracy of the model compared to error estimation that covers the entire learning data. The latter is always an over-optimistic estimation (Hastie et al., 2009:592-593; Strobl et al., 2009b:19, 29; Wu et al., 2003:1642). There are a number of scientific fields where a rigorous model validation is not typically carried out, linguistics being one of them. In linguistics, the data sets are typically fairly small and do not allow data to be split. The optimal situation would be to split the data into three parts: training, validation, and testing. The training data is used to fit the model. The validation data is used to estimate prediction error for model selection and, finally, the test data is used to estimate the generalization error (Hastie et al., 2009:222). Importantly, the OOB-error is only a good estimation of the prediction accuracy of the model as long as variable selection is not carried out, such as training the model with the most important variables estimated with the variable importance measure (Wu et al., 2003:1641-1642). This would only lead to over-optimistic estimation. In addition, there is an important distinction between

[86] Other fractions can be used but these have become fairly standard in resampling.

[87] In regression, the predicted response is the average across the ensemble.

constructing an optimal classifier and finding all the potentially relevant variables. A good classifier need not include all the relevant predictors (Guyon & Elisseeff, 2003). Thus, the RF model used in this study is not necessarily the most optimal RF classifier for the data and certainly not the most parsimonious one. This might be reflected in the classification accuracy of the model. As this is an explanatory study, the model is situated between these two modes. The two modes are prediction accuracy, and the importance of the predictors. In linguistics, the measure of the importance of the predictors is a critical aspect, as it increases the interpretability of the model.

A second source of randomness is injected with the random selection of predictors at each split in the tree, reducing the correlation between the trees that is kept constant in growing the forest (Breiman, 2001a). This is the primary tuning parameter of RF, commonly referred to as *mtry*. The default setting of *mtry* in classification is $\sqrt{p}$, where p is the number of the predictors in the RF model. The second tuning parameter is the number of the grown trees referred to as *ntree,* affecting the stability of RF. These two parameters interact and the goal is to find an optimal value by optimizing the decrease in correlation and maintaining the strength of the individual base-learners. At the same time, RF is not highly sensitive to these tuning parameters compared to certain other machine learning algorithms such as boosted trees (Díaz-Uriarte & Alvarez de Andrés, 2006:4-5; Hastie et al., 2009:591-592). These parameters can be tuned by inspecting the stabilization of the estimated OOB-error in the forest (Breiman, 2001b).

The optimal value of *mtry* also depends on the quality of the predictors. In case the data contains a large number of irrelevant or weak predictors, a larger value of *mtry* may be required, (i.e., to compensate the signal-to-noise ratio). Also, higher-order interaction can only be included with a larger value of *mtry*. Finally, the value of *ntree* is connected to the variable importance measures (Strobl et al., 2009b:32), and to the proximity measure (Shi & Horvart, 2006). A larger value may be required to stabilize them, especially the advanced importance measures because they also introduce another layer of randomness. The final potential tuning parameter is the size of the terminal nodes in RF. In classification, maximally deep trees are gown, (i.e., the default value of the node size is 1). Based on recent studies, there is some indication that the maximum depth of the trees is not necessarily the most optimal value. Examples consist of Segal (2004:10-11) in regression task, andLin and Jeon (2006:582-583) with large but low-dimensional data. In contrast, Díaz-Uriarte and Alvarez de Andrés (2006:4) demonstrate with real and simulated data sets that the node size between 1 and 5 appears to be a fairly inconsequential tuning parameter in terms of prediction accuracy. Their study is one of the few where polytomous models were included.

4.4 Tuning the Parameters of the Random Forest Model

For purposes of modeling, the two random forests algorithms of RF and cRF were considered. There are fundamental differences between them. RF uses bootstrap sampling by default and cRF uses subsampling Strobl et al. demonstrate with simulation studies that the bootstrap sampling may introduce bias. The bootstrap samples may artificially include associations not present in the data. When the base-learners are trained on the samples, the bias may be carried on (Strobl et al., 2007:17-18). A second difference is the choice of the base-learners. RF uses the standard CART without pruning, whereas cRF uses the conditional inference trees discussed in Section 1.4.2. By default, cRF has a stopping criterion in place. This has also a practical component involved with cRF, namely the computational time is reduced when trees are not grown to maximum depth. The algorithm is already computationally intensive.

The standard CART has been shown to be biased towards variables with multiple levels or cut-off points. Variables with these characteristics may be artificially preferred (Kim & Loh, 2001; Strobl et al., 2007). From this perspective, the dichotomization of the categorical variables might appear beneficial. However, there are a number of reasons why dichotomization might not correspond to better performance. First, trees become more complex and sparse simply due to the increased number of required splits with binary variables compared to multiple ones. Second, the interaction between the binary splits may lead to more spread-out structure of the terminal nodes (Kim & Loh, 2001:589-590). Both of these factors are undesirable, especially the latter when the size of the terminal node size is used as proxy to model the distances between the argument construction types.

The potential bias with the standard CART is a complicated matter. Generally, bias and variance are inversely related. Additionally, they interact with the complexity of the model. The training error of the model tends to decrease with increased complexity that can lead to overfitting. Predictions from this model tend to have high variance, (i.e., decreased prediction accuracy). In the opposite case, the model may underfit resulting in increased bias, resulting in decreased prediction accuracy. Thus, the goal is to find a trade-off between them (Hastie et al., 2009:37-38, 52, 220). If certain variables are artificially preferred based on their properties and not relative to the response variable, the ability to generalize to unseen data may be affected. This property of the standard CART may affect the classification accuracy of the model. Sources of this bias are shown in Section 10.1.1.

Before finding the optimal tuning parameters for the data, both RF and cRF were fitted to the data with subsampling. All other parameters were kept at their default values, those being *mtry* at 5 and *ntree* at 500. Another practicality when using random forests is that the models are stochastic. In order to keep results comparable between different runs, random seed was specified.[88] Random seed 12345 was used throughout this study when required. The following models,

[88] By default, random seed is based on system time in R.

with Construction as a function of all the predictors (25), were fitted with default settings:

randomForest(Construction ~ ., replace=F, data = dat)

cforest(Construction ~ ., data = dat)

The algorithms showed a drastic difference. RF required only 31s to fit the data contrasting the required time of 23min with the cRF on a desktop computer (AMD X6 1055T, 3.5 GHz, 16 GB memory). Crucially, the OOB-error with cRF is 28% in contrast to 18% with RF. The drastic difference in performance might be due to data sparseness or the stopping criterion implemented in cRF. The base learners of cRF may simply require more data points. The difference might also be related to the classification problem itself, namely the polytomous response variable. Investigation into this issue would, however, require additional data sets. Another practical aspect is related to the required computational times. When a larger number of trees are fitted, the computational time increases drastically with random forests. In order to arrive at a stable estimation of the advanced variable importance measures, results should be reported based on resampling rather than on the estimation obtained from a single ensemble solution (cf. Nicodemus et al., 2010). Resampling also brings forth the uncertainty in estimations rather than relying on estimations obtained from a single model (Gelman & Hill, 2007:137, 457-459). For these reasons, the classical random forests are used throughout this study as the chosen algorithm.

In order to tune the *mtry* parameter, several models with *mtry* ranging from 1 to 25 were fitted (Strobl et al., 2009b). The *ntree* was kept constant at 500 and subsampling was used. Figure 4.4-1 shows the change in OOB-error across the different *mtry* values.

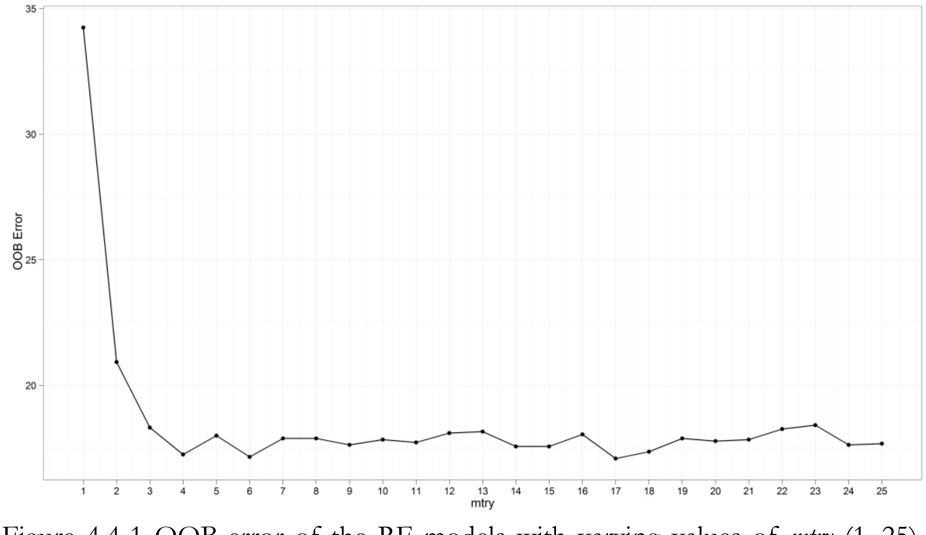


Figure 4.4-1 OOB-error of the RF models with varying values of *mtry* (1–25) and a constant *ntree* of 500.

The OOB-error decreases drastically and plateaus after *mtry* value 3 demonstrating that RF is fairly insensitive to the tuning parameter, at least in this data set. The OOB-error ranges between proximately 17% and 18% throughout the different models. Thus, a new model was fitted with an increased *ntree* parameter of 2,000 to ensure that enough data points are available for the proximity measure. The proximity measure is utilized later in Section 10.2 to construct connections between constructions. Additionally, the *mtry* was increased to eight to enable potential higher-order interactions in the model.

```
randomForest(Construction ~ ., importance=T, proximity=T, ntree=2,000,
        mtry=8, replace=F, data = dat)
```

The additional parameters in the formula, importance and proximity, specify that the variable importance measures are to be calculated and the proximity measure between the data points in the terminal nodes is to be included in the model. Finally, the parameter replace is set to false, specifying that subsampling is used; a random sample of the data is taken without data points appearing multiple times in the sample.

Figure 4.4-2 shows the behavior of the OOB-error during the model fitting when base-learners were added to the ensemble.

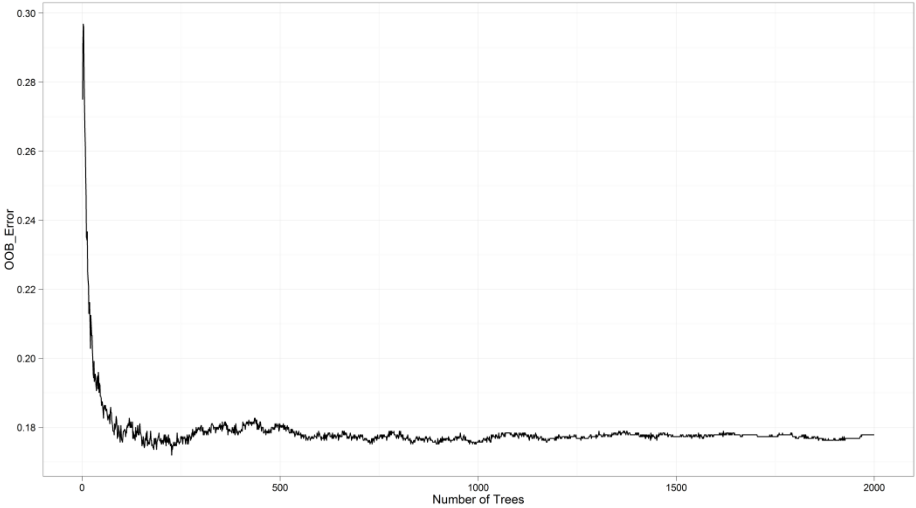


Figure 4.4-2 Behavior of OOB-error during the fitting of the RF model when base leaners are added to the ensemble with *mtry* of 8 and *ntree* of 2000.

Similar to the *mtry*, the OOB-error decreases drastically and plateaus around the default value of 500. An optimal value of *ntree* could be empirically tested within the plateau (between 500 and 2,000 trees). However, the benefit of the procedure would only be related to fasten computations with additional calculations with the RF model, as the stability of the optimal *ntree* value would have to be tested against the advanced importance measures. Because resampling is used in this study, several thousand RF models would have to be fitted with varying *ntree* values. A smaller value than 2,000 would lower the required amount of memory and hard-disk space. The latter is inconsequential in modern times and the former would only be important with larger data sets. Thus, the primary tuning parameters of the RF model appear to be stable and the subsequent discussion of the argument construction types is based on this RF model with the *mtry* value of 8 and the *ntree* value of 2,000.

It is important to remember that the final model is fairly complex based on 2,000 base-learners. Figure 4.4-3 shows the structure of the ensemble with the counts of all the nodes (upper panel) and of the terminal nodes (lower panel).

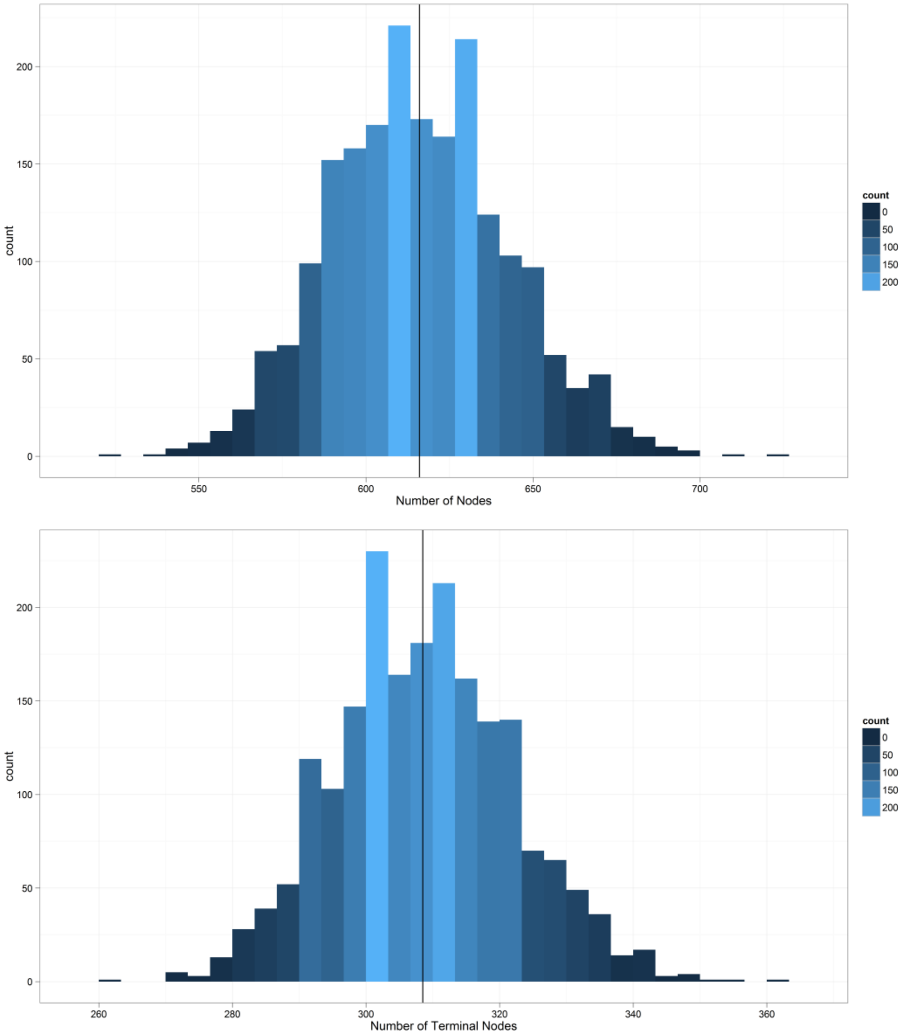


Figure 4.4-3 Histrogram of number of nodes in the final RF model colored by counts. The vertical line shows the mean value. The upper panel shows the total node count (mean = 616) and the lower panel shows the terminal node count (mean = 309).

On average, there are 616 nodes (min. = 523 and max. = 723) and 309 terminal nodes (min. = 262 and max. = 362) per tree. The model operates on imperfect, highly skewed trees embedded in layers of random sampling. The performance of the model can only be assessed as an ensemble. Thus, there is no best tree solution that can be obtained.

The final RF model achieves an estimated prediction accuracy of 82.22% to unseen data, (i.e., OOB-error of 17.78). Although the performance of the model is certainly a goal in itself, the interpretability of the model is important for theoretical linguistics. At the same time, the estimated prediction accuracy also entails that certain construction types do not readily arise from the linguistic model based on the RF algorithm. Certain construction types are also easier to learn than others, and this is not reflected in the estimated prediction accuracy of the model. Furthermore, certain construction types appear to be exhaustively captured by the underlying linguistic model. These types inflate the estimated accuracy of the model. These types are not, removed from the model because the estimated class probabilities are required for building the generalizations, and linking constructions over the argument constructions in Chapter 10. Considering that the response variable consists of 13 argument construction types, validation measures are required to assess the performance of the model in terms of individual argument construction types.

The random forest algorithm with its parameters is defined, adapted after Hastie et al. (2009:588):

```
For $b = 1\ to\ \mathcal{B}$:
    Draw a subsample from the data
    Grow a random-forest tree $t_b$ to the bsampled data
    Recursively repeat the following steps for each terminal node until the
    minimum node size is reached (default 1 in classification)
    Select randomly $m$ (8) predictors at each split from the $p$ (25) predictors
    Select the best predictor among the $m$
    Split the node into two daughter nodes
    Output the ensemble of trees $\{t_b\}_1^{\mathcal{B}}$
    Prediction in classification at data point $x$: the class prediction of $bth$
    random forest tree is $\hat{C}_b\,(x)$. Then $\hat{C}_{rf}^{\mathcal{B}}(x)$ = majority vote $\left\{\hat{C}_b\,(x)\right\}_1^{\mathcal{B}}$
```

In sum, this section demonstrated the process and practicalities of tuning and fitting RF models. Consequently, a usage-based grammar has been implemented that operates on weak input, structural properties, and combines weak base-learners to inductively learn the argument constructions. From this perspective, both the theory and the methodology are, to a degree, in harmony, grounded in weak and memory-based learnability.

4.5 Evaluation Measures for the Random Forest Model

The prediction and classifications models need to be kept separate in evaluating the performance of the RF model. A simple example demonstrates the difference. The example is a response variable consisting of two levels, A (*n* = 90) and B (*n* = 10). A prediction model is fitted to the data and the category A was predicted to occur *n* = 100. The prediction model would achieve a prediction accuracy of 90%. In contrast, the classification model is simply a failure because class B had zero predicted occurrences.[89] Prediction accuracy reflects the overall performance of the model, but in the case where the classes are highly imbalanced,, the prediction accuracy in itself is not a good indicator of the performance of the model. In order to assess the performance of the model in terms of individual argument construction types, a confusion matrix can be constructed to assess the predicted responses. A confusion matrix is a K by K table where the rows contain the observed class labels, *k*, and the columns contain the predicted class labels. The table is 13 by 13 in this case. The confusion matrix is included in the random forests object by default along with the class-wise errors. Importantly, the confusion matrix is based on predictions to the OOB-data and not to the training data. Thus, it is an estimation of the performance of the model for unseen data.

For polytomous response variable, measurements for a single class can be evaluated through dichotomization, (i.e., the one versus all approach). Table 4.5-1 illustrates this approach along with four classification types.

	Predicted	
Observed	Class A	¬ Class A (All other)
Class A	TP = True Positive (correct)	FN = False Negative (incorrect)
¬ Class A (All other)	FP = False Positive (incorrect)	TN = True Negative (correct)

Table 4.5-1 Confusion matrix for dichotomous outcomes.

Class A represents the construction type under inspection, whereas all the remaining types are collapsed into one group (not class A). The counts on the diagonal are correctly classified instances. The true positives are the correctly classified instances of the Class A, and the true negatives are the correctly classified instance not of Class A. The counts off the diagonal are misclassification. The false negatives are instances of the Class A, but they are classified incorrectly, whereas the false positives are incorrectly classified as instances of the Class A.

The four types form the basis of evaluation measures that inform about

[89] The difference between these two models is slightly more nuanced. With the classification model, an assumption is made that the estimated classes are truly different and the goal is obtain a model that would reflect the observed proportions of the classes based on the input. In contrast, the prediction model does not entail this assumption. Thus, we would accept the outcome that there are no differences between the classes. (cf. Arppe, 2008:129-130 for elaborated discussion).

different aspects of the performance of the model. The standard measures of precision and recall in information retrieval are used. They inform us about the performance of the positive classified classes (Sokolova & Lapalme, 2009). Precision is the proportion of the total number of predictions for a certain class that were correctly classified. The proportion is TP/(TP+FP). Recall or true positive rate is the proportion of positive cases for a certain class that were correctly classified. The proportion is TP/(TP+FN). Precision focuses on agreement on the positive cases, whereas recall focuses on identification of the positive cases (Manning & Schütze, 1999:268-269). Thus, these measurements inform us about the performance of the model in terms of classifying and identifying instances.

Other commonly used pairings of performance measures are sensitivity and specificity. The proportions are (TP/(TP+FP) and (TP/(TP+FP), respectively. Precision and sensitivity are identical. On the other hand, specificity informs about the ability to indentify negative classes. For polytomous response variables, specificity is, however, an uninformative measure. The negative class does not constitute a single category with a unifying property, cf. Table 4.5-1. It is the sum of the instances not of class A (Arppe, 2008:131; Sokolova & Lapalme, 2009:432). Thus, the measure is not used.

The previously outlined measures describe the classification either as correct or incorrect, but they do not include the degree of separation of a data point in the model. Binary classification masks the degree of potential competition between different categories. Estimated class probabilities can be used to focus on this aspect of the classification. These plots are constructed by using the fraction of the votes a particular data point received in the ensemble (Liaw & Wiener, 2002). For those that enter into competition, the probabilities start to fluctuate (Lin & Jeon, 2006). Estimated class probability plots are used to visualize the potential competition between different data points. This perspective also meshes with the tenet in Cognitive Linguistics that categories are continuous rather than discrete (Bybee, 2010; Goldberg, 2006; Langacker, 1988b). Usage patterns are not static but dynamic, and the estimated class probabilities enable to highlight this aspect. This visualization technique is introduced in Section 5.2.

5 Focus Constructions

The Transitive Construction constitutes the primary type in constructionist accounts to establish argument constructions (Divjak & Janda, 2008; Goldberg, 2006; Janda, 2008b). Following Langacker (1991:283-284) the basic properties of the Transitive Construction consist of "a person who volitionally initiates physical activity resulting, through physical contact, in the transfer of energy to an external object." This yields an inherently asymmetrical relation and the basic semantic roles of Agent and Patient. In traditional terms, the reconfiguration of the Transitive Construction yields the Passive Construction where the syntactic positions are reversed. These two types occupy a central position in most linguistic theories, while intransitive and impersonal types typically occupy a secondary place. Blevins (2003) calls this a tacit descriptive bias in linguistic theories.

Chomsky's early generative approach posited that the passive construction is a derived construction type formed from the corresponding active structure, (i.e., the transitive), based on a transformation rule. To account for the active/passive distinction, the transformation rule account posits a single lexical entry for a verb with the same underlying syntactic structure, (i.e., deep structure). Thus, the active structure is considered to be a direct projection of the deep structure, whereas the passive is a derived structure from the underlying one (Chomsky, 1957; 1965). The transformation rule would imply that the two structures in question are interchangeable. This same position is also upheld in the early Russian transformational grammar (e.g., Adamec, 1973). A weaker assumption would be to state that they are not necessarily fully interchangeable in every context, instead their propositional content is considered to contain a level of similarity (cf. Коломацкий, 2009:23-24). However, studies on discourse properties demonstrate that the Passive Construction occupies its own niche which cannot be reduced to a simple transformation rule (Croft, 1991; Fried, 2006; Givón, 1994; Israeli, 1997; Shibatani, 1998).

A similar position is already taken by Langacker and Munro (1975) and is analogous to the description also upheld later in Cognitive Grammar in Langacker (1991; 2002). Traditionally in syntactic accounts, the passive is defined in terms of either promotion or demotion. The former refers to the promotion of the object to subject position and the latter to the demotion of the subject to the oblique position (Grimshaw, 1990; Siewierska, 1984). A similar position is taken in the Russian diathesis tradition from the very beginning (cf. Храковский, 1974; 2004). In contrast, the promotion account highlights the opposite pathway of derivation, that is, the object of the transitive verb is promoted to the subject position in the passive. These positions make different predictions about the nature of the passive. The promotion account cannot be used to motivate the existence of the patterns where the object does not obtain the subject position, but instead retains its object morphology, accusative case, and there is no agreement as in 4.5-1.

4.5-1 *Школ-у построе-н-о.*
School-ACC build-PPP-N
The school was built.
(Коломацкий, 2009:30)[90]

These forms are actively used in Russian dialects and demonstrate the impersonal passive construction type attested in other Slavic languages, like Polish and Ukrainian (cf. Зельдович, 2010:8-9; Коломацкий, 2009:30). However, the impersonal type is only attested with the periphrastic passive participle type in Russian. Typically, two prototypical passive constructions are posited for Russian, namely the periphrastic passive participle and the reflexive passive. The former is formed with the markers *-т-,* and *-н-*. These forms are stative and typically include a resultative state in their semantics, following the general tendency observed by Haspelmath that the resultative participles possess a high tendency to develop into passive participles (Haspelmath, 1994:161-162).

Generally, the notion of demotion is employed in functional and cognitive tradition and it is used to cover a relatively wide array of different construction types ranging from the passive and other related changes in argument structure to impersonal types (cf. Divjak & Janda, 2008; Solstad & Lyngfelt, 2006:8). In this vein, Divjak and Janda (2008) regard the demotion as one of the central features of Russian argument constructions. This view on the relatedness of argument constructions subsumes a mapping stemming from the canonical Transitive Construction covering peripheral impersonal types, also. At the same time, both of these notions, demotion and promotion pertain to a single invariant feature of construction types on a highly schematic level, which cannot be used to motivate all the construction types of the Russian Reflexive Marker.[91] Thus, this section addresses argument construction types that are primarily connected to focusing the profile on a certain facet of the action and typically a single entity is construed as occupying the prominent position.

5.1 Definitions of the Passive Construction

In functionally oriented studies, the Reflexive Passive is typically regarded to constitute its own type and not simply reflecting some underlying derived pattern. Fried considers, based on her study of the passive participle and

[90] Translation and glossing was added by the author.

[91] In the Prague school, the concept of passive, is used to cover a wide array of different patterns, which are considered to be related. For example, Adamec (1973:118-124) connects various impersonal patterns with the dative case marking and the reflexive marker in Russian to the category of demipassive. The nominative subject argument of the transitive verb is transformed to the category of demipassive by encoding it with the dative case. Additionally, the label quasi-passive is often evoked in Russian tradition to cover verbs expressing some inherent quality of the subject argument (Geniušienė, 1987:261; Виноградов, 1972:498).

reflexive passives in Russian and Czech that the discrepancy between different approaches to these two passive constructions stems from the fact how they are traditionally framed theoretically. That is two morphologically distinct, albeit unrelated constructions seem to give rise to similar effects. She postulates that the two passive constructions are used to convey different communicative purposes (Fried, 2006:84-85).[92] Similarly, Israeli (1997:181) observes that the passive is not interchangeable with transitive variants in usage, attributing the difference to text cohesion. The latter function, however, has already been established in the Russian functional tradition (cf. Коломацкий, 2009 and references therein).

In addition to the question of function of the Passive Construction, its exact scope is another important aspect. It is considered to be a morphologically distinctive form consisting of a nominative argument (Patient), a verb marked with the Reflexive Marker, and an optional argument profiled with the instrumental case, (i.e., the secondary slot or Agent). This position in defining the Passive Construction builds on the early diathesis tradition (cf. Храковский, 1981:6-11). This morphologically driven approach is also encountered in lexicography (e.g., Тихонов, 1998:244-245) and in Russian Functional Grammar (e.g., Бондарко, А. В., 2002:595-596). Additionally, this definition follows the commonly held view in typological studies that the existence of the Passive Construction presupposes the existence of passive morphology in language (Haspelmath & Müller-Bardey, 2005). A broad definition of the Passive Construction also includes dative arguments as possible candidates for encoding the secondary slot.[93] This position is taken, for example, by Kolomackij (Коломацкий, 2009) as illustrated in 5.1-1. Similar view is expressed in Percov (Перцов, 2003:66) and Hrakovskij (Храковский, 1991). Additionally, Miloslavskij (Милославский, 1978) argues that verbs, such as *нравиться* ‘please, like’ pertain to the Passive Construction, although it does not have a neighbor verb in Russian and always combines with the dative argument.

5.1-1 *Мне* *дума-ет-ся,* *что* *надо* *этот*
I.DAT think-3S.PRS-RM that have.to this.ACC
праздник *упраздни-ть.*
holiday.ACC set.aside-INF
I think this holiday needs to be set aside.
(Коломацкий, 2009:42)[94]

Gerritsen (1990:27-28) distinguishes these instances from the Passive Construction and labels them as Medial-Passive. The verb *думать* ‘think’ and

[92] This mode of analyzing alternating patterns of voice in terms of information or focus structure is illustrated in Haiman (1991).

[93] In addition to dative arguments, various prepositions or prepositional phrases such as *через* ‘over’, *от* ‘from’ and *под влиянием* ‘under the influence’ are also considered to instantiate the Passive Construction (Тихонов, 1998:249-250).

[94] The glosses and the translation were added by the author.

semantically similar ones, like the mental verbs of *волновать* ‘worry’ and *пугать* ‘frighten,’ deviate from the canonical cross-paradigmatic relation between the Transitive and the Passive Construction as they do not appear in the Reflexive Passive Construction in Russian supported with the Nominative-Instrumental pattern (cf. Падучева, 2004). The inclusion of the dative pattern would broaden the definition of the Passive to cover constructions typically classified as impersonal, also. Whether the impersonal types are lumped under the label Passive Construction is a matter of perspective pertaining to the level of imposed granularity. However, the personal Passive Construction is in itself well-established as a type in Russian. Furthermore, the narrower definition maintains the traditional cross-paradigmatic relation between the Transitive and the Passive Construction (Буланин, 1986; Зализняк, А. А., 1980; Исаченко, 1960).

In addition to the cross-paradigmatic relation, the exact profile of the Passive Construction becomes narrower when impersonal patterns are excluded. Janko-Trinickaya (Янко-Триницкая, 1962:125) gives the following ranking of patterns in terms of typicality: *что-кем* ‘what-by-whom’ → *что-чем* ‘what-by-what’ → к*т*о-*кем* ‘who-by-whom’ → *кто- чем* ‘who-by-what.’ According to the ranking, a typical instance of the Passive has a non-human referent type occupying the subject slot. At least for these data, the affinity towards non-human subject referents appears to hold the following proposed ranking (cf. Israeli, 1997:169-170). Out of 210 instances, only 16 have either Person or Animate referent type, as illustrated in 5.1-2 and 5.1-3.

5.1-2 *В собор-е* *коронова-л-и-сь* *польск-ие*
PR cathedral-PREP crown-PST-PL-RM Polish-NOM.PL
правител-и, […].
leader-NOM.PL
The Polish leaders were crowned in the cathedral.
[749, RNC, Польша - добрая соседка: здесь примут по-домашнему! //"Даша," №10," 2004]

5.1-3 *Стар-ый* *суд* *ликвидирова-л-ся,* […].
Old-NOM judge.NOM eliminate-PST.M-RM
The old judge was eliminated.
[213, RNC, Александр Афанасьев. Суд присяжных в России // "Отечественные записки," 2003]

Examples 5.1-2 and 5.1-3 also show the common tendency that the secondary slot is typically left without explicit profile, although it is considered to be an inherent property of the construction type (cf. Коломацкий, 2009:238).[95] The canonical profile is illustrated in 5.1-4 and 5.1-5.

[95] Several possible invariant properties have been offered to account this. For example, Israeli posits a single feature [-responsibility] following Leinonen (1982) and Siewierska (1984).

5.1-4 *Нужно было в кратчайшие сроки обустроить месторождение, получить*
[…]
газ и всё, что здесь создавалось,
[…]
создава-л-о-сь классн-ыми специалист-ами и
create-PST-N-RM first.rate-INS.PL specialist-INS.PL and
культурн-ыми людь-ми.
educated-INS.PL people-INS.PL
It was necessary to arrange as soon as possible a deposit of the gas, to get in place and everything that was created here, was created by first rate specialists and educated people.
[631, RNC, Надежда Шагрова: "Я - мал" ищет единомышленников // "Экран и сцена," 2004.05.06][96]

5.1-5 *Это объясня-ет-ся специфик-ой* […].
It.NOM explain-3S.PRS-RM characteristics-INS
It is explained by the charasteristics
[64, RNC, И.А. Барков. Автоматический синтез структурного описания конструкции // "Информационные технологии," 2004]

Another issue is related to the semantics of the secondary slot, as in 5.1-5. Hrakovskij (Храковский, 1991:180) also considers instrument-like entities as candidates for the secondary slot, for example *синим туманом* 'by the blue mist.' Certainly, the entity profiled in the secondary slot does not instigate the event in 5.1-6.

5.1-6 *Наличи-е в популяци-и человек-а вирус-а*
Presence-NOM PR population-PREP human-GEN virus-GEN
осп-ы сопровожда-л-о-сь клиническ-и
smallpox-GEN accompany-PST-N-RM clinical-ADV
выраженн-ым заболевани-ем.
evident-INS disease-INS
The presence of smallpox virus in the human population was accompanied by clinically evident disease.
[96, RNC, Ликвидация полиомиелита и роль вакцинных вирусов в этом процессе // "Вопросы вирусологии," 2007]

Instead of classifying these instrument-like secondary slots as part of the Passive Construction, Kolomackij (Коломацкий, 2009:232-233) considers that these instances contain a null Agent, left unexpressed, that is still part of the construction. Yet instrument-like entities can compete with Agents, an observation stemming from the very beginning of the studies on semantic roles (Fillmore, 1968; 1970).

Example 5.1-7 was specifically extracted from the Russian National Corpus to illustrate the issue at hand where *дым* 'smoke' can certainly occupy the

[96] Example describes the events of how the city of Novyj Urengoj was established.

subject slot in the Transitive Construction. Paducheva (Падучева, 2001:56) analyzes these patterns as a specific type of derivation, (i.e., Deagentivization), cf. also Section 3.1.10 on causative neighbor verbs.[97]

5.1-7 *Черн-ый* *дым* *заволакива-л* *центр* город-а.
Black-NOM smoke.NOM cover-PST.M center.ACC city-GEN
The black smoke covered the city center.
[RNC, Смерть Человека (2003) // «Вслух о…», 2003.10.24]

Perhaps the most controversial issue related to the Passive Construction is the role of aspect. Whether the Passive can be formed with the perfective aspect along with the imperfective divides the Russian linguistic tradition. Paducheva (Падучева, 2001:53) explicitly states that the perfective aspect is ungrammatical in Russian in the Reflexive Passive Construction. A similar position is also taken by Apresyan (Апресян, Ю. Д., 2002:19) and in most formal approaches (cf. Fehrmann et al., 2010:211; Guhl, 2010:264). At the same time, Paducheva (Падучева, 2001:74) acknowledges the existence of the perfective Passive Construction in the history of Russian, citing Bulahovskij (Булаховский, 1954). In contrast, certain scholars argue for the existence of the perfective Passive Construction even in contemporary Russian, such as Janko-Trinickaya (Янко-Триницкая, 1962:132-134), Percov (Перцов, 2003), and Zel'dovich (Зельдович, 2010).[98]

Percov tracks down the various positions upheld in the Russian linguistic tradition, offering a detailed discussion on this topic. Furthermore, he illustrates that even in contemporary Russian the perfective aspect is used in the Passive Construction and it is by no means an anomaly (cf. Перцов, 2003:66-70). Additional support for this view is offered in Kolomackij's corpus-based study showing a relative high frequency (30%) for the perfective aspect. (Коломацкий, 2009:117-118, 128). Examples 5.1-8 and 5.1-9 illustrate the usage of the perfective aspect in the Passive Construction.

[97] This derivation is not morphologically marked and broadens the definition of diathesis alternation.

[98] Zel'dovich considers that the Passive is typically anomalous when used with the perfective aspect based on a questionnaire study. However, details about the study are not given. The only possible clue is stated in footnote 2 (Зельдович, 2010:5), (i.e., the participants were, primarily, highly qualified scholars). Considering this fact, the results may simply show a prescriptive preference. Similar observations concerning questionnaire studies on low-frequency items and possible prescriptive influences are discussed by Barðdal (cf. 2006:69, 100).

5.1-8 *В Российск-ой Федерац-ии* законодательн-ый
PR Russian-PREP Federation-PREP legislative-NOM
статус ОВОС утверди-л-ся в 1994 г.
status.NOM EIA establish-PST.M-RM PR 1994 y.
The legislative status of EIA was established in 1994 in the Russian Federation.
[396, RNC, Геоинформационное картографирование для оценки воздействия на окружающую среду объектов нефтегазовой промышленности // "Геоинформатика," 2001.03.14][99]

5.1-9 *Конференци-я откры-л-а-сь с обсуждени-я*
Conference-NOM open-PST-F-RM PR discussion-GEN
проблем-ы
problem-GEN
установления границ контроля государства за деятельностью
[…]
религиозных объединений.
[…]
The conference was opened with a discussion on the problem of forming the boundaries of state control over the activites of religious associations.
[528, RNC, Мария Козлова. Свобода совести и светскость государства: проблемы и решения (1 часть) // "Адвокат," 2004.12.01][100]

One solution to the issue is to simply regard them as a separate type, typically labeled as Potential Passive or some derivate from it (Paducheva, 2003:185; Князев, 2007).[101]

Example 5.1-10, specifically extracted from the internet, demonstrates the Passive usage with the reflexive verb *открыться* 'open'.

5.1-10 *После приветственн-ых слов, конференци-я*
PR welcoming-GEN.PL word.GEN.PL conference-NOM
откры-л-а-сь пленарн-ым заседани-ем, […]
open-PST-F-RM plenary-INS talk-INS
After the welcoming words, the conference opened with a plenary talk.
[http://nami.ru/press/news/437/]

The controversial nature of the perfective aspect in the Passive Construction is apparent even in more recent studies. For purposes of the present study, these instances are merged under the label Passive Construction.

[99] OBOC = EIA (Environmental Impact Assessment).

[100] For unknown reasons, the wider context is not available anymore in the Russian National Corpus (26.04.2012).

[101] Knyazev (Князев, 2007:291) has conveniently collected the stock of the most common labels associated with this pattern.

5.2 Passive Construction

The Passive Construction contains 208 data points in the RF model and covers 123 unique reflexive verbs. Traditionally, the Passive Construction is considered to be the most productive of all the types marked with the Reflexive Marker in Russian. The only relative frequent verbs are *подниматься* 'rise,' *делаться* 'make,' and *использоваться* 'use' in the sample, indicating that the Passive is primarily a type with minimal lexical support. The definition of the Passive Construction follows.

Function: Profiles a relation of a focused entity and a backgrounded entity.
Form: Nominative subject and secondary slot in the instrumental case.

The function of the Passive Construction follows the definition given by Fried (2006) in a sense that one entity appears to lie in focus and the second entity is backgrounded. The function is also reflected on the form pole supported by the Nominative-Instrumental pattern. Additionally, the Passive forms a bridge between the verb paradigms that is also supported by semantic similarity across paradigms. The exception to this property appears to be the verb *даваться* 'give,' which was perceived dissimilar to its neighbor verb, *давать* 'give', indicating that the primary sense of the reflexive verb is detached from its neighbor and it leans towards other senses of the verb, possibly 'succeed' or 'acquire.' These senses combine with the dative argument and are delimited to third person (Кузнецов, 2009 [1998]).

Example 5.2-1 illustrates the Passive Construction and the encoding is given in Figure 5.2-1.

5.2-1 *Рассматрива-ют-ся* *вопрос-ы* […].
examine-3P.PRS-RM question-NOM.PL
Questions are being examined.
[10, RNC, Векторная оптимизация в проектировании сложных \
изделий на примерах выбора вариантов реактивного двигателя //
"Информационные технологии," 2004]

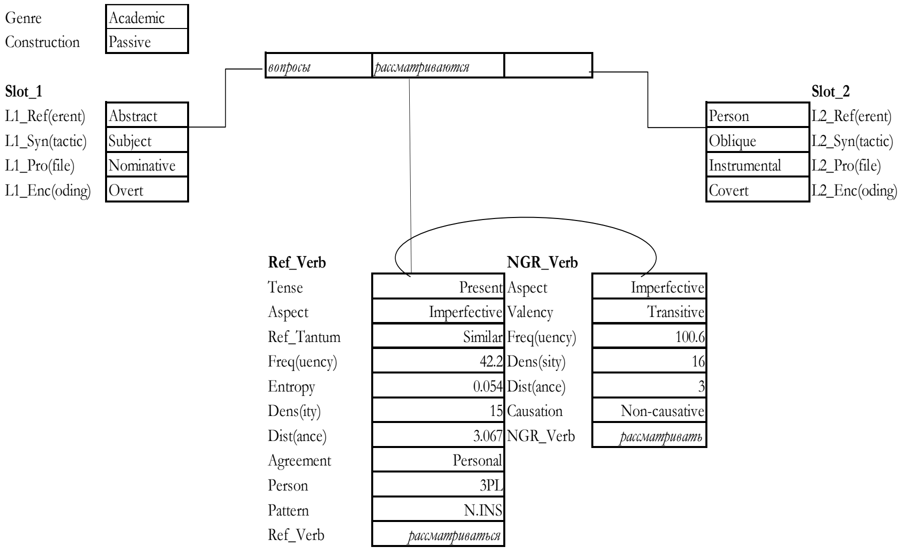


Figure 5.2-1 Layered structure of the canonical Passive Construction.

Table 5.2-1 gives the confusion matrix of the predicted construction types, as outlined in Section 4.5. The observed construction types are given in the rows and the columns give the predicted types by the RF model. The correctly classified instances are located on the diagonal in general. For every instance of the Passive Construction, the RF model predicts a class label out of the total number of labels available in the model (13). The correctly classified instances, 206 in this case, are found in the intersection of the observed and predicted. The misclassifications are located off the diagonal.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	**0**	**0**	**0**	**0**	**0**	**1**	**204**	**1**	**0**	**0**	**2**	**0**	**0**
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 5.2-1 Confusion matrix of the predicted constructions. The Passive Construction is given in bold.

The class-wise error of the Passive is 0.019. However, this error does not reflect how well the category as a type is represented compared to the other available types in the model. Based on the discussion in Section 4.5, in order to evaluate this property of the argument constructions, the Passive Construction is contrasted against all other construction types in the model. The recall is 0.98 and precision is 0.94, indicating that the RF model is able to identify and classify the instances of the Passive Constructions across the global patterns available in the sample.

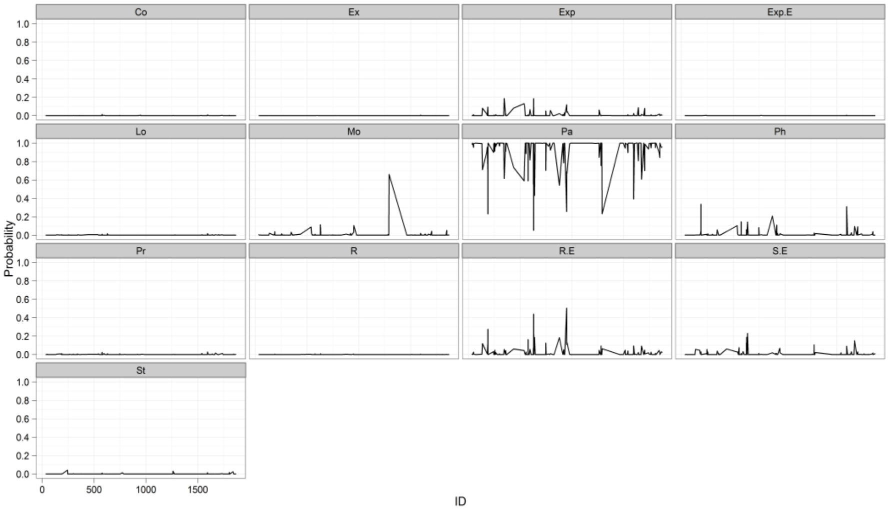


Figure 5.2-2 Faceted estimated class probability plot of the Passive Construction (Pa). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

In addition to estimating the goodness of the classification, the estimated class probabilities were proposed to offer a more fine-grained estimation for individual data points in Section 4.5. From a usage-based perspective, a certain instantiation can still pertain to the Passive, but the exact profile of the instantiation may be further away from the canonical instantiation type of the Passive (a degree membership). In addition to assigning a label to a data point, its estimated class probabilities across all the types available in the data are included in the model. Thus, the RF model estimates a probability, a fraction of votes, of a data point being any of the construction types in the sample. These estimated class probabilities of the Passive were plotted in Figure 5.2-2. The y-axis gives the probability and the facets contain the estimated classes.

The individual data points of the Passive are fairly well separated, facet Pa, but certain individual data points introduce a degree of fluctuation even though as a type the Passive is extremely well separated. The categorical labels mask this behavior of the data. In general, there appears to be a slight competition between the Experiencer (facet Exp) and Reflexive Engagement (facet R.E) types based on the estimated class probabilities across data points. In contrast, the misclassified instances display a clear departure from the Passive. These are visible as troughs in the Pa facet, indicating that the model is unable to recover them based on the input. Diffusive verbs appear to be problematic. The whole semantics of the sentence appears to be more important rather than a specific pattern, as in 5.2-2.

5.2-2 *В основн-ом кадр-ы фильм-а*
PR essential-PREP scene-NOM.PL film-GEN
сопровожда-л-и-сь инструментальн-ой музык-ой
accompany-PST-PL-RM instumental-INS music-INS
Essentially, the scenes of the film were accompanied by instrumental music.
[540, RNC, Александр Смотров.]

Depending on the exact profile, the verb can be used to profile either the Motion Construction (Mo) or the Passive. Example 5.2-2 estimated the Motion with a probability of 0.664 and the Passive with the probability of 0.207. The data point is clearly visible in the estimated class probability plot, the crest in facet Mo.

As a construction type, the Passive appears to be well-separated from the other possible constructions in the data and the structural properties of the verbs capture the essential structure. Additionally, there appears to be very little competition between the Passive and other types available in the data in terms of argument constructions.

5.3 Definitions of the Spontaneous Event Construction

The labels middle, medial, or medio-passive are used especially in the early studies of Indo-European languages and it is commonly used to refer to an inflectional category of verbs (Fagan, 1992; Sturtevant, 1931). These labels figure prominently in the literature of reflexive markers. Moreover, the label middle pertains to the category of voice subsuming a number of different construction types following the classical tripartite structure of active, middle, and passive voice (Kemmer, 1993; Manney, 2000). This tripartite structure is also present in the Russian linguistic tradition, as demonstrated by Shahmatov (Шахматов, 1925). In a more restricted sense, the label is used to cover a specific construction type, especially in the studies of English, exemplified with such instances as *the books sell well* (cf. Davidse & Heyvaert, 2007; Guhl, 2010). In the Russian tradition, similar patterns are commonly labeled as the Modal Quasi-passive. However, at least in their sample, Kalashnikova and Saj (Калашникова & Сай, 2006) demonstrate that the type is infrequent in Russian. Thus, these are considered as verb-specific constructions in this study.

When specific instantiations are concerned, a number of labels are used to refer to a specific type, such as Anticausative, originally proposed by Nedyalkov and Sil'nickij (Недялков & Сильницкий, 1969), Spontaneous Event (cf. Croft, 2001; Kemmer, 1993; Shibatani, 1998; Гаврилова, 1999; Зельдович, 2010;

Перцов, 2003), Decausative (Gerritsen, 1990; Israeli, 1997; Падучева, 2001), and Medial (Gerritsen, 1990; Israeli, 1997). The label Anticausative is sometimes used synonymously for Spontaneous Event. For example, Croft (2001:317) links them together explicitly. Similarly, Siewierska (1984:78) defines Anticausative as an instantiation type profiling an event which is brought about spontaneously. From this perspective, the label Anticausative can be excluded.

In Kemmer's (1993:142) classification, the Spontaneous Event receives the following definition: "A Common use of MM [Middle Marker] across languages is in situations which designate the change of state of an entity, but in which no Agent entity receives encoding." Additionally, Kemmer includes the semantics of the verb as part of the definition by introducing the semantic component of causation (Kemmer, 1993:145-146). These characterizations figure prominently in classifications on the Russian Reflexive Marker. The label Spontaneous Event is proposed in this study and its semantic basis is explicated relative to previous taxonomies. The label allows both the cross-paradigmatic relation between the reflexive and the neighbor verbs to be maintained without excluding reflexiva tantum verbs. Example 5.3-1 illustrates a typical instantiation with the verb *сформироваться* 'form' satisfying the basic definition in Kemmer's taxonomy, (i.e., change of state and a causative neighbor verb *сформировать* 'cause to form').

5.3-1 *Мы сегодня выходим на какой-то путь цивилизационного развития*
[…]
рынка информационных технологий
[…]

когда ужес	*формирова-л-и-сь*	*достаточн-о*
when already	form-PST-PL-RM	sufficient-ADV
известн-ые	*крупн-ые*	*компани-и* […].
known-NOM.PL	large-NOM.PL	company-NOM.PL

We now take a journey on the path of civilized development of the IT market when sufficiently well-known large companies have already been formed.
[1081, RNC, Круглый стол "Взаимодействие бизнеса и государства в ходе реализации проекта "Электронная Россия" (2003)]

Thus, the question becomes how exactly the semantic range of the Spontaneous Event is delimited. Israeli (1997) forms four distinctive groups: Actional Decausative, Emotional Decausative, Medial Decausative, and Medial Proper. However, the definition of the decausative verbs given by Israeli is not compatible with most studies where Decausative is explicitly established relative to causative non-reflexive verbs (Gerritsen, 1990; Князев, 2007; Падучева, 2001).

Israeli (1997:66) states, "However, such reasoning implies that for the all decausative constructions there should be a parallel non-sja causative, which is not the case with most actional decausatives due to the animacy of the subject." This same category appears in Gerritsen's (1990:63-64) classification. Israeli

(1997:65) also considers that the label Actional Decausative corresponds to the translational motion verbs in Kemmer's (1993) classification. Thus, the decausative reflexive verbs are defined relative to a narrow set of semantic classes of verbs, which are motion and emotion.

Consequently, the classification is left with the concept of Medial. Israeli (1997:66) defines the Medial Decausative in terms of negation. "This group includes decausative *-sja* verbs which are neither actional nor emotional. […]. They usually involve inanimate subjects, and the action is presented as if taking place by itself." The last group, Medial Proper, is truly a residual one and only two examples consist of the verbs *содержаться* 'contain' and *иметься* 'exist.' These verbs are also grouped together by Gerritsen (1990:37) under the label Medial. In principle, Israeli's classification allows the inclusion of the reflexiva tantum verbs without arbitrarily excluding them, but the cross-paradigmatic relation of causative ~ decausative is lost. In contrast, Gerritsen makes an explicit distinction between Medial and Decausative based on the relation between the reflexive verb and non-reflexive verb.

In Medial, the reflexive verb depicts an autonomous event and has no corresponding neighbor verb, while in Decausative, the reflexive verb depicts the event from an opposite direction compared to the neighbor verb typically yielding an autonomous event type (Gerritsen, 1990:49). Effectively, we have two types of items that display similar behavior based on surface structure, but the demarcation between them is defined relative to the semantic cross-paradigmatic relation. Another complication factor is that the reflexive verbs, which do not have a causative neighbor verb but still may be considered to profile a spontaneous event type, appear to be a residual class, as in 5.3-2 with *подвергнуться* 'undergo.'

5.3-2 […] *но не единственная причина,*
[…]
по которой молодая, динамичная и агрессивная
[…]
пивн-ая *отрасл-ь* *подверг-л-а-сь*
beer-NOM industry-NOM undergo-PST-F-RM
массированн-ой *атак-е.*
massive-DAT attack-DAT
But it is not the only reason to the young, dynamic and aggressive beer industry has undergone a massive attack.
[918, RNC, Евгений Толстых. Пивка для рывка // "Совершенно секретно," 2003.09.01]

Another complication is the role of verbs in this construction type. According to Geniušienė, the causative component cannot be included in the semantic structure of decausative reflexive verbs. The reasoning for this is the fact that the basic function of the decausative is to mark the removal of the causative component of the corresponding neighbor verb (Geniušienė, 1987:100). Contrary to this position, Paducheva explicitly considers that the causative component is present in the semantic structure of the verb. For Paducheva, the Decausative diathesis contains the slot of Backgrounded Causer that can also be overtly encoded with the prepositional phrase *от*$_{gen}$. Additionally, decausative verbs are connected to perfective aspect (Падучева, 2001:53, 62). Due to these reasons, Paducheva argues against the notion of the Spontaneous Event. If the Backgrounded Causer can be expressed, the spontaneity cannot constitute a sufficient criterion (Paducheva, 2003).

Example 5.3-3 demonstrates the canonical instantiation of the causative ~ decausative alternation with the verb *развалиться* 'fall apart' which are perfective aspect, causative and semantically similar neighbor verb *развалить* 'ruin, destroy,' and an overtly encoded Causer with *от*$_{gen}$.

5.3-3 […] *избушка эта была уже ветхая и*
[…]
однажды от дожд-я развали-л-а-сь.
eventually PR rain-GEN fall.apart-PST-F-RM
The hut was already on the brink of destruction and eventually it fell apart from the rain.
[498, RNC, Татьяна Рик. Про вредную Бабку-Ёжку // "Мурзилка," №6," 2001]

However, the Causer can also be expressed with reflexive verbs that lack the cross-paradigmatic relation, illustrated with *раскраснеться* 'flush' in 5.3-4.

5.3-4 *Лиц-о от напряжени-я раскрасне-л-о-сь.*
Face-NOM PR tension-GEN flush-PST-N-RM
The face flushed from tension.
[1891, RNC, Андрей Геласимов. Фокс Малдер похож на свинью (2001)]

Thus, it appears that the *от*$_{gen}$ is not a distinctive property of the decausative reflexive verbs in Russian, similar to spontaneity.

Instead, the separate pathways gravitate towards similar structures. Additionally, other prepositional phrases can be used to indicate what might be defined as Causer, for instance in 5.3-5 with the complex preposition *в результате* 'as a product of' (cf. Храковский, 1991:164).

5.3-5 *В результат-е сложн-ых физическ-их*
PR product-PREP complex-GEN.PL physical-GEN.PL
процесс-ов образу-ет-ся неоднородн-ое по
process-GEN.PL form-3S.PRS-RM heterogenous-NOM PR
состав-у облачк-о смес-и.
composition-DAT cloud-NOM mixture-GEN
A heterogenous mixture cloud is formed as a product of complex physical processes.
[341, RNC, Ликбез: Что такое непосредственный впрыск бензина // "Автопилот," 2002.02.15]

The semantics of spontaneity can be overtly encoded with the so-called semi-predicates, suc as, *сам* '-self' and *сам собой* 'by itself (cf. Гаврилова, 1999:173; Падучева, 2001:63), that highlight the construal of the event as if taking place by itself. Consequently, the use of semi-predicates functions also as a test for the Passive Construction. Only the Spontaneous Event type is compatible with the semi-predicates. These contradictory criteria are, however, only a serious concern for a theoretical account if one assumes that any given category in a language stems from a single invariant component. In contrast, if a category is assumed to consist of clusters of properties, competing motivations are the norm rather than the exception, as in 5.3-6.[102]

5.3-6 *Как будто небо такое в облаках, и неизвестно,*
[…]
когда распогод-ит-ся.
when clear-3S.FUT-RM
As if the sky is in the clouds and it is unkown when it will become clear.
[1854, RNC, Андрей Геласимов. Ты можешь (2001)]

This is the primary motivation for merging the medial and decausative verbs in this study. Causativity is strictly a cross-paradigmatic relation between the reflexive and the neighbor verb, maintaining the insight acquired in the diathesis tradition that focuses primarily on alternations. At the same time, additional semantic information is required if the clausal meaning is sought. Paducheva (Падучева, 2001) establishes this by introducing semantic verb classes as components leading to partitioning of the reflexive verbs into a number of semantically coherent categories. This operation creates, nonetheless, an interesting issue, namely the possible motivation for the observed clustering of the Decausative around a specific set of the semantic classes of the reflexive verbs. Certainly a rule-based system does not imply any form of semantic clustering.

Nonetheless, another complication is related to the semantic proximity between the Passive and Spontaneous Event. Certain verbs can be considered to be diffusive between these two types, for example, verbs *открываться* 'open and *закрываться* 'close.' This is expected because the preposition *от*$_{\text{gen}}$ is diachronically used to mark the Agent in the Passive Construction and the use of the instrumental case is a later phenomenon (Meyer, 2010:292; Зарицкий, 1961:107-108). Gavrilova (Гаврилова, 1999:159-160, 172) attributes the difference between these two types to the construal of the event (cf. Падучева, 2001:52), illustrated in 5.3-7 with a common expression in metro.

5.3-7 *Осторожн-о, двер-и закрыва-ют-ся.*
Careful-ADV door-NOM.PL close-3P.PRS-RM
Careful, the doors are closing.

[102] Gerritsen (1990:50-58) separates a small category of verbs from the actional decausative verbs if the verb profiles an event taking place in nature.

[(Гаврилова, 1999:160)]

If the driver is understood as intentionally closing the doors, the Passive Construction is profiled. The opposite holds for the Spontaneous Event where the closing of the doors lies in the focus. The latter is certainly the most natural interpretation. Due to this proximity, Paducheva only discusses perfective verbs and deems the Passive Construction ungrammatical with the perfective aspect, preserving the Decausative alternation (Paducheva, 2003; Падучева, 2001).

Related to the construal of the event, Example 5.3-8 illustrates that the exact lexical profile and context are the determining factor between these two types.

5.3-8 *Как говори-л-и, Петербург строи-л-и, а*
how say-PST-PL NAME.ACC build-PST-PL but
Москв-а строи-л-а-сь.
NAME-NOM build-PST-F-RM
As it is said, that Saint Petersburg was built but Moscow grew [lit. built].
(Высоцкий, Gerritsen, 1990:31)][103]

In 5.3-8, the word game along with the combination of the construction types illustrates, *строили* ~ *строилась*, that the interpretation of the Spontaneous Event profiled with the verb *строиться* 'build' arises through the construal of the event rather than through the inherent property of the lexical item.[104]

Another complication of the semantic proximity between the Spontaneous Event and Passive Construction arises with the encoding type of the secondary slot with the usage of *у*$_{\text{gen}}$ prepositional phrase. Typically, the prepositional phrase is used to profile the Possessor. Recently, Kolomackij (Коломацкий, 2009:42-43) includes these patterns under the Passive Construction. However, it seems that the agentivity of the *у*$_{\text{gen}}$ type is rather weak, demonstrated in 5.3-9 and 5.3-10 (cf. Храковский, 1991:164-165).

5.3-9 *У нас провод-ят-ся конференци-и*
PR we.GEN carry.out-3P.PRS-RM conference-NOM.PL
по народн-ому творчеств-у.
PR folk-DAT art-DAT
Conferences on folk art are being carried out here.
[657, RNC, Народный костюм: архаика или современность? //
"Народное творчество", 2004.02.16]

[103] The glosses and translation were added by the author.

[104] In terms of derivational accounts, the form *строили* is commonly referred to as the Unspecified Subject Deletion marked with the plural personal form (Плунгян, 2000:200).

5.3-10 [*ч*]*то у нас всяк-ие их*
that PR we.GEN all.sort-NOM.PL their
товар-ы прода-ют-ся.
goods-NOM.PL sell-3P.PRS-RM
that we have all sorts of goods from them for sale.
[1390, RNC, Беседа с социологом на общественно-политические темы, Санкт Петербург // ФОМ (2003.09.09)]

A clear deviation from agentivity is given in 5.3-11.

5.3-11 *Да и откуда возьм-ёт-ся твёрдост-ь*
Yes and wherefrom take-3S.FUT-RM confidence-NOM
у человек-а,
PR person-GEN
который потерял ориентировку и не знает, куда податься?
[…]
Well how is a person, who has lost his way and does not know where to go, regain his confidence?
[1774, RNC, Василь Быков. Болото (2001)]

Wiemer's (2004:308-310) diachronic study of the Russian periphrastic passive shows that unambiguous agentive uses are rare with the prepositional phrase *у*$_{\text{gen}}$. In relation to synchronic data, this also seems to be evident as well as in 5.3-12.

5.3-12 *Только у него глаз-а совсем не открыва-л-и-сь.*
Only PR his eye-NOM.PL quite NEG open-PST-PL-RM
Only his eyes were not quite opened.
[1809, RNC, Андрей Геласимов. Жанна (2001)]

In the previous examples, the pattern *у*$_{\text{gen}}$ can still be perceived as profiling the Possessor, contrasting 5.3-13 where this interpretation, a possessive relation between the Possessor (*у меня*) and the Possessee (*дверь*) is doubtful. According to Paducheva (2003:188), the *у*$_{\text{gen}}$ does not function as Causer with the Decausative, but further explications are not provided. Nonetheless, Paducheva's (2003:185) observation that the inclusion of the perfective aspect may induce modal interpretation appears to fit Example 5.3-13.

5.3-13 *Я спрашиваю / «Кто там? ».*
[…]
А у меня не откры-л-а-сь двер-ь.
And PR I.GEN NEG open-PST-F-RM door-NOM
I ask:"Who is there?". And I did not manage to open the door.
[1397, RNC, Праздный разговор молодых людей, Московская область //]

Knyazev considers similar patterns as the expression of agentivity (Князев, 2007:295 footnote 1). For instance, the verb *получиться* ‘happen, become’ can be combined with the y_{gen} pattern, but it can also appear with an animate nominative type in the subject slot, as in 5.3-14. The reflexive verb is also perceived as dissimilar to its neighbor *получить* ‘receive, get.’[105]

5.3-14 *О, смотрика, смотрика,*
[…]
ты хорош-о получи-л-ся!
you.NOM good-ADV do-PST.M-RM
O look, look, you did well!
[1566, RNC, Евгений Гришковец. ОдноврЕмЕнно (2004)]

These examples illustrate that the y_{gen} pattern appears to form its own niche cutting across several types. However, the pattern is understudied in the Russian tradition, as well as in corpus-based studies, such as Kolomackij (Коломацкий, 2009). A dedicated study would be required to map the function of the y_{gen} pattern across argument constructions in Russian. Regardless of this, the y_{gen} pattern crosses different construction types and functions as an attractor, bringing them in closer proximity.

In sum, this section introduced the Spontaneous Event Construction connecting it through usage patterns, ranging from decausative verbs to reflexive verbs that lack the cross-paradigmatic relation in contemporary Russian.

5.4 Spontaneous Event Construction

The Spontaneous Event Construction covers 329 instances and contains 190 unique reflexive verbs in the RF model. The construction type has a lexical center and covers relatively frequent verbs that are perceived as semantically similar to their neighbor verbs such as *взяться* ‘set about, come from’ ~ *взять* ‘take,’ *браться* ‘set about, come from’ ~ *брать* ‘take’, *создаться* ‘arise, form’ ~ *создать* ‘create, produce,’ and *открыться* ‘open’ and *открыть* ‘open.’ Whereas another set of relatively frequent verbs are perceived as dissimilar, for instance, *получиться* ‘happen, become’ ~ *получить* ‘receive, get’ or as intermediate, such as *получаться* ‘result, turn out.’ Interestingly, the latter two verbs are aspectual pairs, but apparently the perfective verb form, *получиться*, has become increasingly merged with the reflexive paradigm to the point that the semantic connectivity has been lost to its neighbor verb, contrasting the imperfective verb form which seems to display variation. Additionally, the construction type subsumes the verb-specific constructions, which are connected through the causative semantic component. The definition of the Spontaneous Event follows.

[105] Example 5.3-14 could also be classified as an instantiation of the Modal Quasi-Passive, but it would not, once again, satisfy the derivational rule as the neighbor verb is perceived semantically different.

Function: Profiles an entity undergoing an action.

Form: Nominative subject and a construal-specific secondary slot.

The Spontaneous Event follows the paradigm of the personal construction types. A distinction is made with this type, and the encoding of the secondary slot is considered to be construal-specific, dependent on the mode of profiling the event rather than being an inherent property of the verb-specific construction (cf. Croft, 2001; Langacker, 1988a). Thus, the encoding of the secondary slot forms a continuum ranging from traditional adjuncts such as expressions of location and time, to more verb-specific constructions such as the *у*$_{\text{gen}}$ and *от*$_{\text{gen}}$ patterns. Example 5.4-1 illustrates the Spontaneous Event Construction and the encoding is given in Figure 5.4-1.

5.4-1 *Однако между этими этапами, равно как и внутри них,*

[…]

движени-е	*развива-л-о-сь*	*к*	*райн-е*
movement-NOM	evolve-PST-N-RM	PR	utter-ADV
неравномерн-о.			
irregular-ADV			

However, between these stages, also equal to within them, the movement evolved utterly irregularly.

[223, RNC, [Александр Кацва. Россия 1990-х: Протестное движение //"Отечественные записки," 2003]

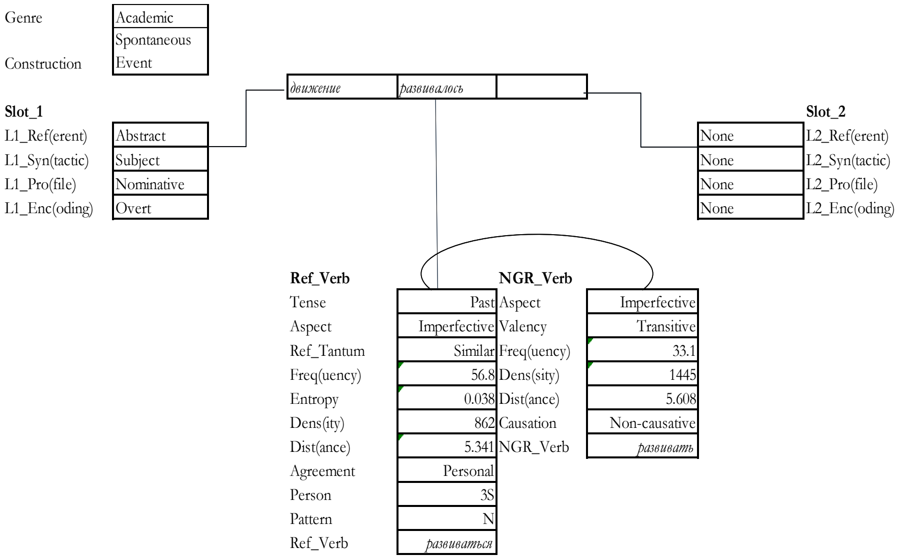


Figure 5.4-1 Layered structure of the canonical Spontaneous Event Construction.

The confusion matrix is given in Table 5.4-1 based on the estimated classes by the RF model. Considering the variation with the secondary slot, the class-wise error, 0.164, is relatively small. Additionally, Recall with the value of 0.836

points to the conclusion that the model is able to identify the construction type when all the types are contrasted globally. However, precision, 0.768, is lower, indicating that the model has difficulties in classifying the instantiations compared to identifying them.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	**1**	**0**	**4**	**0**	**0**	**20**	**2**	**0**	**0**	**3**	**23**	**275**	**1**
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 5.4-1 Confusion matrix of the predicted constructions. The Spontaneous Event Construction is given in bold.

The fluctuation or competition between the different construction types is visible in Figure 5.4-2.

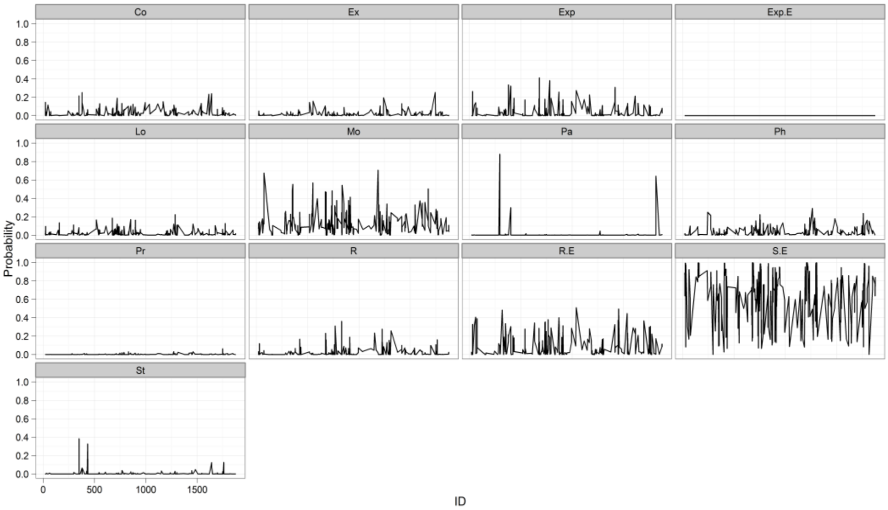


Figure 5.4-2 Faceted estimated class probability plot of the Spontaneous Event Construction (S.E). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The results indicate that certain typical instantiations of reflexive verbs containing the semantics of spontaneity are classified correctly, such as *родиться* 'be born,' *затеяться* 'emerge,' and *измениться* 'alter, change.' Another subtype is formed by deadjectival verbs such as *сузиться* 'narrow,' *увеличиваться* 'grow,' *ухудшиться* 'decay,' *улучшаться* 'improve,' and *уменьшиться* 'diminish.' From a verb-specific perspective, the Spontaneous Event competes, nonetheless, with most verbs, yielding a bleaching effect. Traditionally, the Spontaneous Event is considered to be an intermediate type that connects other realizations of the Reflexive Marker (cf. Kemmer, 1993; Manney, 2000). From this perspective, fluctuation is to be expected and is visualized in Figure 5.4-2. Example 5.4-2 is a prime case of the diffusive nature of the Russian Reflexive Marker with the verb *попасться* 'come up against.'

5.4-2 *Я хочу сказать / что в предстоящем сезоне так же есть еще два*
[…]
резервных *варианта / это бело-красно-зеленый /*
[…]
и *если* *уж* *нам* *попад-ут-ся* *соперник-и*
and if well we.DAT come.up-3P.FUT-RM rival-NOM.PL
как *Монако* […].
like NAME.NOM
с которыми мы не можем играть в наших традиционных цветах / то на
[…]
этот случай у нас есть черный с красным и зеленый с желтым…
[…]
I want to say that in the upcoming season there are still two reserved

variants, white-red-green, and if the rivals, like Monaco, will come up against us, we cannot play in our traditional colors. In this case, we still have black with red and green with yellow.
[1017, RNC, Встреча футбольного клуба "Локомотив" с болельщиками, Москва (2004.02.21)]

The profile follows the Nominative-Dative pattern that merges with the various Experiencer types in Russian, cf. Section 6.1.[106] The estimated class probabilities reflect this (the Spontaneous Event with the probability of 0.07, the Experiencer Construction with the 0.19, and the Stimulus Construction with the 0.327). The instantiation is estimated as the latter type.

Thus, the core of the Spontaneous Event Construction emerges from the data, but the bleaching effect is present with certain instantiations as expected. Nonetheless, the proposed linguistic model appears to be able to capture the typical usage patterns.

5.5 Definitions of the Reflexive Engagement Construction

The label Reflexive Engagement is used in this study as a generalized argument construction type to cover a range of different types proposed in previous studies. In this sense, a motivation for postulating a new label is in order. The shared commonality with the instantiations is that the verb profiles an activity where the subject acts upon a secondary entity. In this regard, the instantiations are between the canonical two-place Transitive Construction and the canonical one-place Intransitive Construction.

Examples 5.5-1–5.5-4 illustrate the construction type with the verbs *касаться* 'touch,' *заниматься* 'do, engage in' and *делиться* 'give, share.' Importantly, the verb *касаться* does not have a neighbor verb and the verb *заниматься* is perceived as semantically intermediate relative to its neighbor *занимать* 'occupy, employ' contrasting *делиться* 'share' perceived as similar to *делить* 'share.'

5.5-1 *Прослед-и-те, чтобы верёвк-а не каса-л-а-сь*
Make.sure-IMP-2P that rope-NOM NEG touch-PST-F-RM
крыш-и.
roof-GEN
Make sure that the rope does not touch the roof.
[289, RNC, Навыки: Как правильно перевозить груз на крыше // "Автопилот", 2002.04.15]

5.5-2 *Наш-а организаци-я занима-ет-ся исследовани-ем*
Our-NOM organization-NOM do-3S.PRS-RM research-INS
общественн-ого мнени-я,
public-GEN opinion-GEN
в том числе в сфере политики.

[106] This particular type comes in proximity to a variation with the Reciprocal Construction that also follows the Nominative-Dative pattern (Israeli, 1997). It seems that the subtype is fairly infrequent, as it is not attested in the sample.

[…]
Our organization does research on public opinion, including in politics.
[1128, RNC, Беседа в Новосибирске (2000.08.15)]

5.5-3 *"Даш-а" дел-ит-ся секрет-ами:* […].
NAME-NOM give-3S.PRS-RM secret-INS.PL
Dasha magazine gives away secrets
[786, RNC, "Даша" делится секретами - побалуй близких вкусным обедом! // "Даша," №10," 2004]

Perhaps the most prominent subtype used in the previous studies is the Benefactive or Indirect Reflexive (cf. Kemmer, 1993; Виноградов, 1972), exemplified in 5.5-4 with the verb *строиться* 'build'.

5.5-4 *Буд-ете сам-и строи-ть-ся - учт-и-те.*
be-2P.FUT self-NOM.PL build-INF-RM - consider-IMP-2P
You will build your own house by yourselves - consider that.
[295, RNC, Мария Пупшева. На крыше дома твоего // "Вечерняя Москва," 2002.04.11]

This class is proposed for verbs denoting a situation type where the primary entity performs an activity for one's benefit. This semantic component is occasionally used to divide the Indirect Reflexive into more fine-grained classes, (i.e., the Benefactive). Example 5.5-4 illustrates the vagueness of reflexive verbs in general. The cues originating in the profile change the construction type. The animacy of the referent, although omitted in the example, excludes the Passive and Spontaneous Event types.

The referent of the subject argument is not being built by someone nor is the building happening to them spontaneously. The semantic structure of the example does not contain the coreference of Agent and Patient. The referents are not building themselves. Traditionally, the basic component used to establish this subtype is a paraphrase test with the reflexive pronoun *себе* 'for oneself,' or with other semantically similar prepositional phrases, such as *для себя* 'for oneself.' From a derivational perspective, this type could also be classified as pertaining to the category of Omitted Object (Сай, 2007). The object of the non-reflexive verb is omitted and incorporated into the semantic structure of the reflexive verb, (i.e., *дом* 'house').

Gerritsen (1990:80-87) makes a more fine-grained distinction based on these paraphrase tests, yielding a separation between benefactive and possessive reflexives.[107] The category Benefactive is also adapted by Israeli (1997), whereas Knyazev (Князев, 2007) uses the label Possessive.[108] This mode of analysis relies heavily on the exact profile of the instantiations and originates in the composition of the whole expression. This is present in Gerritsen's analysis with

[107] Interestingly, malefactive has not been proposed as a class, at least to my knowledge, although the benefactive analysis is akin to binary categories.

[108] Israeli also considers this type to be productive.

the benefactive reflexives given in 5.5-5:

5.5-5 […] *сколько* *нам* *понадоб-ит-ся* *хлеб-а* […].
how.much we.DAT necessary-3S.FUT-RM bread-GEN
How much bread do we need?
(Gerritsen, 1990:86)[109]

Although Gerritsen's study is among the few which also take into account for case patterns, Example 5.5-5 deviates in a number of ways from what one might expect a typical benefactive reflexive to be. First, an inanimate entity is the subject referent. The genitive case marking is due to *сколько* 'how much.' Second, the proposed benefactive is in the dative case (*нам*). Consequently, this pattern is analyzed as an instance of the Stimulus Construction, as its form aligns with those patterns as do its semantics in Section 6.4. This analysis does not deny the obvious interpretation that the bread is the beneficiary for the Experiencer, for instance, by preventing starvation. This mode of analysis, however, relates to the question on the exact range of information considered to be relevant for establishing abstractions over instances.

Furthermore, the Benefactive analysis does not distinguish between semantically different notions, such as Recipient and Benefactive. Although both of these semantic types share a number of properties, a division is drawn between the modes of transfer in typical cases. Shibatani considers that the Recipient is characterized as being the target of a direct transfer and the Recipient of an actual transfer of some entity. In contrast, the Benefactive is a target of indirect transfer, that is the intended target of some activity and not a target of some transferred entity (Shibatani, 1996). At the same time, the most prominent shared feature among these two notions is benefit. Both notions imply that the entity gains something. From a semantic perspective, Gerritsen's label Benefactive captures this semantic component.

Another typical verb which can be considered to contain the semantic component of Benefactive is *готовиться* 'prepare,' given in 5.5-6.

5.5-6 *В холодную зимнюю ночь увидел он на опушке леса большую толпу людей,*
[…]
поклонников Тора, которые стояли вокруг священного дуба — "дерева грозы"
[…]
и *готови-л-и-сь* *к* *человеческ-ому*
and prepair-PST-PL-RM PR human-DAT
жертвоприношени-ю *в* *чест-ь* *"Смерт-и* *солнц-а."*
sacrifice-DAT PR honor-ACC NAMR-GEN NAME-GEN
In the cold winter night, he saw a large group of people, worshipers of Thor, at the edge of the forest. They stood around the holy oak, the Thunder Oak, and people prepared for human sacrifice in honor of the "Death of the Sun."

[109] The translation and glossing were added by the author.

[93, RNC, Н. Ю. Феоктистова. Новогородняя ёлка // "Первое сентября," 2003]

Certainly against a specific culturally motivated background, the human sacrifice can be interpreted as beneficial activity. Nonetheless, Example 5.5-6 may also be interpreted as consequential reflexive proposed by Gerritsen. For Gerritsen, the essential semantic component of the consequential reflexive verb lies in the manner the activity and is depicted by the verb bearing consequences to the Agent (Gerritsen, 1990:88).

However, there is a fine line between the benefactive and consequential interpretations. The consequential might be involved in such verbs as *держаться* 'hold' and *добиться* 'reach, attain, achieve,' as in 5.5-7 and 5.5-8.

5.5-7 *Они* *держа-л-и-сь* *из последн-их* *сил* [...].
They.NOM keep-PST-PL-RM PR last-GEN.PL strength.GEN.PL
They hanged on with their last strength.
[1582, RNC, Александр Дорофеев. Эле-Фантик // "Мурзилка," №1 5,"2003]

5.5-8 *Правительств-о* *доби-л-о-сь* *того,*
Goverment-NOM achieve-PST-N-RM that.GEN
что оборот остался на крайне ограниченном уровне.
[...]
The goverment achieved to keep the turnover at a very limited level.
[829, RNC, Денис Викторов. Стена // "Бизнес-журнал," 2003.10.23]

The verb *цепляться* 'cling' serves to demonstrate in 5.5-9 that the interpretation between the benefactive or consequential changes depending on the adapted perspective. Similar observation holds for the reflexive verb *ориентироваться* 'orientate' in 5.5-10.

5.5-9 *Я только одного не пойму никак.*
[...]
[*ч*]*его* *к* *нему* *баб-ы* *так цепля-ю-тся?*
What.GEN PR he.DAT lady-NOM.PL so cling-3P.PRS-RM
There is only one thing that I cannot understand in any way. Why the ladies cling on him so much?
[1228, RNC, Обсуждение реалити-шоу "Дом-2" (2006.04)]

5.5-10 *Конечно, пар-ы, котор-ые хорош-о*
Of.course couple-NOM.PL who-NOM.PL good-ADV
ориентиру-ют-ся в сексуальн-ых реакци-ях
orientate-3P.PRS-RM PR sexual-PREP.PL reaction-PREP.PL
партнёр-а и сво-их собственн-ых,
partner-GEN and own-GEN.PL own-GEN.PL
могут достичь кульминации одновременно, но не каждый раз.
[…]
Of course, couples, who are able to orientate well in the sexual reactions of the partner and in their own, can reach culmination simultaneously but not every time.
[714, RNC, Вероника Стрельникова. Опять акробатика, милый? // “Даша,” №10," 2004]

In my view, the Reflexive Engagement Construction captures the essential semantics in 5.5-9 and 5.5-10, a directed activity towards a secondary entity.

Additionally, the benefactive and the consequential semantics are closer to semantic component of manner. Thus, the proposed analysis follows the configuration of who-did-what-to-whom rather than the manner of did-how. The verb *поддаться* ‘give away, fall for’ is another example, as in 5.5-11. In this regard, a stark contrast to previous studies is the analysis proposed by Geniušienė. She considers similar verbs as agentive labeled as the Agentive Reflexive (Geniušienė, 1987:78).

5.5-11 *Георги-й Иван-ов тоже подда-л-ся*
NAME-NOM NAME-NOM also fall.for-PST.M-RM
этому очаровани-ю.
this.DAT spell-DAT
Georgij Ivanov also fell for this spell.
[202, RNC, Вадим Крейд. Георгий Иванов в Йере // “Звезда," №6,” 2003]

Thus, the consequential semantics along with the agentive verbs proposed by Geniušienė are analyzed in this study as part of the verb-construction analogously to the Benefactive.

In addition to the categories proposed in the previous studies, the Reflexive Engagement can be used to motivate a change in the profile with verbs typically pertaining to the category of the Reciprocal Construction, for example *бороться* 'fight,' as illustrated in 5.5-12.

5.5-12 *Предновогоднее выступление команды станет своеобразной*
[…]
репетицией перед Сочинским фестивалем
[…]

где наш-и земляк-и буд-ут
where our-NOM.PL countryman-NOM.PL be-3P.FUT
боро-ть-ся за выход в высш-ую лиг-у.
fight-INF-RM PR entry.ACC PR high.SUP-ACC league-ACC
The performance of the team before New Year's Eve becomes a distinctive rehearsal before the Sochi festival where our countrymen will fight for the entry to the top league.
[520, RNC, Юбилейный концерт будет благотворительным // Московский комсомолец в Саранске, 2004.12.23]

In this configuration, the verb *бороться* 'fight' does not retain its typical reciprocal structure in terms of co-reference between multiple Agent and Patient roles (Kemmer, 1993), cf. Section 5.9. Additionally, the secondary entity is an abstract Goal rather than Patient. Thus, the semantic structure resembles the consequential analysis proposed by Gerritsen (1990). Within the Reflexive Engagement Construction, the verb *приняться* 'proceed' displays a similar pattern to the extensions of the Reciprocal Construction with the *за*$_{\text{acc}}$ pattern as in 5.5-13.

5.5-13 *Потом приня-л-и-сь за постройк-у дом-а*
Then proceed-PST-PL-RM PR building-ACC house-GEN
Then they proceeded to build the house.
[814, RNC, Красна изба углами // "Народное творчество," 2003.12.22]

Another small cluster of verbs resembles the consequential reflexive verbs given in 5.5-14–5.5-17.

5.5-14 *Миш-а над вами издева-л-ся.*
NAME-NOM PR you.INS.PL bully-PST.M-RM
Misha bullied you.
[1170, RNC, Разговор в офисе страховой компании (2006.11)]

Similar analysis can be used to motivate the extensions of the Experiencer Construction, cf. Sections 6.1 and 6.2. The verb *смеяться* 'laugh' is a typical example of undergoing some mental event (Янко-Триницкая, 1962). A change in the profile establishes a different sense 'mock,' as in 5.5-15 with the *над*$_{\text{ins}}$ pattern.

5.5-15 *А потом бил мальчишек,*
[..]
котор-ые над ним смея-л-и-сь.
who-NOM.PL PR he.INS laugh-PST-P-RM
And then he beat the boys who mocked him.
[585, RNC, Андрей Геласимов. Жанна (2001)]

The verb *смеяться* 'laugh' and *издеваться* 'bully' share the same morphological marking *над*$_{\text{ins}}$ of the secondary slot. The dative case is considered archaic with *смеяться* (Кузнецов, 2009 [1998]), although it is still used with *усмехнуться* 'smirk.' The verbs are related through the shared root form, -*сме*-.[110]

In addition to the previously mentioned family of types, this generalized argument construction is also used to cover instantiations, such as 5.5-16. This particular pattern is typically included in taxonomies under different labels. Examples include Antipassive (Guhl, 2010), Absolutive in the diathesis tradition (Князев, 2007), Potential Active in Gerritsen (1990), and Aggressive in Israeli (1997).

5.5-16 [*я*] *не* *куса-ю-сь.*
I.NOM NEG bite-1S.PRS-RM
I do not bite.
[1325, RNC, Лекция по культурологии (2006.04)]

The label Potential Active captures the semantics of the instantiation in a sense that the reflexive verb in 5.5-16 does not profile a single action carried out by the subject. It does not profile the Semantic Reflexive Construction either, where the action is performed upon oneself, (i.e., to bite oneself). Instead, the meaning is closer to a habitual sense.

However, *кусаться* 'bite' can also combine with the *на*$_{\text{acc}}$ pattern; unfortunately, the pattern is not attested in the sample. This is simply due to the infrequency of the reflexive verb in general. Nonetheless, this directed activity with the *на*$_{\text{acc}}$ pattern is attested with other reflexive verbs, forming yet another small verb-specific cluster, as in 5.5-17 and 5.5-18.

5.5-17 *Философ* (*замахива-ет-ся* *на* него)
Philosopher.NOM threaten-3S.PRS-RM PR he.ACC
The philosopher threatens him
[1532, RNC, Ordinamenti // "Экран и сцена," 2004.05.06]

5.5-18 *То,* *что* *ты* *не* *веша-ешь-ся* *мне*
That that you.NOM NEG hang-2S.PRS-RM I.DAT
на *ше-ю* *с* *требовани-ем* *жени-ть-ся.*
PR neck-ACC PR demand-INS marry-INF-RM
The fact that you do not cling to my neck with demands of marriage.
[685, RNC, Ольга Зуева. Скажи что я тебе нужна… // "Даша," №10," 2004]

In this verb-specific construction, *замахиваться* 'threaten' and *вешаться* 'hang' might fit the Agentive Reflexive proposed by Geniušienė (1987:78).

Additionally, the Reflexive Engagement Construction can be used to motivate such instantiations like in 5.5-19 with *жениться* 'marry' formed with the

[110] From a morphological perspective, the verb *смеяться* is simpler compared to *усмехнуться*. Interestingly, the simpler form has undergone leveling, whereas the complex form retains the older pattern.

на$_{\text{prep}}$ pattern.

5.5-19 *Робин* *жени-л-ся* *на* *принцесс-е.*
NAME.NOM marry-PST.M-RM PR princess-PREP
Robin married the princess.
[1737, Сергей Седов. Доброе сердце Робина // "Мурзилка," №7," 2002]

In sum, the label Reflexive Engagement is proposed as a unifying abstraction subsuming a number of different semantically- and pattern-driven types. The unification does not imply that the more fine-grained level is not available. On the contrary, the abstraction is meant to capture a generalization, functioning as the gravitational center for more specific subtypes, (i.e., verb-specific constructions). The label Reflexive Engagement covers such traditional categories as Benefactive, Indirect Reflexive, Possessive and Consequential (Побочно-возвратные), (Gerritsen, 1990; Israeli, 1997; Виноградов, 1972; Князев, 2007). Thus, verb-specific constructions can be used, if deemed necessary, to capture the traditional labels and the proposed argument construction holds a generalization over them as a schematic type. The final motivation for proposing a new label is that it enables, once again, to bring together verbs that do not form the cross-paradigmatic relation.

5.6 Reflexive Engagement Construction

The Reflexive Engagement Construction covers 215 instances in the RF model. Additionally, the type contains 126 unique reflexive verbs. The lexical center of the Reflexive Engagement is supported with the following frequent verbs: *заниматься* 'do, engage,' *касаться* 'touch,' and *смеяться* 'mock.' The verbs in question do not have neighbor verbs in Russian. Another set of frequent verbs supporting this type is attested with *встретиться* 'meet,' *учиться* 'study,' *пользоваться* 'use,' and *держаться* 'hold.' These verbs have neighbor verbs which are perceived as semantically similar. In this sense, the construction is supported by both cross-paradigmatic and paradigmatic reflexive verbs. The definition of the Reflexive Engagement follows.

Function: Profiles an entity engaged in action towards a second entity.
Form: Nominative subject and verb-specific constructions for the secondary slot.

The Reflexive Engagement Construction has a weak alignment in terms of the form pole and it is primarily manifested in small clusters of verb-specific constructions such as the *на*$_{\text{acc}}$, *за*$_{\text{acc}}$ and *над*$_{\text{ins}}$ patterns. Importantly, these patterns can be viewed as anchors for extension types ranging from the Reciprocal Construction to Experiencer. The definition also implies that typically the subject is either a person or, at least, animate, and the initiator of the event (cf. Geniušienė, 1987:78). However, abstract entities can be construed as initiators through metaphorization, illustrated as in 5.6-1.

5.6-1 *Теори-и* *бор-ют-ся* *за* *“захват”* *и*
Theory-NOM.PL fight-3P.PRS-RM PR capture.ACC and
“охват” *сам-ыми* *разн-ыми* *метод-ами.*
coverage.ACC very-INS.PL different-INS.PL method-INS.PL
Theories are fighting for the capture and coverage with a variaty of different methods.
[Александр Ослон. Мир теорий в эпоху "охвата" // "Отечественные аписки," 2003]

The layered structure is given in Figure 5.6-1 and illustrated with Example 5.6-2.

5.6-2 *Власт-ь* *“дел-ом”* *занима-ет-ся.*
Regime-NOM matter-INS attend-3S.PRS-RM
The regime is atteding the matter.
[959, RNC, Иосиф Гальперин. Власть "делом" занимается // “Совершенно секретно”, 2003.08.09]

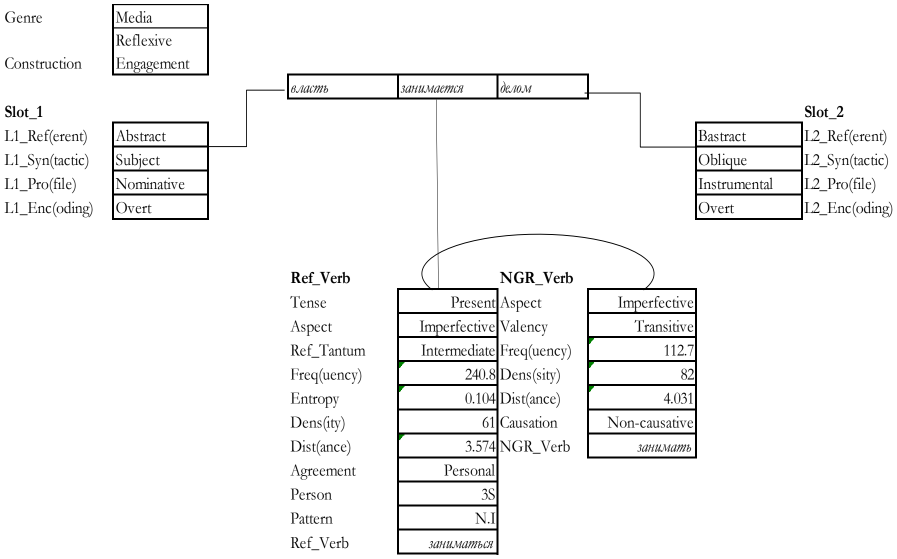


Figure 5.6-1 Layered structure of the canonical Reflexive Engamgent Construction.

The confusion matrix is given in Table 5.6-1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	**4**	**0**	**9**	**0**	**0**	**28**	**4**	**1**	**0**	**4**	**145**	**20**	**0**
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 5.6-1 Confusion matrix of the predicted constructions. The Reflexive Engagement Construction is given in bold.

The class-wise error, 0.326, is fairly high and belongs to the five construction types that display a high degree fluctuation. The recall value of 0.674 indicates that the model is able, up to a certain extent, indentify the type globally. But with the precision of 0.621, the model has difficulties in classifying them. The competition between the instantiations is visualized in Figure 5.6-2.

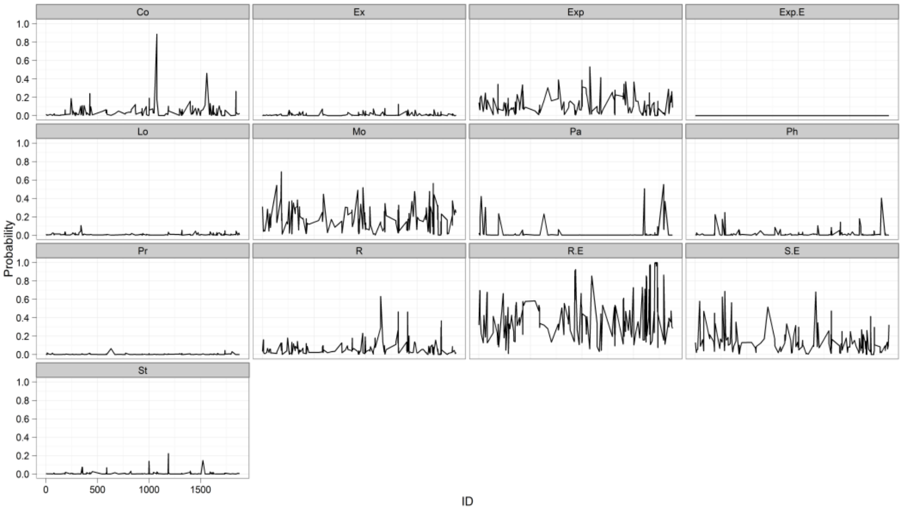


Figure 5.6-2 Faceted estimated class probability plot of the Reflexive Engagement Construction (R.E). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The competition is centered on three Construction types: the Experiencer, Translational Motion, and Spontaneous Event. The strongest fluctuation appears to be with the Experiencer Construction. For example, there are four instantiations with the verbs *бороться* 'fight' that pertain to the category of the Reflexive Engagement. Two of them are estimated to be instances of the Experiencer Construction, cf. Example 5.6-1. A possible motivation behind this competition is the fact that both of the misclassified instances have an abstract secondary entity aligning more strongly with the Experiencer Construction. Example 5.6-3 illustrates the issue with the verb *бояться* 'be afraid.'

5.6-3	*Я*	*бо-ю-сь*	*за*	*сво-е*	*украшени-е.*
	I.NOM	be.afraid-1S.PRS-RM	PR	own-ACC	jewelry-ACC

I am afraid for my jewelry.
[1241, RNC; Передача "С утра пораньше" на телеканале "Домашний" (2006.04)]

The structural properties of these verbs align in a specific configuration with the *за*$_{acc}$ pattern. From a global perspective, the results suggest that the Reflexive Engagement Constructions is primarily a locally driven type and the cue validity of the structural properties of the verbs are lower, amounting to stronger competition between different instantiations, specifically between the Experiencer and the Motion Constructions. Thus, the higher competition leads to a situation where semantics of the verbs play an important role for these three types rather than the structural properties.

5.7 Definitions of the Phase Construction

As a general type, the Phase Construction is well-separated in the Russian verbal system. The basic semantic property of this type is the modification of the event structure by introducing a temporal segmentation(s) (Падучева, 2004; Храковский, 1987). The following non-reflexive verbs constitute the canonical phasal verbs based on Paducheva (Падучева, 2004:180): *начать* 'begin,' *продолжать* 'continue,' and *закончить* 'end.' Another set of labels applied to this category are inchoative and ingressive. These labels have their roots in the aspectual tradition in linguistics. A third and rarely used label is inceptive (Flank, 1987). The interconnection of these categories in relation to aspect is discussed in detail, for example, by Hrakovskij (Храковский, 1987:188-195). If the phase is taken as a semantic category in its own right, generally speaking, Russian has invested extensively in profiling this domain. This is done by covering it, not only with verbs, but also with prefixes, such as *за-* and *по-* (cf. Janda, 1986; Храковский, 1987:157-162).[111]

In this respect, the notion of phase is used in a narrow sense in this study

[111] There is an interesting connection between these prefixes. They typically exclude each other when they profile a beginning point of a directed motion, (i.e., a verb can only be conjoined with one of them). Hrakovskij gives few exceptions to this. However, an example is *подуть* and *задуть* 'start to blow.'

and it is applied to a construction type supported by a set of reflexive verbs. Additionally, only instances containing the temporal segmentation without prefixation are considered. Thus, such verbs as *забеспокоиться* 'start to worry, become worried' and *засмеяться* 'start to laugh' are not classified as instances of the Phase Construction type, as the temporal segmentation is introduced with the prefix *за*. At the same time, this only shows the interconnectedness of categories in Russian, ranging from prefixes to verbs. Recently, Langacker analyzes infinitive complements under the broad category of Phase. Such verbs as *know* and *believe* introduce a result phase, whereas *learn* and *calculate* yield an action phase when combined with an infinitive (Langacker, 2009:312-313). Langacker's concept of Phase allows the inclusion of the canonical phasal verbs with a larger group of reflexive verbs that combine with the Nominative Infinitive pattern. The latter group has not been systematically included in previous taxonomies. For example, Israeli (cf. 1997:64, 66-67) gives the verb *собираться* 'going to' and, presumably, it is classified as an instance of the Medial Decausative. This is not explicitly stated.

The temporal segmentation of an event is linked to the profile of the phase verbs through the Nominative Infinitive pattern. A comprehensive list of Russian verbs compatible with this pattern is given by Divjak (2004). From a temporal perspective, the semantics of phase contains either: a beginning, a continuation, or an end point. In this sense, the previously stated verbs lexically specify segmentation by imposing an interval upon the profiled event type (cf. Золотова, Г. А., 2005 [1973]:212-214). Kustova considers that the interval is lexically profiled and is a crucial component of the profile for the canonical phasal verbs. The extreme ends ranging from phasal to non-phasal can be illustrated with the verb *состояться* 'take place.' Kustova considers that with the later verb the whole event is profiled in it's totality compared to phasal verbs, such as *начаться* 'begin' (Кустова, Г. И. , 2002:73-74).

The phasal verbs also impose restrictions on the infinitive, namely that the infinitive can only be in the imperfective aspect (cf. Divjak, 2009 for discussion on aspectual properties of infinitives in Russian). Paducheva motivates this restriction in terms of aspectual semantics. The definition of a phase assumes that the event type profiled by the infinitive can be divided into temporally coherent parts. Traditional accounts on Russian aspect define the imperfective as a process, while the perfective depicts a single, unified action. Thus, the invariant basis of aspect is used to motivate this behavior of the canonical phasal verbs because a holistic action is indispensable into parts. At the same time, this is hardly a new observation, as it is stated by Paducheva (Падучева, 2004:179) citing the work of Peshkovskij (Пешковский, 1938). In contrast, the Nominative Infinitive pattern is not available for the canonical phasal reflexive verbs. Examples 5.7-1–5.7-3 illustrate the usage patterns.

5.7-1 *Итак, ссор-а* *нача-л-а-сь* *с* *разговор-а*
So quarrel-NOM begin-PST-F-RM PR discussion-GEN
о *семь-е.*
PR family-PREP
So the quarrel started from a discussion about the family.
[686, RNC, Ольга Зуева. Скажи что я тебе нужна… // "Даша", №10", 2004]

5.7-2 *Полёт* *их* *продолжа-л-ся,* […].
Flight.NOM their continue-PST.M-RM
Their flight continued.
[1789, RNC, Василь Быков. Болото (2001)]

5.7-3 *Нет,* - *говор-ит,* - *вс-я* *кончи-л-а-сь.*
No - say-3S.PRS - everything-NOM end-PST-F-RM
No, it is said that everything ended.
[1995, RNC. Олег Тихомиров. Про козла Тихомира // "Мурзилка," №2," 2001]

From a derivational perspective, this behavior can be partly motivated by appealing to the argument structure of the neighbor verb. It is commonly argued that the phasal verbs have semantically a one-place argument structure and syntactically a two-place one.

To support this claim in derivational models, impersonal verbs are used as a test to display this behavior (cf. Апресян, Ю. Д., 1980:26; Храковский, 1987:163). Hrakovskij shows this behavior with the impersonal verb *смеркаться* ‘to be dusk’ and the phasal verb *начать* ‘begin,’ given in 5.7-4 and 5.7-5.

5.7-4 *Смерка-ет-ся.*
to.be.dusk-3S.PRS-RM
It is getting to be dusk.

5.7-5 *Нача-л-о* *смерка-ть-ся*
begin-PST-N to.be.dusk-INF-RM
It is beginning to be dusk.

It is argued that a zero subject occupies the subject position when the phasal verb is conjoined with the impersonal verb, as in 5.7-5. (Храковский, 1987:163). From a constructionist perspective, the behavior of the neighbor verbs can be accounted for through verb-specific constructions; namely, they have multiple patterns available. Examples 5.7-6 and 5.7-7 were specifically extracted from the Russian National Corpus.

5.7-6 *Ванг,* *китайск-ий* *иммигрант,*
NAME.NOM Chinese-NOM immigrant.NOM
начина-л *сво-ю* *карьер-у* *довольн-о* *стандартн-о,* […].
begin-PST.M own-ACC career-ACC fair-ADV conventional-ADV
Wang, a Chinese immigrant, began his career fairly conventionally.
[Леонид Черняк. Три ошибки Доктора // «Computerworld», 2004]

5.7-7 *Ребёнок начина-ет капризнича-ть,* […].
Child.NOM begin-3S.PRS act.up-INF
The child begins to act up.
[О. Г. Баринов. Зоологический сад // «Первое сентября», 2003]

In contrast, the behavior of the canonical phasal reflexive verbs can be motivated by stating that they gravitate towards the patterns associated with processual semantics. Based on corpus-data, Divjak and Gries (2006) show that the phasal reflexive verbs profiling the beginning-type typically appear with a process or a temporal event occupying the subject position (cf. Падучева, 2004:182-187).

Paducheva (Падучева, 2004:183) formalizes the semantics as P (process) has its own initial phase Q (cf. Богуславский, 1998). These phases can be lexically profiled, as in 5.7-8 and 5.7-9 with the instrumental case.

5.7-8 *Длинн-ый коридор заканчива-ет-ся окн-ом,* […].
Long-NOM corridor.NOM end-3S.PRS-RM window-INS
The long corridor ends to the window.
[635, RNC, Иваново. Детство // "Экран и сцена," 2004.05.06]

5.7-9 *Там дел-о кончи-л-о-сь банан-ом.*
There situation-NOM end-PST-N-RM banana-INS
All of that ended with a banana.
[1202, RNC, Обсуждение компьютерной игры (2006.11)]

At the same time, the exact profile is a determining factor and the canonical phasal verbs can also appear in the Passive Construction as illustrated with the verb *прерываться* 'interrupt' in 5.7-10.

5.7-10 […], *как будто течени-е времени не прерыва-л-о-сь*
as if course-NOM time.GEN NEG interrupt-PST-N-RM
общественн-ыми бур-ями перестроечн-ых времён.
social-INS.PL tempest-INS.PL NAME-GEN.PL time.GEN.PL
As if the course of time was not interrupted by the social tempests related to the times of Perestroika.
[635, RNC, О свойствах постоянных величин // "Экран и сцена," 2004.05.06]

In addition to the canonical instantiations of the phasal reflexive verbs, the semantic component of the phasal segmentation can be used to motivate the patterns with such reflexive verbs as *браться* 'undertake, take on' and *приняться* 'proceed.' The latter verb appears in different argument constructions types in the sample, ranging from the Reflexive Engagement, when combined with the *за*$_{\text{acc}}$ pattern, to the Phasal Construction, as in 5.7-11.

5.7-11 *И* *пчёлк-и* *приня-л-и-сь* препира-ть-ся.
And bee-NOM.PL proceed-PST-PL-RM bicker-INF-RM
And the bees proceeded to bicker.
[1984, RNC, Татьяна Рик. Про вредную Бабку-Ёжку // "Мурзилка," №6," 2001]

Paducheva (Падучева, 2004:212) considers, for example, that the semantics of the verb *бросаться* 'rush' contains a temporal state. Moreover, the verbs *бросаться* 'rush,' *браться* 'undertake,' and *приняться* 'proceed' illustrate that the potentiality to combine with an infinitive cannot be straightforwardly derided from the neighbor verb. The cross-paradigmatic verbs *бросаться* 'rush'~ *бросать* 'throw, cast' both combine with an infinitive contrasting the *браться* 'undertake' ~ *брать* 'take, get' and *приняться* 'proceed' ~ *принять* 'take.' For the latter two cross-paradigmatic verbs, only the reflexive verbs readily combine with an infinitive (Divjak, 2004:21-28; Денисов & Морковкин, 2002).[112] This shows that multiple pathways are available for verbs and they can gravitate towards a common structure.

Divjak and Gries analyze nine Russian near-synonymous verbs pertaining to the semantic domain of trying. The following verbs are relevant for the purposes of the present study: *стараться*, *пытаться,* and *тщиться*. Based on corpus-data, the following labels are proposed for them. 'You could succeed' is profiled with the verbs *стараться* and *пытаться*. The verb *тщиться* is labeled as 'you can't succeed.' Thus, the proposed labels are used to differentiate the near-synonymous verbs in terms of partitioning the trying-event type (Divjak & Gries, 2006). For the purposes of the present study, this illustrates the level of granularity that can be imposed upon items when verb-specific constructions are the primary goal. For the present study, the label Phase Construction provides a coarse-grain level of proximity between them based on the Reflexive Marker.

The Phase Construction is also supported with decision verbs, for example, *решаться* 'decide' and *намереваться* 'intend.' Following Paducheva's (Падучева, 2004:310) definition, these verbs contain the phase of *make decision*. However, Gerritsen classifies the verb *решаться* as part of the consequential type discussed in relation to the Reflexive Engagement Construction. In contrast, Israeli (1997:84) considers that only the *на*$_{acc}$ pattern contains the consequential semantics, as in 5.7-12.

[112] A corpus search based on the disambiguated subcorpus of the Russian National Corpus gave one hit with the combination of *брать* + INF assuming a distance of one between the reflexive verb and the infinitive [RNC, Алексей Варламов. Купавна // Новый Мир, № 11-12, 2000].

5.7-12 *Я* *реши-л-ся* *на* *убийств-о.*
I.NOM decide-PST.M-RM PR murder-ACC
I decided on murder.
[Israeli 1997: 84][113]

This pattern is not attested in the database, but it falls under the label Reflexive Engagement Construction supported by the *на*$_{\text{acc}}$ pattern.

The other end of this proximity based on patterns is illustrated by such verbs as *готовиться* 'prepare' and *разучиться* 'be out of practice, forget how.' The former has a semantically similar neighbor verb, *готовить* 'prepare.' The former verb is also a prime example of the clustering of multiple patterns. Example 5.7-13 gives the Reflexive Engagement Construction profiled with the *к*$_{\text{dat}}$ pattern, whereas Example 5.7-14 demonstrates the infinitive pattern.

5.7-13 *Все были безумно рады и*
[…]
готови-л-и-сь *к* *больш-ому* *весель-ю.*
prepare-PST-PL-RM PR big-DAT festivity-DAT
Everybody was insanely happy and prepared for a big celebration.
[1727, RNC, Сергей Седов. Доброе сердце Робина // "Мурзилка," №7," 2002]

5.7-14 *Собинов-а* *готов-ит-ся* *отмети-ть* *ещё*
NAME-NOM prepare-3S.PRS-RM celebrate-INF more
одн-у *больш-ую* *дат-у.*
one-ACC big-ACC day-ACC
Sobinova is preparing to celebrate one more big day.
[854, RNC, Воспитываем подвижников-этномузыкантов // "Народное творчество," 2003.10.20]

The last attested construction type with this verb is the Passive Construction, as in 5.7-15.

5.7-15 *Сейчас как* *раз* *готов-я-тся* *различн-ые*
Now as just prepare-3P.PRS-RM various-NOM.PL
документ-ы.
document-NOM.PL
Just now various documents are being prepared.
[1450, RNC, Беседа на радио о государственной службе // 2004.11.15)]

Importantly, the neighbor verb *готовить* does not readily combine with an infinitive creating a gap in the otherwise smooth cross-paradigmatic relation. This particular verb demonstrates the problematic nature of grammatical and lexical structures when they are portrayed as a binary and mutually exclusive relation.

[113] The glossing was added by the author.

An opposite pathway can be illustrated with the verb *разучиться* 'be out of practice, forget how,' as given in 5.7-16.

5.7-16 *Мне* *каж-ет-ся* *я* *уж* *и* *езди-ть-то*
I.DAT seem-3S.PRS-RM I.NOM PART so drive-INF
разучи-л-а-сь.
be.out.of.practice-PST-F-RM
It seems to be that I have so forgotten how to drive.
[1282, RNC, Праздные разговоры (2006.04)][114]

The neighbor verb *разучить* 'learn' is perceived as semantically dissimilar and it does not readily combine with an infinitive, constituting a double gap in the cross-paradigmatic relation. At the same time, another complexity is involved with this particular verb, namely the prefix *раз-*. From a morphological perspective, the base verb is formed with the cross-paradigmatic verbs *учить* ~ *учиться* 'learn, study,' both of which can combine with an infinitive. However, if the reflexive verbs are simply always derived from the non-reflexive verb, one would assume that *разучить* also appears with an infinitive. In contrast, when the Reflexive Marker is assumed to constitute a system of its own, this particular verb can be motivated. The verb *разучиться* displays an intermediate level of detachment from the cross-paradigmatic relation, and a stronger gravitation towards the system of the Reflexive Marker. Excluding the Reflexive Marker, the only shared commonality is the phonological form.

5.8 Phase Construction

The Phase Construction contains 145 instances covering 42 unique reflexive verbs in the RF model. It is a generalized argument construction type that combines canonical phasal reflexive verbs with the Nominative Infinitive pattern. The semantics of the phase does not constitute a monolithic type (Divjak, 2004:125-126, 140-141). Instead, the integration of the phase forms a continuum ranging from verbs combining the Nominative Infinitive pattern to the traditional phasal reflexive verbs. Such verbs as *намереваться* 'intend' and *тщиться* 'try' form one extreme end of the continuum. They do not have neighbor verbs and systematically combine with an infinitive. Another set of verbs such as *стараться* 'try,' *надеяться* 'hope,' and *бояться* 'be afraid' is in close proximity based on structural properties, but they have a wider range of patterns available. The other end of the continuum can be considered to be populated by verbs that have semantically similar neighbor verbs. Examples are *браться* 'undertake,' *решаться* 'decide,' and *повторяться* 'repeat' followed by semantically intermediate reflexive verbs, like *стремиться* 'strive' and *собираться* 'intend.' Finally, the canonical phasal reflexives, for instance *продолжаться* 'continue' and *начинаться* 'begin,' are related to the previously mentioned continuum from a semantic point of view where the semantic component of phase is fully lexicalized. The definition of the Phase Construction follows.

[114] The postfix *-то* is commonly used in spoken Russian for emphatic purposes.

Function: Profiles an entity relative to phase.
Form: Nominative subject and verb-specific constructions for the secondary slot.

The function of the construction assumes that an entity is profiled relative to phase. The encoding of the argument construction displays verb-specific constructions covered by two subtypes. The fully lexicalized reflexive verbs form the traditional semantic group of phasal verbs and combine with traditional adjuncts, such as space and time. The second subtype supports the Nominative-Infinitive pattern but are differentiated based on the degree of integration with the infinitive, ranging from loosely connected to tightly connected ones. Example 5.8-1 illustrates the canonical Phase Construction and the encoding is given in Figure 5.8-1.

5.8-1 *Обычн-о* *вс-ё* *начина-ет-ся* *с*
Typical-ADV everything-NOM begin-3S.PRS-RM PR
реальн-ого *или* *электронн-ого* *письм-а* […].
real-GEN or electronic-GEN letter-GEN
Typically, everything begins from a real mail or e-mail.
[850, RNC, Игорь Сирин. Свой путь // "Бизнес-журнал," 2003.10.23]

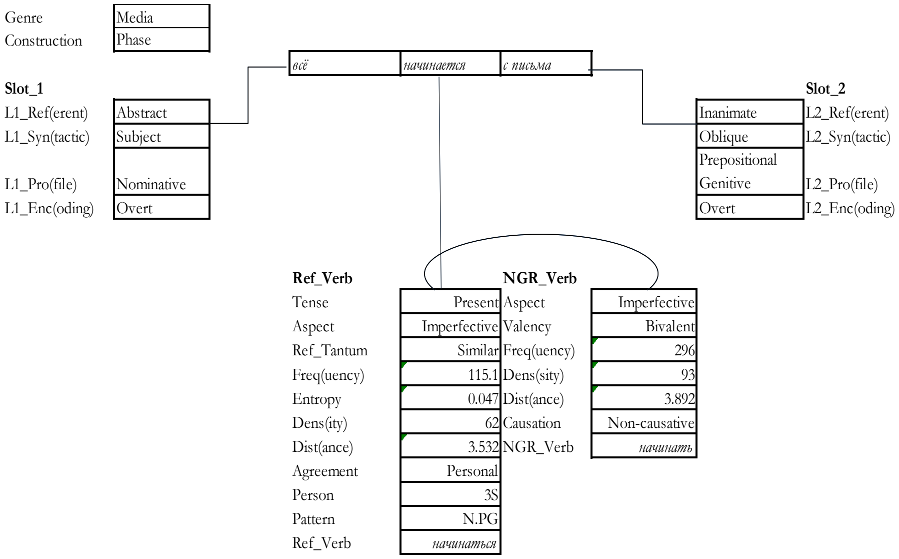


Figure 5.8-1 Layered structure of the canonical Phase Construction.

The confusion matrix is given in Table 5.8-1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	**0**	**0**	**1**	**0**	**0**	**0**	**4**	**127**	**0**	**0**	**2**	**11**	**0**
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 5.8-1 Confusion matrix of the predicted constructions. The Phase Construction is given in bold.

The class-wise error is fairly small with the value of 0.124. Additionally, both the recall of 0.876 and the precision of 0.977 indicate that the construction type is well-separated globally. The performance of the model is higher in classifying the instances compared to indentifying them.

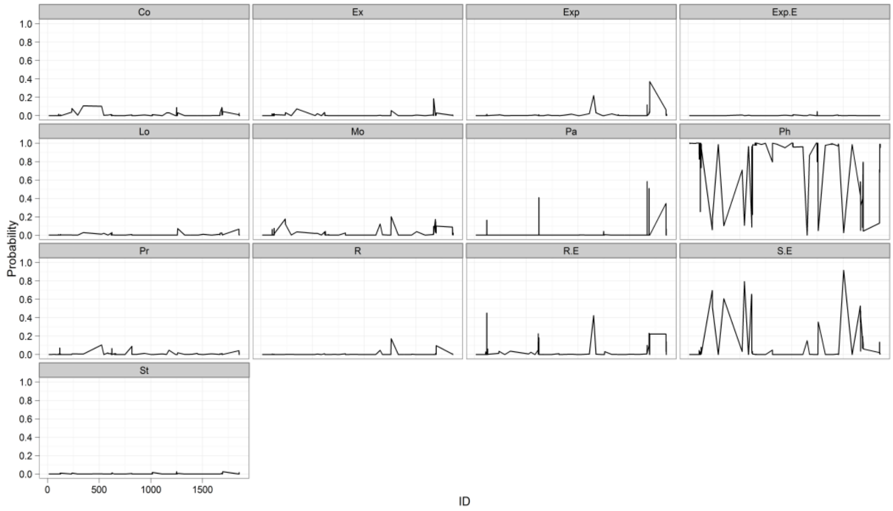


Figure 5.8-2 Faceted estimated class probability plot of the Phase Construction (Ph). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

Figure 5.8-2 illustrates that the competition between different construction types appears to be generally weak, but a slight competition with the Passive and the Spontaneous Event Constructions is, nonetheless, present. For example, *завершиться* 'end, be over' is classified as an instance of the Passive, possibly because the N.I pattern is strongly associated with the latter type, leading to over-generalization with these instantiations. The fluctuation is strongly present with the canonical phasal verbs, as their profile is more dependent on the construal of the event.

5.9 Definitions of the Reciprocal Construction

Reciprocal Construction is typologically associated with reflexive markers (Bostoen & Nzang-Bie, 2010; Kemmer, 1993; König & Gast, 2008). Lichtenberk (1985:21) gives the following definition for the canonical reciprocal construction: "There are two participants, A and B, and the relation in which A stands to B is the same as that in which B stands for A." This definition is taken as a starting point for example in studies by Kemmer (1993:96-97) and Knyazev (Князев, 2007:316-317). Extensive survey on expressing reciprocity in Russian is offered by Knyazev (Knjazev, 2007). Example 5.9-1 demonstrates this canonical pattern in Russian.

5.9-1 [*м*]ы *никогда* *не* *ссор-им-ся,* [...].
We.NOM never NEG argue-1P.PRS-RM
We never argue.
[780, RNC, Я желанна. Разве это стыдно? // "Даша," №10," 2004]

The subject argument is in the nominative case and plural form by necessity, and the configuration consists of two referents, implying a symmetrical relation holding between them, cf. Examples 5.9-2 and 5.9-3.

5.9-2 *Они* *познакоми-л-и-сь* *на* *танцплощадк-е.*
They.NOM become.friends-PST-PL-RM PR dance.floor-PREP
They became friends on the dance floor.
[1650, RNC, Токарева Виктория. Своя правда // ""Новый Мир", №9", 2002]

5.9-3 *И* *они* *нача-л-и* *дра-ть-ся.*
And they.NOM begin-PST-PL fight-INF-RM
And they began to fight (each other).
[1815, RNC, Андрей Геласимов. Жанна (2001)]

Knyazev establishes five basic types for expressing reciprocity in Russian (Князев, 2007:320-322; cf. Янко-Триницкая, 1962:190). These subtypes are illustrated with reflexive verbs attested in the sample.

1) Reflexiva tantum verbs: *бороться* 'fight'
2) Reflexive and non-reflexive verbs designate reciprocity but differ in causation: *драться* 'scuffle'

3) Canonical reciprocal verbs: *целоваться* ‘kiss (each other)’[115]
4) Multiplicative verbs based on prefix, suffix and a Reflexive Marker: not attested
5) Prefix and Reflexive Marker: *спеться* ‘come to agreement’

The first type shows that the Reciprocal Construction has a lexical basis supported by the reflexiva tantum verbs similar to most construction types marked with the Reflexive Marker. At the same time, the first type is dependent on the definition of reflexiva tantum. For instance, Knyazev (Князев, 2007:321) considers that the verb *согласиться* ‘agree’ differs in meaning compared to its neighbor, whereas it is regarded as intermediate in this study. On the other hand, the neighbor verb of *бороться* ‘fight’ is regarded to be archaic in contemporary Russian. The same reservation also applies to the second type where the definition of causative neighbor verb is most likely to vary across different studies. Knyazev (Князев, 2007:321) states that the neighbor verb of *ссориться* ‘argue (with each other)’ is causative in a sense that *ссорить* ‘argue’ profiles an internally caused change of state of the subject argument. In contrast, the neighbor verb is not tagged as causative in the Russian National Corpus. Hence, verbs containing the semantics of the internal causation are not considered as causative in this study.

The third type is closest to the traditional derivational definition where the contribution of the Reflexive Marker to reciprocity is clearest, as in 5.9-4 with the verb *видеться* ‘see (each other)’ ~ *видеть* ‘see.’

5.9-4	*Не*	*виде-л-и-сь*	*уже,*	*наверное,*	*лет*	*семь.*
	NEG	see-PST-PL-RM	even	probably	year.GEN.PL	seven.

They probably have not seen each other in about seven years.
[1851, RNC, Андрей Геласимов. Ты можешь (2001)]

The fourth type is not attested in this study. This type is strictly defined in terms of derivational potential of a verb and is established through the following three components: prefix *пере-*, the polyfunctional suffix *-ива-*, and the Reflexive Marker *-ся* (Князев, 2007:337). The only reflexive verb in the sample close to the formal definition of the fourth type is *переглянуться* ‘exchange looks.’[116] However, the verb is formed with the semelfactive suffix *-ну-*, profiling an event type where the activity is performed once (Dickey & Janda, 2009).

The fifth subtype is typically classified as a separate class in the early Russian tradition (cf. Виноградов, 1972:500). However, the combination with the prefix *раз-* and motion verbs creates diffusive usage patterns (cf. Janda & Nesset, 2010). When we move away from the canonical instances, the boundaries between construction types get blurred, as illustrated in 5.9-5

[115] The neighbour verb *целовать* ’kiss’ is tagged as causative in the Russian National Corpus.

[116] The multiplicative reflexive verb is *переглядываться* (Князев, 2007:338).

5.9-5 […] *и* *они* *разбежа-л-и-сь.*
and they.NOM run.away-PST-PL-RM
And they ran away.
[1840, RNC, Андрей Геласимов. Нежный возраст (2001)]

In terms of structural properties, the reflexive verb *разбежаться* 'run away' does not have a neighbor verb in Russian displaying simultaneously multiple properties - detachment from the cross-paradigmatic relation and prefixation. Similar behavior is attested with the reflexive verbs *разбегаться* 'scatter,' *расстаться* 'leave, part ways,' *срастаться* 'grow together,' and *созвониться* 'get in touch by phone.' Thus, the prefix and the Reflexive Marker are fused together. From a semantic point of view, the verb pertains to the category of collective reciprocals (Lichtenberk, 1985). Kemmer follows this distinction and characterizes the latter as an event type in which the action is carried out jointly but it lacks a clear endpoint, for example *the guests left* (Kemmer, 1993:98-99).

Another important subtype of reciprocals is established with the preposition *с*$_{\text{ins}}$, commonly labeled as Comitative. The inclusion of the Comitative breaks away from the symmetrical relation and a single entity. The subject argument, occupies a more prominent position (cf. Janda, 1993b:184). Knyazev (Князев, 2007) considers this pattern as semi-symmetrical. This subtype is exemplified in 5.9-6–5.9-8.

5.9-6 [*я*] *не* *мог-у* *с* *вами* *согласи-ть-ся.*
I.NOM NEG can-1S.PRS PR you.INS.PL agree-INF-RM
I cannot agree with you.
[1415, RNC, Беседа с социологом на общественно-политические темы, Санкт- Петербург // ФОМ (2004.01.27)]

5.9-7 […] *А. Т. Ф. действует как самый обыкновенный гуманитарий:*
[…]
выдвигает гипотезы и указывает факты,
[…]
котор-ые *согласу-ют-ся* *с* *этими* *гипотез-ами.*
which-NOM.PL agree-3P.PRS-RM PR this.INS.PL hypothesis-INS.PL
A. T. F. acts like an ordinary scholar: proposes hypotheses and indicates facts that agree with these hypotheses.
[424, RNC, А. А. Зализняк. Лингвистика по А. Т. Фоменко]

5.9-8 *Ну, она типа сказала,*
[…]
что *я* *с* *ней* *не* *больше* *хоч-у*
that I.NOM PR she.INS NEG more want-1S.PRS
встреча-ть-ся *из* *чувств-а* *мест-и.*
meet-INF-RM PR sense-GEN revenge-GEN
Well, she said like, that I do not want to meet her any more of sense of revenge.
[1045, RNC, Разговор на улице между мужчиной и женщиной (2005.04.13)]

5.9-9 *Они* *пожени-л-и-сь.*
They.NOM marry-PST-PL-RM
They got married
[1659, RNC, Токарева Виктория. Своя правда // ""Новый Мир," №9," 2002]

In contrast, the reflexive verb *пожениться* 'get married' gravitates towards the canonical Reciprocal Construction, as in 5.9-9.

5.10 Reciprocal Construction

The Reciprocal Construction contains 103 data points and covers 56 unique reflexive verbs in the RF model. The Reciprocal Construction displays similar structural behavior as most construction types marked with the Reflexive Marker in Russian. It is a lexically supported center with reflexive verbs that have become detached from the cross-paradigmatic relation, (e.g., *бороться* 'fight,' *общаться* 'communicate,' and *здороваться* 'greet'). This is contrasted with the cross-paradigmatic verbs functioning as anchors across paradigms, for instance *целоваться* 'kiss,' and *шептаться* 'whisper.' Finally, a mixed type is established in interaction with prefixation that creates gaps in the cross-paradigmatic relation, for example *разбежаться* 'run away' and *уживаться* 'get along.' The definition of the Reciprocal Construction follows.

Function: Profiles a symmetrical relation between entities.
Form: Nominative subject and verb-specific constructions for the secondary slot.

The function of the Reciprocal Construction assumes that multiple entities are profiled and the profiled relation holding between them is symmetrical. The Reciprocal Constructions have two specific subtypes for encoding the secondary slot as either the second referent is incorporated or the secondary referent is encoded with the c_{ins} pattern. In the case of the former type, the secondary slot is encoded with the traditional adjunct that is used to profile the localization of the event in time or space, for instance. Examples 5.10-1 and 5.10-2 illustrate the Reciprocal Construction along with the two subtypes supported through the reflexive verb *собраться* 'gather.'

5.10-1 *Мы* *с* *вами* *собра-л-и-сь* для того
We.NOM PR you.INS.PL gather-PST-PL-RM PR that.GEN
чтобы обсудить некоторые политические проблемы и события […]
We gathered with you to discuss certain political problems and developments.
[1127, RNC, Беседа в Новосибирске (2000.08.15)]

5.10-2 *Тысяч-и* *людей* *собра-л-и-сь*
Thousand-NOM.PL people.GEN.PL gather-PST-PL-RM
на *Трафальгарск-ой* *площад-и,* […].
PR Trafalgar-PREP Square-PREP
Thousands of people gathtered at Trafalgar Square.

[539, RNC, Александр Смотров. Тысячи людей собрались на Трафальгарской площади, чтобы увидеть новую версию знаменитого фильма"Броненосец Потемкин" // "РИА "Новости," 2004.09.13]

The encoding is given in Figure 5.10-1 and the confusion matrix in Table 5.10-1.

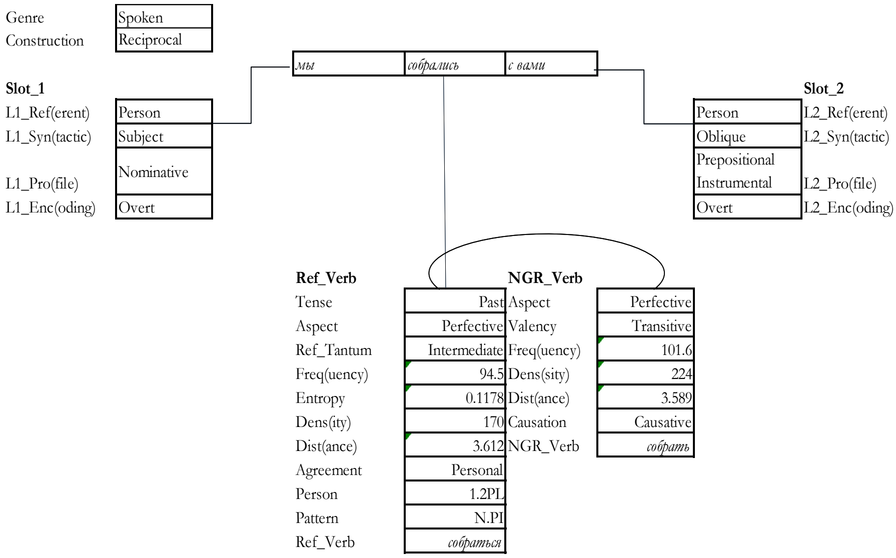


Figure 5.10-1 Layered structure of the canonical Reciprocal Construction.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	**0**	**0**	**2**	**0**	**1**	**13**	**0**	**0**	**0**	**70**	**13**	**4**	**0**
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 5.10-1 Confusion matrix of the predicted constructions. The Reciprocal Construction is given in bold.

The class-wise error of 0.3256 indicates that the Reciprocal Construction does not readily arise from the input similar to the Reflexive Engagement Construction. A more detailed analysis is achieved with the precision and recall values. The recall value of 0.679 indicates that the model has difficulty in identifying the instantiations of the Reciprocal Construction in usage. On the other hand, the precision value of 0.777 indicates that the model is better at classifying them.

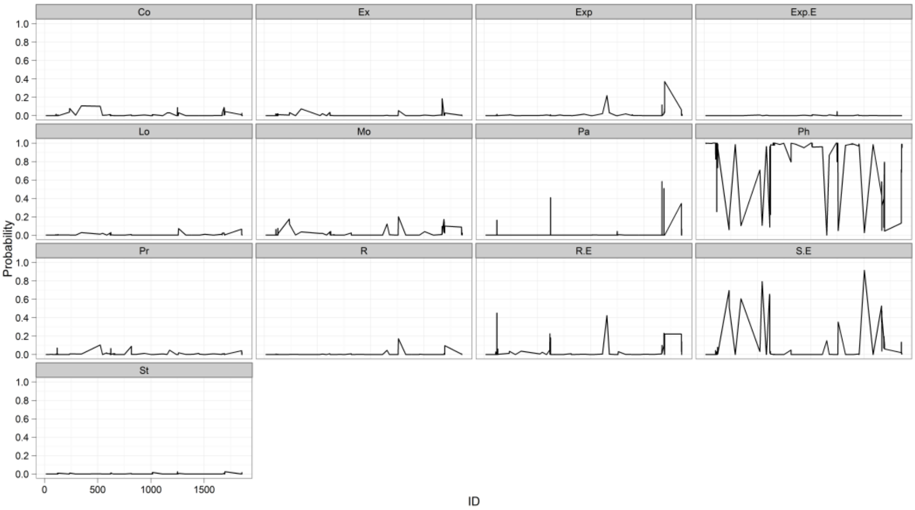


Figure 5.10-2 Faceted estimated class probability plot of the Reciprocal Construction (R). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

Figure 5.10-2 indicates a slight fluctuation with the Reciprocal Construction (R). A competition is present between the Reflexive Engagement and the Motion Constructions. A misclassification pertaining to the former type is illustrated in 5.10-3 and the latter one in 5.10-4.

5.10-3 *Кремы для тела, в состав которых входит ментол,*
[…]

успешн-о	*бор-ют-ся*	*с*	“апельсинов-ой	корк-ой”.
succesful-ADV	fight-3P.PRS-RM	PR	orange-INS	rind-INS

Body lotions that contain menthol sucessfully fight against the “orange rind effect”.
[723, RNC, На заметку // "Даша," №10," 2004]

5.10-4

Холмогор-ов	*сродни-л-ся*	*с*	*глух-им,* […].
NAME-NOM	become.friends-PST.M-RM	PR	deaf-INS

Holmogorov became friends with a deaf.
[1976, RNC, Олег Павлов. Карагандинские девятины, или Повесть последних дней // ""Октябрь", №8", 2001]

When usage patterns deviate from the canonical instances in the sample, they become in close proximity with other types. For instance, metaphorical extensions expand the range of structural properties, as in 5.10-3, where the estimated class probability for the Reciprocal becomes low (0.11) compared to the Reflexive Engagement (0.251). This shows that the structural properties offer cue validity in canonical cases, whereas deviations are handled locally by the semantics of a particular verb or, ultimately, the whole context.

5.11 Definitions of the Content Construction

The Content Construction is used in this study to unify certain reflexive verbs and patterns that have been previously either excluded or classified as instances of some form of impersonal type. From a semantic point of view, the Content Construction converges towards profiling a content either communicated or perceived. Assuming that activities performed by a human participant constitute a basic type in language, the Content Construction confines to the pattern who-communicated-what in its basic configuration (cf. Падучева, 2004:355-356). Reflexive verbs pertaining to this pattern have not been included systematically in previous taxonomies. Most likely they would fall under some form of middle construction.[117] Examples 5.11-1–5.11-4 illustrate the basic pattern.

5.11-1 *Я извиня-ю-сь что задержа-л.*
I.NOM apologize-1S.PRS-RM that hold.up-PST.M
I apologize that I hold you up.
[1171, RNC, Разговор в офисе страховой компании (2006.11)]

5.11-2 *Предложи-л закури-ть, но* я отказа-л-ся.
offer-PST.M smoke-INF but I.NOM refuse-PST.M-RM
He offered a cigarette but I refused.
[1833, RNC, Андрей Геласимов. Нежный возраст (2001)]

5.11-3 *Бы-л-о врем-я, когда мы жалова-л-и-сь*
be-PST-N time-NOM when we.NOM complain-PST-PL-RM
на отсутстви-е дебют-ов.
PR lack-ACC debut-GEN.PL
потом на отсутствие качественных дебютов.
[…]
There was a time, when we complained about the lack of debuts but now about the quality of debuts.
[611, RNC, Весенний призыв // "Экран и сцена," 2004.05.06]

5.11-4 *Я слыша-л что Путин положительн-о*
I.NOM hear-PST.M that NAME.NOM positive-ADV
отзыва-ет-ся о НАТО.
speak-3S.PRS-RM PR NAME.PREP
I heard that Putin speaks positively about NATO.
[1011, RNC, Беседа в Новосибирске (2004.04.06)]

[117] Paducheva (Падучева, 2004:370) considers that the verb *отказаться* 'refuse' belongs to the category of Semantic Reflexive, co-reference of Agent and Patient.

These verbs do not conveniently follow any established derivational patterns. Another important property that points to the conclusion that these reflexive verbs are basic is the frequency of use relative to their neighbor verbs: *извиняться* 'apologize' (14.1 freq) ~ *извинять* 'forgive' (1.6 freq), *отказаться* 'refuse' (115.8 freq) ~ *отказать* 'refuse' (32.3 freq), *отзываться* 'respond, answer' (14.3 freq) ~ *отзывать* 'respond, answer' (1.1 freq), and *жаловаться* 'complain' (46.1 freq) ~ *жаловать* 'accord, like' (4.2 freq). Another complication for a derivational motivation is the perceived semantic similarity. The first two reflexive verbs are perceived as similar contrasting the latter two verbs as *отзываться* is intermediate, whereas *жаловаться* is dissimilar. Additionally, certain verbs of communication, such as *откликнуться* 'answer, shout back,' lack a neighbor verb. Thus, these particular verbs show that the center can be supported by both the cross-paradigmatic relation and the detached reflexive verbs.

In addition to this basic pattern, the semantics of the Content Construction appear to be highly augmentable in Russian, leading to a range of divergent patterns. These patterns have been explored in previous studies. The motivation behind this is most likely the fact that an augmentation by definition is more amiable for a derivational analysis. They are often labeled under the Impersonal Passive, simply under the Passive or some form of generalized impersonal type (Виноградов, 1972; Шахматов, 1925). For example, Israeli (1997:166-168) considers that infinitives and subordinate clauses also occupy the subject slot in the Passive Construction, as illustrated in 5.11-5.

5.11-5 *И* *ещёг* *овор-ит-ся,* *что* [...].
And also speak-3S.PRS-RM that [...]
And it is also said that.
[506, RNC, В.Н. Комаров. Тайны пространства и времени (1995
2000)]

The communicator in the event is backgrounded and is not present in the profile in Example 5.11-5. Plungyan (Плунгян, 2000:203) considers similar patterns as pertaining to the semantic function of highlighting the process.

Knyazev (Князев, 2007:302-303) states that the category of the Impersonal Passive is available for verbs which typically lack the Patient argument, and it is formed with intransitive verbs or transitive verbs which are used intransitively. A similar point of view is already expressed by Janko-Trinickaya (cf. Янко-Триницкая, 1962:72-76). Gerritsen distinguishes three types based on the argument structure of the neighbor verb: 1) oblique object instead of accusative, 2) subordinate clause instead of accusative, and 3) infinitive instead of accusative. Thus, several types are labeled as improper impersonal (Gerritsen, 1990:127). These augmentations are demonstrated with the verb *предлагаться* 'propose.' The reflexive verb has a semantically similar and bivalent neighbor verb *предлагать* 'propose.' A dative argument can be included in the profile with this reflexive verb given in 5.11-6.

5.11-6 *Им предлага-л-ся, например, коротк-ий*
They.DAT propose-PST.M-RM for.example short-NOM
рассказ о женщин-е […].
tale.NOM PR woman-PREP
For example, a short tale about a woman was proposed for them.
[37, RNC, Михаил Арапов. Когда текст обретает смысл // "Знание — сила," №1," 2003]

The profile can be further augmented by excluding the nominative subject and including an infinitive, as in 5.11-8 and by contrasting Examples 5.11-6 and 5.11-7.

5.11-7 *В стать-е предлага-ет-ся подход к*
PR article-PREP propose-3S.PRS-RM approach.NOM PR
решени-ю задач синтез-а топологи-и
solution-DAT problem.GEN.PL synthesis-GEN topology-GEN
An approach to solve problems of synthesis of topology is proposed i the article.
[17, RNC, Задачи синтеза сетей синхронной иерархии // "Информационные технологии," 2003]

5.11-8 *Помимо этого предлага-ет-ся раздели-ть*
PR this.GEN propose-3P.PRS-RM divide-INF
местн-ые бюджет-ы […].
local-ACC.PL budget-ACC.PL
Besides this, the local budgets are proposed to be divided.
[993, RNC, Минфин корректирует Бюджетный кодекс // "Время МН," 2003.08.06]

Paducheva (Падучева, 2004:197-198) considers that verbs containing the component of receiving or giving information pertain to the category of perception. These instantiations are illustrated in 5.11-9 and 5.11-10.

5.11-9 *Оказыва-ет-ся, вс-ё нача-л-о-сь*
Appear-3S.PRS-RM everything-NOM begin-PST-N-RM
в Германи-и […].
PR NAME-PREP
It seems that everything began in Germany.
[91, RNC, Н. Ю. Феоктистова. Новогородняя ёлка // "Первое сентября," 2003]

The person whose perspective is imposed upon the content can be included in the profile with the dative case, as in 5.11-10 (cf. Падучева, 2004:210-212).

5.11-10 *Мне каж-ет-ся это* стил-ь так-ой.
I.DAT seem-3S.PRS-RM it.NOM style.NOM that.kind-NOM
It seems to me that it is that kind of style
[1211, RNC, Праздный разговор (2006.11)]

These instantiations also support the reflexive verb *видеться* 'seem, appear' when used in the Content Construction. At the same time, these particular instances resemble units. They are used as discourse connectors, especially when combined with the first person pronoun (cf. Scheibman, 2002:64-67). Additionally, mental verbs can gravitate towards the Content Construction when used in the impersonal subtype, as in 5.11-11.

5.11-11 *Положени-е странно-е, но ведь*
Situation-NOM strange-NOM but indeed
чувству-ет-ся, что эт-о действительн-о так.
feel-3S-PRS-RM that it.NOM really-ADV so
The situation is strange but it, indeed, feels that it is so.
[464, RNC, С. Г. Бочаров. Из истории понимания Пушкина (1998)]

5.11-12 *Мне прост-о не вери-л-о-сь,*
I.DAT simple-ADV NEG believe-PST-N-RM
что такое можно сочинить.
[…]
I simply could not believe that it was possible to invent something like that.
[616, RNC, Джим Кэрри - изнутри и снаружи // "Экран и сцена," 2004.05.06]

The Content Construction can be used to motivate lexicalized patterns which function as discourse connectors. For example, the reflexive verb *касаться* ranges from 'touch', (the Reflexive Engagement Construction), to 'concern,' (the Content Construction), as given in 5.11-13.

5.11-13 *Что же каса-ет-ся синхронистическ-ого мышлени-я,*
It.NOM still concern-3S.PRS-RM contemporary-GEN thinking-GEN
то его можно назвать "пространственным"; […].
[…]
As for the contemporary thinking, it can be called "spatial."
[503, RNC, В.Н. Комаров. Тайны пространства и времени (1995-2000)]

Another lexicalized pattern with the Content Construction is profiled with *имеется в виду* 'mean,' as in 5.11-14.

5.11-14 [*и*]*ме-ет-ся в вид-у что*
have-3S.PRS-RM PR view-PREP that
есть достаточно большое количество должностей.
[…]
Meaning that there is a sufficiently large number of posts.
[1448, RNC, Беседа на радио о государственной службе // (2004.11.15)]

Another idiomatic usage pattern is attested with the reflexive verb *огрызнуться* 'snap' that simultaneously resembles the semantic class of verbs of sound (Падучева, 2004:401-402, 420).

5.11-15 *Как, как? огрызну-л-а-сь Зин-а.*
How How snap-PST-F-RM NAME-NOM
How, how? snapped Zina.
[1685, RNC, Виктор Кологрив. Медовый луг // "Мурзилка", №5," 2002]

In the non-communicative usage, the verb could be construed as an instance of the incorporated object type, (i.e., 'show one's teeth). However, the reflexive verb lacks a neighbor verb, at least in Contemporary Russian.

The recognition of the Content Construction forms a nexus between the impersonal types, with the fully frozen units that have been primarily excluded in the previous studies. The Content Construction also displays another semantic connectivity between verbs. For instance, certain reflexive verbs profiling the Reciprocal Construction, such as *общаться* 'communicate' and *шептаться* 'whisper,' align with the instantiations of the Content Construction from a semantic perspective.

5.12 Content Construction

The Content Construction contains 110 instances and covers 43 unique reflexive verbs in the RF model. This particular construction type is supported by the high frequency verbs of perception, such as *оказаться* 'seem appear,' and *казаться* 'seem, appear.' Another small cluster is formed with the verbs of communication such as *говориться* 'speak' and *общаться* 'communicate.' These verbs illustrate the divergence of the cross-paradigmatic, as *оказаться* is perceived semantically dissimilar to its neighbor, contrasting *казаться* that is intermediate. The verbs of communication differ also, as *говориться* is perceived as similar to its neighbor, whereas *общаться* does not have a neighbor. The definition of the Content Construction follows.

Function: Profiles a content of communication or perception.
Form: Subject and verb-specific constructions for the secondary slot.

There are three primary verb-specific construction types. First, the personal type is primarily established with the verbs of communication, such as *извиняться* 'apologize,' and *жаловаться* 'complain.' Second, the impersonal Clausal subject type is supported with *говориться* 'speak,' *оказаться* 'seem appear,' and *выясниться* 'turn out.' Third, the Infinitive subject is attested with the following verbs: *предлагаться* 'propose, suggest,' and *полагаться* 'suppose, consider.'[118] These patterns appear to be more centered on the reflexive verbs rather than general patterns associated with the neighbor verbs and less dependent on the construal of the event compared to the Spontaneous Event. This demarcation is indented

[118] Depending on the solution for these verbs in this particular configuration, they are also often labeled under the Passive.

to be captured with the differentiation between the construal-specific and the verb-specific configuration. Example 5.12-1 illustrates the basic usage pattern and the encoding is given in Figure 5.12-1.

5.12-1 *Родител-и*	*во*	*вс-ём*	*мир-е*	*жалу-ют-ся*
Parent-NOM.PL	PR	all-PREP	world-PREP	complain-3P.PRS-RM
на	*то,*			
PR	that.ACC			

что детей невозможно оторвать от компьютера, […].

[…]

Parents around the whole world are complaining that it is impossible take away the children from the computer.

[689, RNC, Юлия Ковалева. Комментарий психолога // "Даша", №10," 2004]

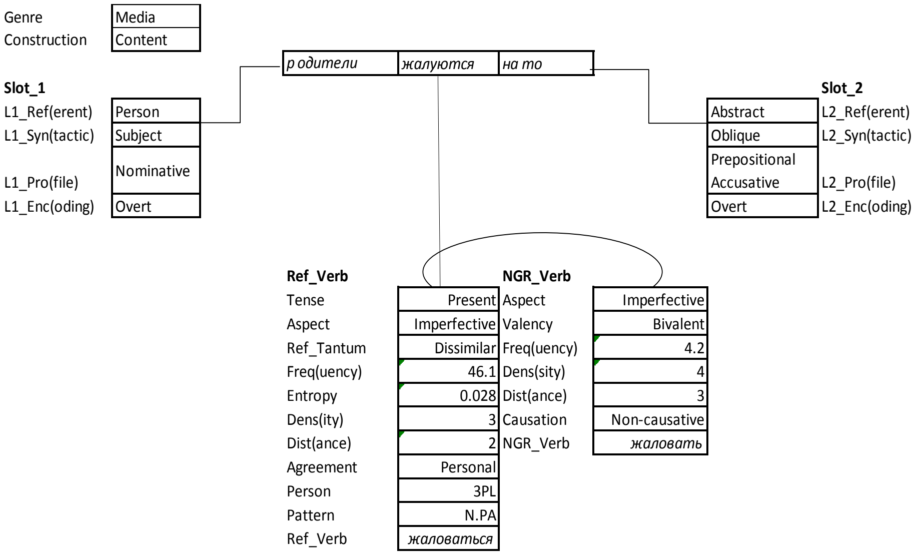


Figure 5.12-1 Layered structure of the canonical Content Construction.

The confusion matrix is given in Table 5.12-1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	**93**	**0**	**5**	**0**	**0**	**3**	**0**	**0**	**0**	**0**	**4**	**5**	**0**
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 5.12-1 Confusion matrix of the predicted constructions. The Content Construction is given in bold.

The class-wise error of 0.155 is fairly low in the model. Additionally, both the recall of 0.845 and precision of 0.939 indicate that the model is able to identify and classify these instances based on the structural properties. The precision value also shows that the structural properties offer a slight advantage in classifying the instances compared to identifying the instantiations.

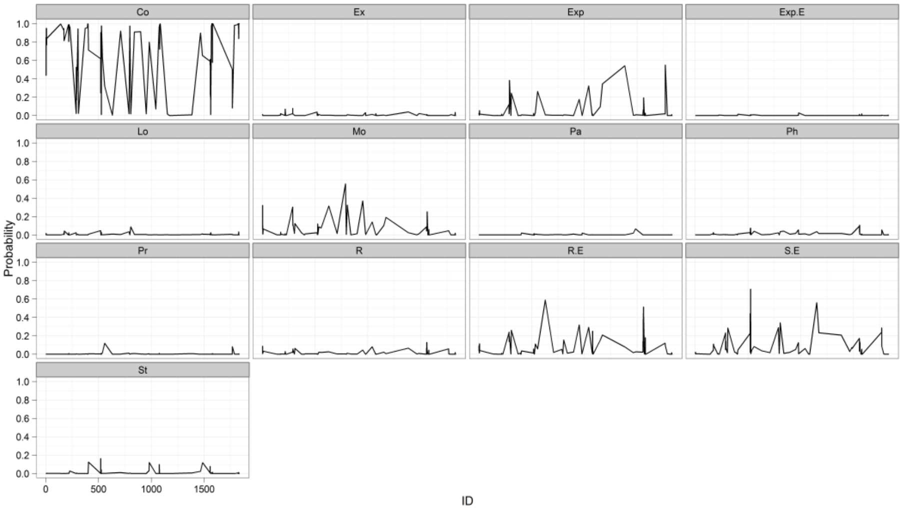


Figure 5.12-2 Faceted estimated class probability plot of the Content Construction (Co). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The estimated class probabilities show that the misclassified instances appear to be poor candidates for the Content Construction, displaying a strong deviation from the other attested patterns within this construction. For instance, Example 5.12-2 is classified as the Experiencer Construction with the estimated class probability of 0.541 compared to the estimated class probability of 0.008 with the Content Construction.

5.12-2 *Толстяк* *месяц* *уже* *обща-ет-ся*
Fatty.NOM month.ACC already talk-3S.PRS-RM
и тут узнает что это мальчик на самом деле.
[…]
The fatty has been talking for a month already and then learns that it is a boy in fact.
[1057, RNC, Праздный разговор молодых людей, Московская область (2005)

This behavior illustrates the strong effect of alignment. When the reflexive verbs appear in their typical configuration, the structural properties appear to offer sufficient cue validity. However, deviations lead to mixed properties and the estimated class probabilities are in flux. It seems that the semantics of the reflexive verb anchor these cases over the structural properties when compared globally across different construction types.

6 Mental Constructions

Mental events are among the basic conceptual types, in addition to spatial events, in describing human experience. On a fine-grained level, mental verbs are often partitioned into verbs of emotion, perception, and cognition (Croft, 1991:213-214; Падучева, 2004:197-198, 269-270, 273-274). Related to this, Kemmer makes a similar division between mental verbs that are marked with the Reflexive Marker. She also divides these verbs into simple and complex structures. The latter type is profiled with a dependent event type, like a subordinate clause, or with an infinitive (Kemmer, 1993:127-136). In the early Russian tradition, the majority of these verbs have been grouped under the label General Reflexive, cf. Section 1.2.1. The type depicts an activity which is not directed towards the object but stays within the sphere of the subject (Виноградов, 1975:496). In contrast, Janko-Trinickaya (Янко-Триницкая, 1962:149-153) labels certain verbs of emotion such as *волноваться* 'worry,' as expressions of internal experience.

The primary configuration of the mental event is related to the encoding of the basic semantic structure of the argument constructions. Generally, mental verbs contain two distinct roles: an entity undergoing the mental activity and a target of that mental activity. For the verbs of emotion, this configuration yields the traditional distinction between the Experiencer and Stimulus roles (Croft, 1991:216-219; Падучева, 2004:278). For purposes of the present study, the encoding of this typical configuration is taken as a starting point. Typicality refers to the possibility to extend the basic Experiencer/Stimulus configuration to cover the elaboration of the participant construed in the event (cf. Langacker, 1991:194-195). Thus, verbs of cognition, such as *удивляться* 'wonder' and *разобраться* 'grasp,' are analyzed under the label of the Mental Event. Section 6.1 introduces the semantic characterization of the argument constructions and its relation to usage-patterns. Sections 6.2–6.4 are used to establish the realizations and anchor them to three general argument construction types: the Experiencer, the Experiencer Extension, and the Stimulus Construction.

6.1 Definitions of the Mental Constructions

The primary motivation for unifying the mental verbs in this study is the observation that verbs of emotion and cognition can fluctuate between different interpretations. For example, Paducheva (Падучева, 2004:274) shows this behavior with the reflexive verb *бояться* 'be afraid' being labeled as emotion in 6.1-1, and cognition in 6.1-2.[119]

[119] Paducheva uses the term mental.

6.1-1 *Я* *тебя* *бо-ю-сь.*
I.NOM you.GEN be.afraid-1S.PRS-RM
I am afraid of you.
(Падучева, 2004:274)[120]

6.1-2 *Бо-ю-сь,* *что* *ты* *не* *прав*
Be.afraid-1S.PRS-RM that you.NOM NEG correct.NOM
I'm afraid that you are wrong.
(Падучева, 2004:274)

Similarly, Mel'chuk (Мельчук, 1995:89) considers that both the reflexive verbs *бояться* 'be afraid,' and *надеяться* 'hope' are weak emotional verbs. Importantly, both of these verbs lack a neighbor verb. This subtype contains 255 data points, and covers 89 unique reflexive verbs in the final sample. Importantly, 34 of these unique reflexive verbs do not have a neighbor verb, leading to a situation where a strong lexical support is present. Examples are, *гневаться* 'be angry,' *соскучиться* 'become bored,' *поглянуться* 'like,' and *любоваться* admire.' Thus, these verbs lacking the cross-paradigmatic relation can be considered as forming a lexical gravitational center for the paradigm of the Reflexive Marker within this sutype.

Another complication in terms of cross-paradigmatic relations is connected to the reflexive verbs that are perceived semantically as either intermediate or dissimilar compared to their neighbor verb. Although the number of these verbs is low in the sample, (15 unique reflexive verbs), a degree of detachment from the cross-paradigmatic relation is present. Such verbs as *стесняться* 'be ashamed,' *прийтись* 'be necessary,' and *чудиться* 'fancy' exemplify this latter group. Consequently, these verbs are typically either excluded from the analysis or labeled under some form of middle in previous studies.

The reflexive verbs confining the full cross-paradigmatic relation are often labeled under the Decausative (Князев, 2007:283-287; Падучева, 2001). Israeli (1997:65-66) uses the label Emotional Decausative and Gerritsen (1990:58-63) posits the category of Reactional Decausative. In derivational accounts, the causative verbs are often regarded as establishing the core of the mental verbs (Падучева, 2004:276). Examples 6.1-1–6.1-5 illustrate these verbs: *встревожиться* 'grow suspicious,' *успокоиться* 'settle down,' and *волноваться* 'worry.' They have a semantically similar causative neighbor verb.

6.1-3 *Перв-ый* *раз* *он* *встревожи-л-ся,*
First-ACC time.ACC he.NOM grow.suspicious-PST.M-RM
когда на огромном скошенном лугу, где они приземлились никто их
[…]
не встретил, никакого, партизанского дозора там не было
[…]

[120] The glossing and translations were added to the examples by the author.

First time he grew suspicious, when on the large, tampered field, where they landed, no one was there to meet them and there was not any partisan patrol.
[RNC, Василь Быков. Болото (2001)]

6.1-4 *Малань-я вздохну-л-а.*
[…]
Понемногу успокои-л-а-сь.
Little.bit settle.down-PST-F-RM
Malanja drew breath. She settled down a little bit.
[1993, RNC, Олег Тихомиров. Про козла Тихомира // "Мурзилка",№2", 2001]

6.1-5 *Я уж волну-ю-сь.*
I.NOM really worry-1S.PRS-RM
I am really worried.
[1144, RNC, Телефонные разговоры // М. В. Китайгородская, Н. Н. Розанова. Речь москвичей: Коммуникативно-культурологический аспект. М.: ИРЯ РАН, 1999]

The causative neighbor verbs are often postulated to constitute the basic type and the reflexive verbs are derived from them (cf. Grimshaw, 1990; Sonnenhauser, 2010). The frequency distribution of these verbs, however, portrays a different perspective. Only four neighbor verbs have a higher frequency compared to the reflexive verb: *взволноваться* 'work up,' *порадоваться* 'rejoice,' *встревожиться* 'grow suspicious,' and *пугаться* 'become frightened.' Although there are only 22 unique verbs left in this partition that satisfy the properties of causative cross-paradigmatic relation, the distributions lean towards the usage-based model where the log ratio is one of the determining factors in establishing the perceived basic type (Bybee, 2010; Hay, 2001). Thus, the decausative alternation is included by proxy through the cross-paradigmatic relation.

In addition to these prototypical instances, there a number of usage patterns that can become closer to the mental argument constructions (cf. Падучева, 2004:384, 391). Example 6.1-6 shows an extension with the verb *врубиться*, ranging from the spatial sense 'enter into, deepen,' to mental 'grasp.'

6.1-6 *И вообще не врубл-ю-сь*
And at.all NEG grasp-1S.FUT-RM
[з]*ачем ему такая нужна.* [о]*на ж ничего не знает и не умеет!*
[…]
I do not grasp it at all. Why would he need her? She does not know or is not able to do anything.
[1215, RNC, Разговор друзей (2006.11)]

Additional borderline cases are given in 6.1-7 and 6.1-8 with the verbs *теряться* 'be at loss' and *разбираться* 'sort out, grasp.'

6.1-7 *Однако, согласно данным исследований, проведённых на Западе,*
[…]
подавляющ-ее большинств-о мужчин теря-ют-ся при
Vast-NOM majority-NOM man.GEN.PL be.at.loss-3P.PRS-RM PR
вид-е женщин-ы в ярк-ом наряд-е.
sight-PREP woman-GEN PR striking-PREP dress-PREP
However, according to a study conducted in the West, the vast majority of men are at loss at the sight of woman in a striking dress.
[737, RNC, Обрати внимание // "Даша", №10", 2004]

6.1-8 *Вы, ваш-е величеств-о,* совершенн-о не
You.NOM your-NOM Majesty-NOM complete-ADV NEG
разбира-ете-сь в люд-ях.
grasp-2P.PRS-RM PR people-PREP
You, your Majesty, do not grasp people at all.
[1736, RNC, Сергей Седов. Доброе сердце Робина // "Мурзилка", №7," 2002]

Another type of extension is given in 6.1-9 with the reflexive verb *пользоваться* ‘use.’

6.1-9 *Какие мать или отец не хотят,*
[…]
чтобы их ребёнок с удовольствием посещал школу
[…]
и пользова-л-ся уважени-ем и любов-ью
and receive-PST.M-RM respect-INS and love-INS
не только педагог-ов,
NEG only teacher-GEN.PL
но и, что очень важно, сверстников?
[…]
What kind of mother or father would not want that their child would to happily attend school and receive respect and love not only from the teachers but importantly from their schoolmates?
[360, RNC, Алевтина Луговская. Если ребенок боится ходить в школу (2002)]

Probably, the typical instantiation of the verb *пользоваться* ‘use’ is in the Reflexive Engagement Construction illustrated in 6.1-10.

6.1-10 *Учён-ые* *полага-ют,* *что* *тукан-ы*
Scientist-NOM.PL think-3P.PRS that toucan-NOM.PL
пользу-ют-ся *сво-ими* *ярк-ими* *клюв-ами*
use-3P.PRS-RM own-INS.PL bright-INS.PL beak-INS.PL
как *сигнальн-ыми* *знак-ами.*
as signal-INS.PL sign-INS.PL
Scientists think that toucans are using their bright beaks as signaling signs.
[164, RNC, Туканы // "Мурзилка," №2," 1999]

In contrast, the reflexive verb *пользоваться* also extends to the mental argument constructions, (i.e., the 'receive' sense). The latter two instances demonstrate the diffusive nature of the reflexive verbs. Both the verb and the pattern can have low cue validity and it is through the argument slots that the profile is anchored in these cases.

6.2 Experiencer Construction

The Experiencer Construction covers 137 data points and 67 unique reflexive verbs in the RF model. The construction type is supported by two verb-specific constructions. First, the contribution of the reflexive verbs lacking a neighbor verb is fairly high with 23 unique reflexive verbs. Examples are *бояться* 'be afraid', *нуждаться* 'need,' and *надеяться* 'hope,' demonstrating a strong lexical basis of this particular argument construction type. Second, the reflexive verbs pertaining to the cross-paradigmatic relation form another cluster with 33 unique reflexive verbs, such as *увлекаться* 'be fascinated,' *волноваться* 'worry,' and *рассердиться* 'become angry.' A residual group is established with reflexive verbs that are either dissimilar such as *стесняться* 'be ashamed,' and *разобраться* 'grasp,' or intermediate such as *врубиться* 'grasp,' and *настороиться* 'become concerned.' The definition of the Experiencer Construction appears below.

Function: Profiles a mental relation between entities.
Form: Nominative subject and verb-specific constructions for the secondary slot.

The function of the Experiencer Construction captures the semantics associated with mental verbs consisting of the Experiencer and the Stimulus roles. The mental reflexive verbs display a strong verb-specific clustering in terms of patterns (cf. Янко-Триницкая, 1962:165). Examples 6.2-1–6.2-3 illustrate the divergence with verbs formed with the non-cross-paradigmatic verbs, ranging from the genitive and *на*$_{\text{acc}}$ to *в*$_{\text{prep}}$ patterns.

6.2-1 *Она* *холод-а* *не* *бо-ит-ся.*
She.NOM cold-GEN NEG fear-3S.PRS-RM
She is not afraid of cold.
[1396, RNC, RNC, Беседа психолога с ребенком // (2005.06)]

6.2-2 *Мы наде-ем-ся на участи-е*
We.NOM hope-1P.PRS-RM PR participation-ACC
специалист-ов по нейронн-ым сет-ям [...].
specialist-GEN.PL PR neural-DAT.PL net-DAT.PL
We look forward to the participation of specialists on neural networks.
[73, RNC, Конференция по когнитивной науке (2003)]

6.2-3 *Однако сейчас русск-ий народ* как
But now Russian-NOM people.NOM like
никогда нужда-ет-ся в настоящ-ей
never need-3S.PRS-RM PR genuine-PREP
правовой защит-е.
legal-PREP protection-PREP
But now the Russian people need like never before a genuine legal propection
[932, RNC, Андрей Андреев. БУДУЩЕЕ ПРИНАДЛЕЖИТ НАМ! // "Завтра," 2003.08.22]

Similar verb-specific constructions are present with the cross-paradigmatic verbs. Examples 6.2-4 and 6.2-5 illustrate the issue with *волноваться* 'worry,' and *испугаться* 'be scared.' Both verbs have causative neighbor verbs and are perceived to be semantically similar to their neighbors.

6.2-4 *Не дума-ю, что волнова-л-и-сь за нравственност-ь.*
NEG think-1S.PRS that worry-PST-PL-RM PR morality-ACC
I do not think that worried about morality.
[1852, RNC, Андрей Геласимов. Ты можешь (2001)]

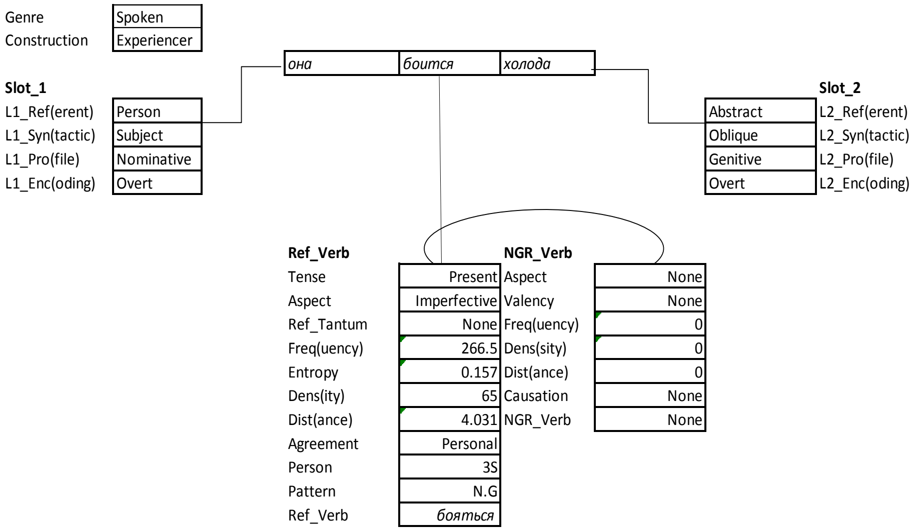
6.2-5 *Несмотря на неопровержимые улики,*
[...]
присяжн-ые испуга-л-и-сь ответственност-и и
juror-NOM.PL be.scared-PST-PL-RM responsibility-GEN and
не смогли лишить человека свободы.
[...]
Regardless of the overwhelming evidence, the jurors were scared of responsibility and could not deprive person's freedom.
[588, RNC, Убийцу не смогли опознать только присяжные // "Московский комсомолец в Нижнем Новгороде," 2004.07.30]

Certainly, one can always argue that the secondary slot is syntactically optional as these are intransitive verbs, as illustrated in 6.2-6.

6.2-6 *Рассерди-л-а-сь Лизавет-а.*
Become.angry-PST-F-RM NAME-NOM
Lizaveta became angry.
[932, RNC, Юрий Макаров. Про зайца // "Мурзилка," №12," 2001]

At the same, any account concerned with usage has to be able to incorporate these patterns in some manner. These examples serve to illustrate that the semantics of these verbs alone does not account for the patterns (cf. Faulhaber, 2011). It is worth pointing out that a degree of overlap in terms of patterns and semantics can be, nonetheless, brought forward. The verb *бояться* 'be afraid' aligns with *спугаться* 'be scared' through the Nominative-Genitive pattern. Additionally, *бояться* also aligns with *волноваться* 'worry' when used in the *за*$_{\text{acc}}$ pattern, (i.e., in the 'be afraid for' sense). Thus, a degree of overlap in terms of both semantics and patterns is available. Importantly, the overlap cannot be captured through derivation, as the gravitation is with a non-cross-paradigmatic reflexive verb.

The basic encoding of the Experiencer Construction is illustrated in Figure 6.2-1, based on Example 6.2-1.



Genre	Spoken
Construction	Experiencer

она	*боится*	*холода*

Slot_1	
L1_Ref(erent)	Person
L1_Syn(tactic)	Subject
L1_Pro(file)	Nominative
L1_Enc(oding)	Overt

	Slot_2
Abstract	L2_Ref(erent)
Oblique	L2_Syn(tactic)
Genitive	L2_Pro(file)
Overt	L2_Enc(oding)

Ref_Verb		NGR_Verb	
Tense	Present	Aspect	None
Aspect	Imperfective	Valency	None
Ref_Tantum	None	Freq(uency)	0
Freq(uency)	266.5	Dens(sity)	0
Entropy	0.157	Dist(ance)	0
Dens(ity)	65	Causation	None
Dist(ance)	4.031	NGR_Verb	None
Agreement	Personal		
Person	3S		
Pattern	N.G		
Ref_Verb	*бояться*		

Figure 6.2-1 Layered Structure of the canonical Experiencer Constructions.

The confusion matrix shows the predicted classes based on the RF model in Table 6.2-1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	**1**	**0**	**93**	**0**	**0**	**12**	**2**	**1**	**0**	**4**	**21**	**2**	**1**
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 6.2-1 Confusion matrix of the predicted constructions. The Experiencer Construction is given in bold.

The class-wise error of 0.321 is fairly high with this particular type. Furthermore, both the recall of 0.679 and precision of 0.775 indicate that the model has difficulties in identifying and classifying these instantiations, but the model appears to be better at classifying the instances compared to identifying them based on the input. This construction type also attracts others, primarily instances of the Reflexive Engagement Construction, as indicated in Figure 6.2-2.

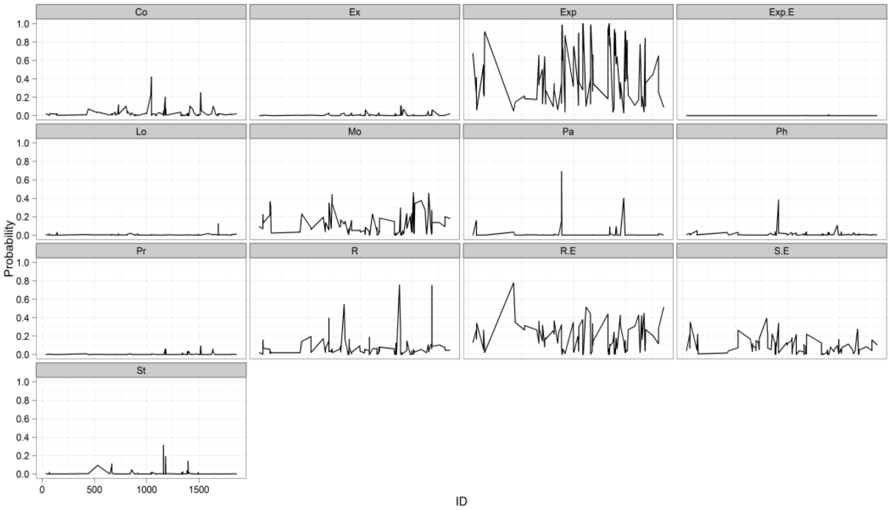


Figure 6.2-2 Faceted estimated class probability plot of the Experiencer Construction (Exp). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

As indicated in the confusion matrix, the strongest competition appears to be with the Reflexive Engagement and Spontaneous Event Construction. Although only six instances of the Spontaneous Event are predicted to be instances of the Experiencer Construction, the estimated class probabilities, nonetheless, show that the Spontaneous Event is also activated throughout the data points. One possible motivation behind the fluctuation is the configurational similarity. In general, the results suggest that the classification of the Experiencer Construction is only partly associated with the structural properties of the verbs when global comparison is made. The semantics of the reflexive verbs appears to contribute considerably to this argument construction.

6.3 Experiencer Extension Construction

The Experiencer Extension Construction covers 54 instances and 11 unique reflexive verbs, such as *хотеться* 'want,' *удаться* 'manage,' *оставаться* 'stay, remain,' and *понадобиться* 'necessary.' This type is similar to the Experiencer Construction but they differ in the encoding of the subject argument and the inclusion of the infinitive as part of the structure. This property is considered to constitute an augmentation in this study. Various infinitive patterns have not been systematically included in previous taxonomies on the Russian Reflexive Marker. These usually fall under some generalized impersonal type, such as the Impersonal Intensive meaning in Vinogradov (Виноградов, 1972:500). The definition of the Experiencer Extension follows.

Function: Profiles an augmented mental relation between entities.
Form: Dative subject and Infinitive secondary slot.

Generally, this pattern is supported by a relative small number of verbs in Russian (cf. Divjak & Janda, 2008). Additionally, these reflexive verbs appear to be detached from the cross-paradigmatic relation. Only the verbs *хотеться* 'want' and *мечтаться* 'dream' are perceived to be semantically similar to their neighbor verbs. Examples 6.3-1 and 6.3-2 illustrate the usage pattern of the construction.

6.3-1 *И* *мне* *чаще* *вс-его* *совсем* *не*
And I.DAT often.COMP all-GEN quite NEG
хоч-ет-ся *зна-ть,* […].
want-3S.PRS-RM know.INF
And Most often I do not quite want to know.
[1546, RNC, Евгений Гришковец. ОдноврЕмЕнно (2004)]

6.3-2 *Нам* *оста-ёт-ся* *лишь* *правильн-о*
We.DAT remain-3S.PRS-RM only correct-ADV
расстави-ть *их,* […].
arrange-INF they.ACC
We can only correctly arrange them.
[261, RNC, Александр Зайцев. Загадки эволюции: Краткая история глаза // "Знание — сила," 2003]

In contrast, Gerritsen (1990:152) considers that the verb *мечтаться* 'dream' is an instance of the so-called Medial-Passive. Certainly a possible analysis but, at least, in this specific configuration the profile of the verb aligns with other instances of this construction type as in 6.3-3.

6.3-3 [*м*]*не* *мечта-л-о-сь* *работа-ть* *в* *дружн-ой*
I.DAT dream-PST-N-RM work-INF PR friend-PREP
команд-е.
team-*PREP*
I dreamed of working in a friendly team.
[1191, RNC, Интервью с руководителем отдела (2006.11)]

These instances also display variation in terms of how strongly the infinitive is integrated with the reflexive verbs. For example, *хотеться* 'want' can also appear with the genitive case, as in 6.3-4. This is the only verb to display this variation in the sample where there is little doubt that the dative would not occupy the subject slot. Thus, it is included under the label Experiencer Extension. A larger number of verbs would be required to determine whether this pattern should be considered as a separate type.

6.3-4 Я была счастлива с ними работать,
[…]
но актёр-у всегдахоч-ет-ся чего-то нов-ого.
but actor-DAT alwayswant-3S.PRS-RM something new-GEN.
I was happy to work with them but the actor always wants something new.

[623, RNC, Кейт Уинслет: "Наше прошлое должно быть с нами" // "Экран и сцена," 2004.05.06]

Figure 6.3-1 illustrates the encoding of the Experiencer Extension based on Example 6.3-1.

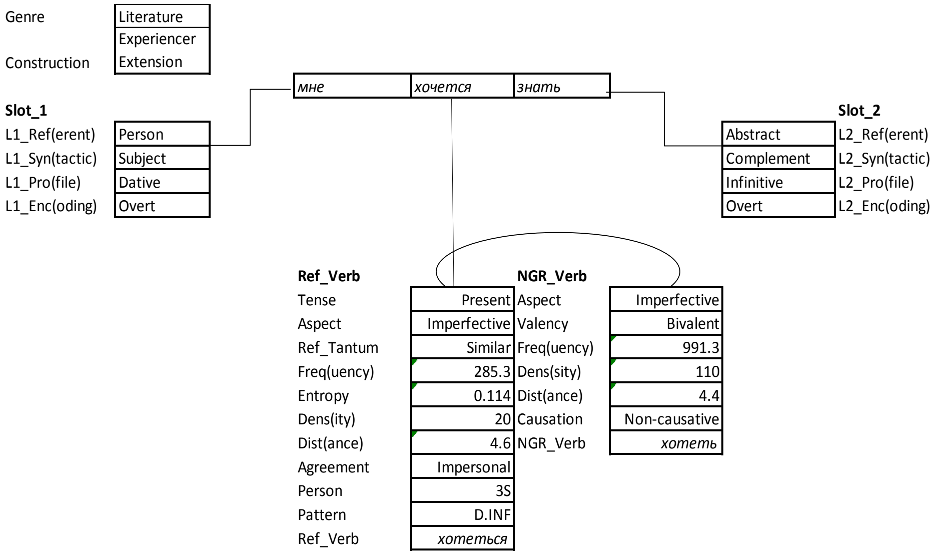


Figure 6.3-1 Layered structure of the canonical Experiencer Extension Construction.

The confusion matrix is given in Table 6.3-1. It is clear based on the confusion matrix that a classification task is fairly trivial with these three instances: class-wise error of 0, recall of 1, and precision of 1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	**0**	**0**	**0**	**54**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 6.3-1 Confusion matrix of the predicted constructions. The Experiencer Extension Construction is given in bold.

The estimated class probabilities point to the same results in Figure 6.3-2. The construction type is essentially without any competition between the other instances in the sample.

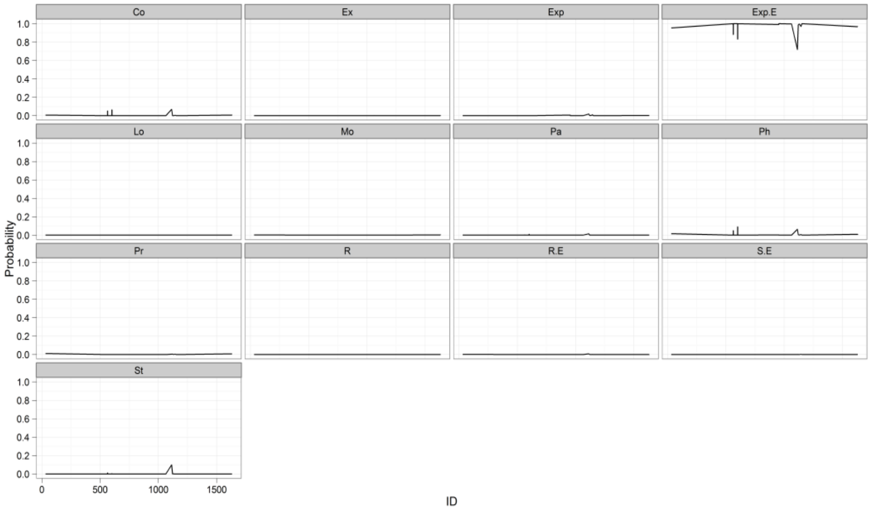


Figure 6.3-2 Faceted estimated class probability plot of the Experiencer Extension Construction (Exp.E). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The results suggest that the number of false positives in a syntactically tagged corpus should be minimal with a simple search: dative subject, reflexive verb and infinitive. In contrast, an argument is put forward later that the structure of this construction type can be regarded to be structurally salient (cf. Geeraerts et al., 1994). It can be considered that it is this property that contributes to the maintenance of small and deviating types in a network.

6.4 Stimulus Construction

The Stimulus Construction is another small construction type in the sample covering 65 instances, and 13 unique reflexive verbs. Such reflexive verbs are *нравиться* ‘please, like,’ *требоваться* ‘require,’ and *запомниться* ‘remember.’ From a semantic perspective, the verb *пригодиться* ‘come useful’ is further away from the canonical ones, but it nonetheless, follows the pattern. Generally, the argument constructions profiling the mental relation have a fairly high type frequency, but it seems that the lexical concentration of reflexive verbs is fairly sparse, covering 90 unique reflexive verbs in the sample. The basic type is established with the Experiencer Construction. The Experiencer Extension establishes an augmentation of the basic relation. The Stimulus Construction arises against this background by introducing a focus rather than an augmentation. The Stimulus occupies the subject slot with this type. The definition of the Stimulus Construction follows.

Function: Profiles a focus on the mental relation between entities.

Form: Subject and verb-specific constructions for the secondary slot.

This construction type also displays variation in terms of encoding the subject slot either with the Nominative or the Infinitive subject, leading to verb-specific constructions. Examples 6.4-1 and 6.4-2 illustrate the canonical pattern with the nominative subject.

6.4-1 *Сразу скажем: создать сад,*
[…]
котор-ому не требу-ет-ся уход,
which-DAT NEG require-3S.PRS-RM maintenance.NOM
невозможн-о!
impossible-ADV
We will say straight away: to create a garden which does not require maintenance, is impossible!
[875, RNC, Татьяна Ефимова. Скажи: легко! // "Сад своими руками," 2003.09.15]

6.4-2 *Шест-ь лет понадоби-л-о-сь*
Six-NOM year.GEN.PL be.necessary-PST-N-RM
руководств-у миров-ого футбол-а, […].
management-DAT world-GEN football-GEN
Six years was necessary for the management of the world football.
[538, RNC, Борис Зайцев. ФИФА выиграла судебную тяжбу за свои права в Интернете // "ИТАР-ТАСС," 2004.09.15]

The variation with the encoding of the subject slot is illustrated with the verb *нравиться* ‘please, like’ in 6.4-3 and 6.4-4, ranging from the canonical

Nominative subject to Infinitive as was outlined in Sections 3.2.5 and 3.2.6.

6.4-3 *Вот вам нрав-ят-ся ее песн-и?*
Well You.DAT like-3P.PRS-RM her song-NOM.PL
Well do you like her songs?
[1029, RNC, Беседа с Д. Арбениной, лидером группы "Ночные снайперы", "Школа злословия", канал "Культура" (2003.12.08)]

6.4-4 *Потому что нам нрави-л-о-сь лази-ть*
Therefore that we.DAT like-PST-N-RM climb-INF
на трет-ий этаж.
PR third-ACC floor.ACC
Because we liked to climb on the third floor.
[1816, RNC, Андрей Геласимов. Жанна (2001)]

Figure 6.4-1 gives the encoding of the Stimulus Construction based on Example 6.4-4.

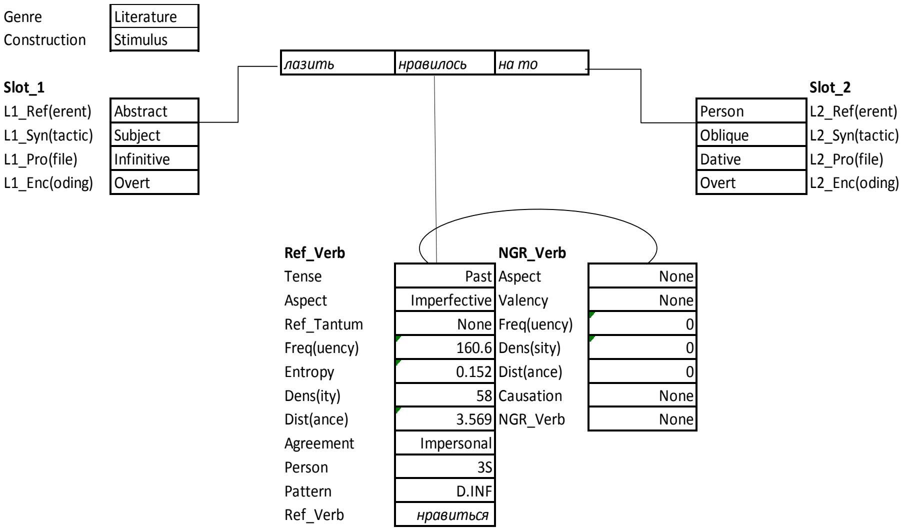


Figure 6.4-1 Layered Structure of the canonical Stimulus Construction.

Another important aspect of this construction type is that the gravitation between the paradigmatic and non-cross-paradigmatic verbs is present, as illustrated with the verb of perception *потребоваться* ‘require’ in 6.4-5. Traditionally, these instantiations are analyzed as involving reduced volitionality (Gerritsen, 1990).

6.4-5 *Для этого тебе потребу-ют-ся*
PR this.GEN you.DAT require-3P.PRS-RM
бигуд-и разн-ого размер-а.
curler-NOM.PL different-GEN size-GEN
For that you require curlers with different sizes.
[771, RNC, Укладки для весенних дней: да здравствуют перемены! // "Даша", №10," 2004].

Importantly, the cross-paradigmatic verbs align with the reflexive ones that do not have a neighbor verb, as in 6.4-6 (cf. Золотова, Г. А., 2005 [1973]:187-188). This kind of gravitation once again imposes difficulties if the derivation is the primary motivation behind the various reflexive construction types (cf. Gerritsen, 1990:151).

6.4-6 *"Сн-ит-ся мне вс-ё это*
Dream-3S.PRS-RM I.DAT all-NOM this.NOM
или на сам-ом дел-е?" подума-л Медвежонок.
or PR actual-PREP fact-PREP think-PST.M NAME.NOM
Did I dream all of this in fact? – Bear thought.
[1612, RNC, Сергей Козлов. Новогодняя сказка // "Мурзилка," №1", 2003]

The confusion matrix based on the RF model is given in Table 6.4-1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**65**

Table 6.4-1 Confusion matrix of the predicted constructions. The Stimulus Construction is given in bold.

The Stimulus Construction appears to display similar behavior to the Experiencer Extension Construction. The class-wise error is 0, and recall is 1 although the precision of 0.97 indicates that the Stimulus Construction can attract other types. The misclassified instantiation of the Spontaneous Event was already illustrated in Section 5.4.

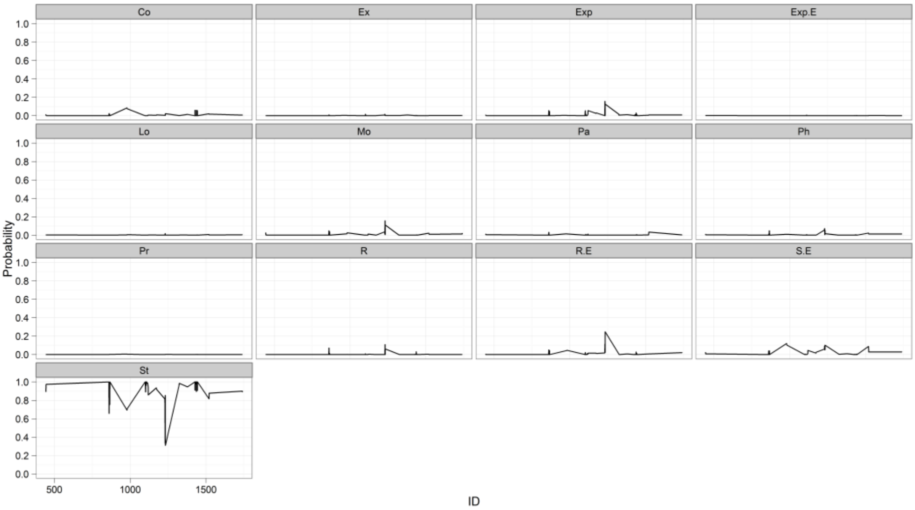


Figure 6.4-2 Faceted estimated class probability plot of the Stimulus Construction (St). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The estimated class probabilities show that the majority of the instantiations are well separated. However, a degree of fluctuation present. Example 6.4-7 is estimated to instantiate the Stimulus Construction but the estimated class probability is low (0.31), competing with such types as the Reflexive Engagement (0.248), the Experiencer (0.126), and the Spontaneous Event (0.101).

6.4-7 *Голосова-л-а* *за* *Путин-а* на фон-е
Vote-PST-F PR NAME-ACC PR background-PREP
стар-ого *он* *смотре-л-ся* *намного* *лучше.*
old-GEN he.NOM look-PST.M-RM much good.COMP
She voted for Putin. He looked much better against the background of the old.
[1498, RNC, Беседа с социологом на общественно-политические темы, Самара // ФОМ (2004.02.17)]

Globally, the instantiations of the Stimulus Construction appear to be well-separated. The type is primarily a configurational one and mental verbs can appear in different construction types depending on the imposed perspective. For instance, the verb *понадобиться* ‘necessary’ appears in the Experiencer Extension with the Dative Subject and in the Stimulus Construction with the Nominative Subject. Although the variation is highly verb-specific, it gives rise to the argument construction.

7 Spatial Constructions

In Cognitive Linguistics, the central claim is that our concepts are structured internally and relative to one another and it is this structuring that allows us to reason and comprehend. In the strongest position, the meaningfulness of these structures arises due to our bodily experience (Johnson, 1987; Lakoff, 1987:266-267). Although the universal aspect of the body has been challenged by Levinson's typological studies on spatial categorization (see especially Levinson & Wilkins, 2006), there are certain fundamental properties which are shared, namely the frame of reference we use to locate ourselves, others, and objects. Levinson´s studies have shown that the cross-linguistic variation in spatial conceptualization is wider than expected.[121] However, what languages have inherently in common is that regardless of the system employed in a given language, the dominant spatial conceptualization imposes its structure virtually on every spatial scene. Languages utilize a closed-class set of markers for any given spatial scene, and speakers of a given language community need to select among them when conceptualizing a certain spatial scene. Due to their closed-class nature, these markers carry their own configurational structure.

A basic spatial configuration consists of portioning the scene into parts. One element is conceptualized as being the primary object, which is further characterized relative to some other secondary object. This mode of selection displays an asymmetric conceptualization where the primary object is viewed as dispositional relative to some other object, yielding a distinction between three basic configurational patterns in terms of the disposition: the location of the primary object when stationary, its path when in motion, and its orientation during either of these states. (cf. Talmy, 2000a:181-183; 2000b:25-26). In Talmy´s terminology, the primary object is labeled as Figure and the secondary object as Ground. Whereas in Cognitive Grammar, they are called Trajector and Landmark, respectively (Langacker, 1987:231-232, 237).[122] Both of these concepts rest on the principles laid out in Gestalt psychology (Koffka, 1935). For linguistic purposes, the Figure is associated with features such as the conceptually more dependent entity in the event, salient and its orientation or motion is perceived to be more relevant.[123] The Ground is the opposite of

[121] Levinson´s typological studies depart from the canon in a sense that the divergence of encoding spatial scene in languages are broader than the model based on spatial domain alone would predict (for a detailed discussion see Levinson & Meira, 2003).

[122] The primacy of the spatial semantics figures prominently in Cognitive Linguistics. For example, Janda (2008a) anchors aspectual distinctions to spatial semantics, specifically to the verbs of motion in Russian. Similar argumentation is laid out in Croft (1991:192-194).

[123] Because the concept of Figure typically aligns with the concept of subject, Talmy's definition of the Figure as the more dependent entity in the event goes against most definitions of subjects as the more independent entity in the event as in Dowty

Figure, and its primary function is that of reference, that is, it is used to locate the Figure in a spatial configuration. (Talmy, 2000a:182-187, 211-212).[124] Thus, this chapter covers the basic spatial configurations marked with the Russian Reflexive Marker. Section 7.1 lays out the configurations associated with motion, whereas the spatial location and its extension to existential semantics are discussed in Sections 7.3.1 and 7.3.2.

7.1 Definitions of the Motion Construction

Traditionally, motion is defined relative to a moving entity. At a given time an entity occupies a location L_1 and at a following moment in time the entity occupies another location L_2. In this description, L_1 is defined as the starting point of the motion event and L_2 as the end point (Cardini, 2008; Fillmore, 1983; Langacker, 2002:152-153, 155-156; Lyons, 1977; Miller, G. & Johnson-Laird, 1976). In the Russian tradition, the primary semantic component in defining the semantics of motion relative to semantic classes is velocity (Падучева, 1999:87-88; Розина, 1999). Another set of important components to differentiate verbs of motion are the manner and the path (Beavers, Levin, Rappaport Hovav & Tham, 2010; Slobin, 2006; Talmy, 1975; 2000b).[125] Examples 7.1-1–7.1-3 illustrate the primary properties of the motion event in general.

7.1-1 *Остальн-ые* *носи-л-и-сь* *по* *вс-ему* *двор-у*.
Rest-NOM.PL rush-PSR-PL-RM PR whole-DAT yard-DAT
The rest rushed around the whole yard.
[1841, RNC. Андрей Геласимов. Нежный возраст (2001)]

7.1-2 *А* *слонёнок* *Эле-Фантик* слоня-л-ся
And baby.elephat.NOM NAME.NOM wander-PST.M-RM
в *основн-ом* *по* *дик-ому* *пляж-у*, […].
PR main-PREP PR wild-DAT beach-DAT
The baby elephant, Ele-Fantik, wandered mainly on the wild beach.
[1575, RNC, Александр Дорофеев. Эле-Фантик // "Мурзилка", №1 5," 2003]

(1991:572), Keenan (1976:312-320), Langacker (1987:288-290; 1991:282-283), and Primus (1999:266-268).

[124] The definition of the Figure and the Ground also connects to semantic roles, especially in the Localist tradition where semantic roles are defined relative spatial semantics (Gruber, 1965; Jackendoff, 1990).

[125] Generally, *идти* 'go by foot' is the neutral verb of motion in contemporary Russian, (i.e., the manner component is bleached, especially in spoken Russian, for example *машина идет* 'the car goes (by foot)' instead of *едет* 'goes (by transport), *корабль идет* 'the ship goes' instead of *плывет* 'swims,' *снег идет* 'it is snowing, [lit. the snow goes]' instead of *падает* 'falls down') (Майсак & Рахилина, 1999:61).

7.1-3	*Робин*	*между*	*тем*	*уже*	*стремительн-о*
	NAME.NOM	mean	while	already	quick-ADV

мча-л-ся	*к*	*замк-у.*
race-PST.M-RM	PR	castle-DAT

Meanwhile, Robin had quickly raced to the castle.
[1716, RNC, Сергей Седов. Доброе сердце Робина // "Мурзилка", №7," 2002]

In contrast, the concept of motion does not appear in the diathesis tradition. The category of Autocausative covers the reflexive verbs of motion if they happen to have a semantically similar neighbor verb and the subject arguments are the same across the derivation, the subject-oriented type. In the object-oriented type, these verbs would appear under the label Decausative (cf. Geniušienė, 1987:86-87; Князев, 2007:275-278).[126] However, the verbs *носиться* 'rush,' *слоняться* 'wander,' and *мчаться* 'race, rush' serve to demonstrate the gravitation of this semantic domain rather than derivations. The verb *носиться* has a causative neighbor, *носить* 'carry' but it is perceived semantically dissimilar whereas *слоняться* 'wander' has a semantically similar intransitive neighbor, *слонять* 'wander.' Thus, the neighbor verb is inherently non-causative.[127]

On the other hand, *мчаться* 'rush, race' displays all the properties of a full cross-paradigmatic relation, a semantically similar causative neighbor verb *мчать* 'rush, race.' In terms of surface generalizations, it is argued that these and similar verbs gravitate towards a specific argument construction type in Russian albeit through different pathways labeled as the Motion Construction similar to Barðdal (2008:68) who also posits a single abstract construction type with subtypes. [128]

From a constructionist perspective, Examples 7.1-1 and 7.1-2 share the same profile. Both have the Nominative subject and a secondary slot with the *по*$_{\text{dat}}$ prepositional phrase, but differ in the manner of velocity and how the path of the motion is profiled. Thus, they follow the general property of argument

[126] In the early Russian tradition, the Middle Reflexive Meaning covers expressions of change of state as well as change in physical position subsuming, at least partly, reflexive verbs of motion (Виноградов, 1972:496).

[127] Another difficulty for strict derivation is the type of orientation. The Autocausative is typically considered to be subject-oriented, (i.e., the subject arguments are the same between the base form and the derived form). However, most inanimate entities can be construed as moving by themselves in Russian as is noted by Paducheva (2003:182). With inanimate or abstract subject arguments the potential derivational pathway is less than clear-cut and they could also be interpreted as decausative reflexive verbs (cf. Князев, 2007:277-278).

[128] The Autocausative type could simply be defined relative to agentivity and defined so broadly that also intransitive motion verbs are included as it is done in Knyazev (Князев, 2007:276).

constructions that they do not typically encode manner (cf. Goldberg, 2006:106; Tomasello, 2003:126). Consequently, the manner of velocity is a verb-specific property, and the primary concern is the relation and encoding of the argument slots. In the Russian tradition, the prepositions involving directionality are analyzed as a series of spatial localization features: 'inside,' 'on the surface' and 'proximity:' 1) *в*$_{\text{acc}}$ *в*$_{\text{prep}}$ and *из*$_{\text{gen}}$ 2) *на*$_{\text{acc}}$ *на*$_{\text{loc}}$ and *с*$_{\text{gen}}$ and 3) *к*$_{\text{dat}}$ *у*$_{\text{gen}}$ and *от*$_{\text{gen}}$ (Князев, 1999:183). Thus, a degree of cue-validity originates from the encoding of the secondary slot.

In typological studies, the motion verbs are typically divided into two types: translational and nontranslational. These types have become a fairly standard descriptive tool in functionally oriented studies (Bostoen & Nzang-Bie, 2010; Kemmer, 1993; Talmy, 2000b).[129] The Translational Construction pertains to an event type which necessarily contains a trajectory along which the motion is taking place. Thus, the figure changes its location. The previously given Examples in 7.1-1–7.1-3 profile a motion event with trajectory along which the motion is taking place. In terms of frequency and semantics, the canonical instantiation of the translational motion type is *двигаться* 'move,' as given in 7.1-1. Additionally, the profiling of the full trajectory is illustrated with the verb *уткнуться* 'nuzzle' in 7.1-5.

7.1-4 *Молодые люди довольно охотно идут к нам / они работают год /*
[…]
два / приобретают какую-то капитализацию
[…]
и дальше уже движ-ут-ся в бизнес.
and far.COMP already move-3P.PRS-RM PR business.ACC
Young people are quite happy to come to us. They work a year or two and they acquire some capitalization and then move onward to business.
[1451, RNC, Беседа на радио о государственной службе // (2004.11.15)]

7.1-5 *Заяц с налёт-у нос-ом в валенок*
Rabbit.NOM PR air-GEN nose-INS PR felt.boot.ACC
Дед-а Мороз-а уткну-л-ся.
Santa-GEN Claus-GEN nuzzle-PST.M-RM
The rabbit nuzzled his nose into Santa Claus's felt boot from the air.
[1940, RNC. Юрий Макаров. Про зайца // "Мурзилка," №12," 2001]

The nontranslational type pertains to events where the motion along the trajectory is not included in the profile. Thus, the figure retains its relative position.[130] Examples 7.1-6 and 7.1-7 illustrate typical instantiations with the verbs *колебаться* 'oscillate, fluctuate', *качаться* 'swing, rock, dance.'

[129] Talmy (2000b:35) uses the labels Translational and Self-Contained Motion. The label nontranslational motion is adopted from Kemmer (1993).

[130] Certain motion verbs can also extend to the Mental Domain. For example, *колебаться* can also be used to profile experiencing doubt (cf. Розина, 1999).

7.1-6 […] *и* *колебл-ю-тся* *песчинк-и,* […].
and oscillate-3P.PRS-RM sand.grain-NOM.PL
And the grains of sand are oscillating.
[1978, RNC, Олег Павлов. Карагандинские девятины, или Повесть последних дней // ""Октябрь", №8", 2001]

7.1-7 […] *где* *на* *колюч-их* ветк-ах
where PR thorny-PREP.PL branch-PREP.PL
звёзд-ы *кача-ют-ся,* […].
star-NOM.PL dance-3P.PRS-RM
Where stars are dancing on thorny branches.
[1739, RNC, С Новым годом! // "Мурзилка", №12", 2002]

Additionally, Kemmer (1993) separates verbs that profile a change in the body posture, for example *ложиться* 'lie down', *садиться* 'sit down', and *подниматься* 'rise,' as in 7.1-8 and 7.1-9.[131]

7.1-8 [*п*]*редставители руководства выходят по одному из-за кулис и*
[…]
сад-ят-ся *за* *стол* *посередине* сцен-ы.
sit.down-3P.PRS-RM PR table.ACC middle scene-GEN
The representatives of the management leave one at a time behind the scene and sit down at the table in the middle of the scene.
[1016, RNC, Встреча футбольного клуба "Локомотив" с болельщиками, Москва (2004.02.21)]

7.1-9 *В* *ярост-и* *поднима-ет-ся* *и* *ид-ёт* *к* *окн-у.*
PR rage-PREP rise-3S.PRS-RM and walk-3S.PRS PR window-DAT
He rises in rage and walks towards the window.
[1526, RNC, Ordinamenti // "Экран и сцена", 2004.05.06]

Considering that Kemmer (1993:56) groups such verbs as *нагнуться* 'bend down' under the label nontranslational motion, exemplified in 7.1-10, the category of change in the body posture can be considered to constitute a set of verb-specific constructions.

7.1-10 *Елен-а* *Андреевн-а* *нагн-ёт-ся* *над* *стол-ом.*
NAME-NOM NAME-NOM bend-3S.FUT-RM PR table-INS
Elena Adreevna bends down under the table.
[628, RNC. Легкое дыхание // "Экран и сцена", 2004.05.06]

From a semantic perspective, these verbs once again display the gravitation rather than a semantic derivation. For example *ложиться* 'lie down' does not have a neighbor verb contrasting *садиться* 'sit down' which is perceived as semantically intermediate to its neighbor verb.[132]

[131] Dixon (1991:94-95) classifies such English verbs as *sit* and *lie* belonging to the rest type of motion events.

[132] Kolomackij (Коломацкий, 2009:37) considers that the verbs *садиться* 'sit down'

The canonical instances are well-separated, but certain reflexive verbs are fairly bleached. The shared commonality for these verbs is the neutrality of velocity focusing on the offset of the motion, for example *остановиться* 'stop' and *вернуться* 'return,' illustrated in 7.1-11 (cf. Падучева, 2004:46).

7.1-11 *По* *дорог-е* *пят-ь* *раз* *останавлива-л-ся* *отдыха-л.*
PR road-DAT five-NOM time.GEN-PL stop-PST.M-RM rest-PST.M
On the road he stopped five times to rest.
[1990, RNC. Олег Тихомиров. Про козла Тихомира // "Мурзилка", №2", 2001]

These verbs typically include some form of trajectory. However, the verb itself implies a motion event along a path prior to the rest state. From this perspective, these expressions can be considered pertaining to the category of Motion Construction. In contrast, the verb *задержаться* 'stay, delay' illustrates another bleached type in 7.1-12.[133]

7.1-12 *В* *тот* *ден-ь* *я* *не* *задержа-л-а-сь*
PR that.ACC day-ACC I.NOM NEG stay-PST-F-R
на *работ-е,*
PR work-PREP
быстренько оделась и убежала домой.
[…]
That day I did not stay at work, I changed quickly and ran home.
[785, RNC, Я желанна. Разве это стыдно? // "Даша", №10", 2004]

The verb lacks a clear trajectory although it displays similar propensity towards the rest state as the previously mentioned reflexive verbs. Thus, it can be considered to instantiate the category of Motion Construction.

Certain verbs of motion are closer to the typical Semantic Reflexive Construction, (e.g., *мыться* 'wash oneself' and *одеться* 'dress oneself'). Certainly washing oneself and dressing oneself are typically localized in space, the spatial semantics is integrated to a lesser degree with the Semantic Reflexive compared to such instances as *запутаться* 'tangle,' in 7.1-13. At the same time, these instances are also problematic for derivational accounts, especially for the Autocausative type, because the semantics of the subject is closer to non-volitional and uncontrolled event types (cf. Князев, 2007:278).

and *садить* 'seat, put' are not semantically related any more in contemporary Russian. Thus, the label intermediate appears to be most appropriate for this particular cross-paradigmatic relation in terms of semantics.

[133] Paducheva (Падучева, 2004:278) considers that the neighbor verb, *задержать* 'keep, delay' profiles a state within the category of verbs of emotion.

7.1-13 *Я* *запута-л-ся* *в* *поводк-ах,*[…].
I.NOM tangle.PST.M-RM PR harness-PREP.PL
I tangled in harnesses.
[1939, RNC, Андрей Геласимов. Нежный возраст (2001)]

Another complication when dealing with usage patterns is the degree of motion. Examples 7.1-14 illustrates the issue with the verbs *колебаться* ‘oscillate, fluctuate.’

7.1-14 […] *числ-о* *котор-ых,* *по* *различн-ым*
number-NOM which-GEN.PL PR different-DAT.PL
оценк-ам, *колебл-ет-ся* *в* *диапазон-е*
value-DAT.PL hover-3S.PRS-RM PR range-PREP
от *одн-ого* *до* *нескольк-их* *миллион-ов.*
PR one-GEN PR several-GEN.PL million-GEN.PL
The number of which, depending on the values, hovers in the range of one million to several millions.
[245, RNC. Янис Астафьев. Кто будет работать в России в 2015 году? // "Отечественные записки", 2003]

The profile includes prepositions *от*$_{\text{gen}}$ ‘from’ and *до*$_{\text{gen}}$ ‘to’ which are used to highlight the boundaries. At the same time, the definition of the Motion Construction crucially depends on the whole profile. The verb *рваться* can be used to illustrate the issue at hand. In 7.1-15, *рваться* profiles the translational motion, (i.e., the ‘sweep along’ sense).

7.1-15 *Они* *рв-ут-ся* *в* *разн-ые*
They.NOM sweep.along.3P.PRS-RM PR different-ACC.PL
сторон-ы *как* *сумасшедш-ие.*
direction-ACC.PL like madman-NOM.PL
They sweep along to different directions like madmen.
[1838, RNC, Андрей Геласимов. Нежный возраст (2001)]

In contrast, Example 7.1-16 deviates from the motion type. The prepositional phrase indicates a setting where the event is unfolding (cf. Langacker, 2002:230-232; 2009:118).

7.1-16 *А в начале третьего финального акта*
[…]
на *тёмн-ом* *экран-е* *рв-ут-ся*
PR dark-PREP screen-PREP break.down-3P.PRS-RM
молнии-и.
lightning-NOM.PL
In the beginning of the third and final act, lightning bolts are breaking down on the dark screen.
[650, RNC, Спасительная эстафета игры // "Экран и сцена", 2004.05.06]

Consequently, the instantiation is pertains to the Spontaneous Event Construction, (i.e., the 'break down' sense). Similar fluctuation is connected to certain verbs such as *разбегаться* 'scatter' as given in 7.1-17.

7.1-17 *Что* *ни* *утр-о* *пуглив-о* *разбега-л-и-сь*
That NEG morning-NOM timid-ADV scatter-PST-PL-RM
облак-а,
cloud-NOM.PL
повылезшие за ночь как из щелей на чёрствые звёздные крошки.
[…]
That the clouds timidly scattered that had crawled out as if from an
opening to the callous stellars crumbs overnight.
[1945, RNC, Олег Павлов. Карагандинские девятины, или Повесть
последних дней // ""Октябрь", №8", 2001]

The configuration is classified as an instance of the Reciprocal Construction, the collective subtype in Section 5.9, but it also partly contains the semantics of motion by including movement along a trajectory and velocity. This may be one of the contributing factors for the semantic gravitation of the Reflexive Marker in general. The Reciprocal Construction contains strong exemplars such as *бороться* 'fight,' *поссориться* 'argue' and *видеться* 'see each other,' but certain verbs such as *разбегаться* 'scatter' subsume simultaneously several properties across different argument constructions.

7.2 Motion Construction

The Motion Construction contains 212 data points and covers 140 unique reflexive verbs. This is one of the few constructions that display a substantial proportion of unique reflexive verbs relative to the type frequency of the construction based on the sample and the imposed classification. Based on frequency, the following reflexive verbs can be considered to be the canonical instantiations forming the most prominent verb-specific constructions: 1) the rest type, *вернуться*$_{\text{perf}}$ 'return,' and *остановиться* 'stop,' 2) translational motion, *двигаться* 'move,' *отправиться* 'leave,' and *собираться* 'be going to,' 3) the change in body posture, *садиться* 'sit down,' and 4) nontranslational motion, *колебаться* 'oscillate, fluctuate.'[134] The first set has semantically similar but non-causative neighbor verbs: *вернуть* 'return, give back,' *остановить* 'stop.' The second set also has semantically similar neighbors: *двигать* 'move,' *отправить* 'send,' and *собирать* 'collect, assemble.' The latter neighbor verb is the only non-causative of the three. In contrast, the cross-paradigmatic relation between *садиться* 'sit down' ~ *садить* 'seat, put' is perceived to be semantically intermediate. Finally, the fourth group seems to be supported by a semantically similar causative neighbor verb *колебать* 'oscillate, fluctuate.

[134] The sense 'be going to' or 'intend' of *собираться* was separated based on the form. The Nominative-Infinitive pattern was classified as the Phase Construction and the combination with a prepositional phase, 'where' as the Motion Construction.

The center of the Motion Construction is primarily supported with the cross-paradigmatic relation and semantic similarity simultaneously supporting such infrequent reflexive verbs not confined to the cross-paradigmatic relation, (e.g., *выситься* 'arise', *потусоваться* 'mingle' and *промыкаться* 'linger on'). The definition of the Motion Construction follows.

Function: Profiles a spatial movement of an entity.

Form: Nominative Subject and a construal-specific secondary slot.

The definition of the Motion Construction assumes that the manner component is underspecified and it is part of the verb-specific constructions. The encoding of the secondary slot is primarily determined by the mode of construing the movement of an entity as a whole rather than being an inherent property of the reflexive verb. At the same time, a degree of specificity is present with the verbs of motion primarily related to prefixation. Certain prefixes form reduplication, such as *добраться* 'reach,' where the secondary slot is encoded with the *до*$_{\text{gen}}$ preposition (cf. Князев, 1999:184-185). On one hand, these and similar instances pertain to the prefix-constructions and, on the other, to the category of verb-specific constructions. Example 7.2-1 illustrates an instantiation of the Motion Construction and the encoding is given in Figure 7.2-1.

7.2-1 *Для проверки справедливости данной гипотезы мы пошли по пути*

[…]

распределения массы центрального тела по всему объёму

[…]

в	*котор-ом*	*движ-ут-ся*	*друг-ие*
PR	which-PREP	move-3P.PRS-RM	other-NOM.PL

небесн-ые	*тел-а.*
celestial-NOM.PL	object-NOM.PL

In order to verify the validity of this hypothesis, we opted for the mass distribution of the central object in the whole volume in which other celestial objects are moving.

[319, RNC, В.В. Ахияров. Гравитация в Солнечной системе // "Геоинформатика", 2002.03.20]

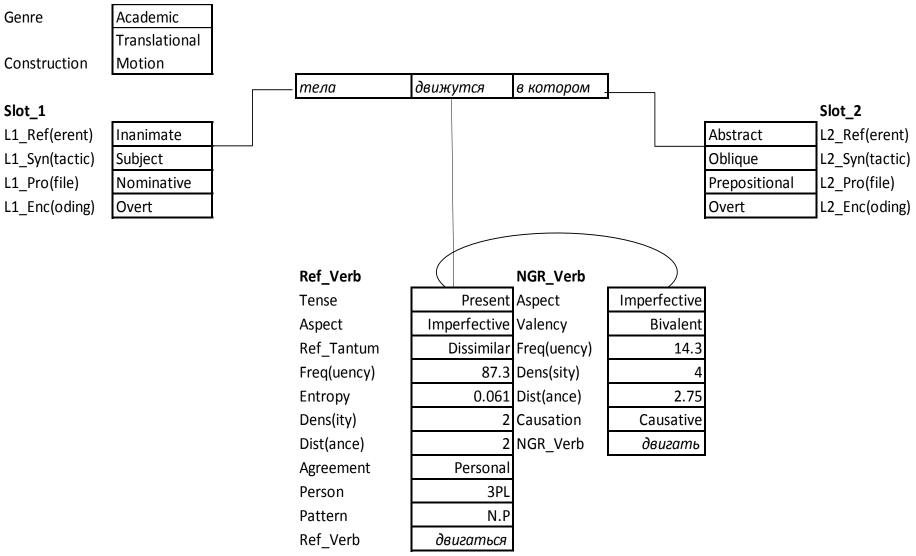


Figure 7.2-1 Layered structure of the canonical Motion Construction.

The confusion matrix is given in Table 7.2-1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	**0**	**1**	**6**	**0**	**0**	**141**	**0**	**0**	**0**	**9**	**23**	**32**	**0**
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 7.2-1 Confusion matrix of the predicted constructions. The Motion Construction is given in bold.

The class-wise error of 0.335 is fairly high with the instantiations of the Motion. The recall of 0.665 is better at the identification of the instantiations compared to the classification with the precision of 0.61.

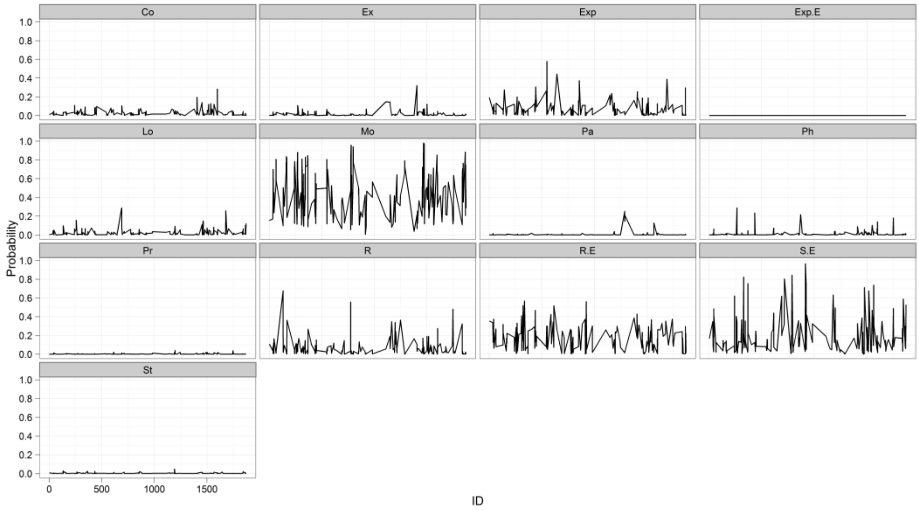


Figure 7.2-2 Faceted estimated class probability plot of the Motion Construction (Mo). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The structural properties of the reflexive verbs create a situation where the Reflexive Engagement and Spontaneous Event are activated based on the estimated class probabilities, as demonstrated in Figure 7.2-2. Particularly, the competition with the Spontaneous Event and the Translational Motion is evident with inanimate and abstract subject referents. Typically, transports are construed as moving by their own force but they become based on their structural properties closer to the Spontaneous Event, as in 7.2-2 with the verb *тронуться* 'set out.' This is reflected in the estimated class probabilities: Translational Motion with the estimated probability of 0.204 and the Spontaneous Event with the probability of 0.313. A more complex instantiation is given in 7.2-3 with the verb *передвинуться* 'move, shift.' Inherently, the verb certainly contains the canonical properties of a verb of motion, (i.e., a path and movement along a trajectory), but the construal of the situation may bring the Translational Motion and the Spontaneous Event in close proximity. Based on the class probabilities, Example 7.2-3 is estimated to be a poor instantiation of the Translational Motion (0.058) relative to the Spontaneous Event (0.872).

7.2-2 *И* *когда* *эта* *флотили-я* *трону-л-а-сь*
And when this.NOM fleet-NOM set.out-PST-F-RM
в *пут-ь*
PR journey-ACC
And when this fleet set out on a journey.
[369, RNC, Олег Тихомиров. Подвиг Магеллана // "Мурзилка", №1", 2002]

7.2-3 *Осенн-ий* *сезон* *самоубийств*
Autumn-NOM season.NOM suicide.GEN.PL

передвину-л-ся *на* *декабр-ь.*
move-PST.M-RM PR December-ACC
The autumn season of suicides moved to December.
[515, RNC, Осенний сезон самоубийств передвинулся на декабрь // "Московский комсомолец" в Саранске", 2004.12.23]

The canonical instantiations appear to emerge from the data based on the structural properties of the verbs alone. However, even a slight deviation in the profile appears to create a fluctuation in the estimated class probabilities indicating that this particular type is strongly connected to the semantics of the situation.

7.3 Definitions of the Location and Existence Constructions

There is a tight connection between the Location and Existence Constructions. Additionally, possessive constructions are considered to be related to the previously mentioned types (Afonso, 2008; Freeze, 1992; 2001; Jespersen, 1924; Langacker, 2009; Leinonen, 1984). However, most of the theoretical frameworks on this matter are based on the relation between locations and the copula verb *быть* 'be' (cf. Babby & Comrie, 1980; Paducheva, 2008; Арутюнова, 1976). In terms of the Russian Reflexive Marker, the verbs *находиться* 'be located, be,' *располагаться* 'be situated,' and *уместиться* 'fit' being inherently stative, can be considered constituting the canonical instantiations supporting the Location Construction (Золотова, Г. А., 2005 [1973]:231-232), as in 7.3-1.

7.3-1 [*а*] *пят-ая* *част-ь* *жител-ей*
and fifth-NOM part-NOM inhabitant-GEN.PL
наход-ят-ся *за* *черт-ой* *бедност-и.*
be-3P.PRS-RM PR line-INS poverty-GEN
And a fifth of the inhabitans are below the poverty line.
[989, RNC, Бедные беднеют, богатые богатеют // "Московский комсомолец в Сыктывкаре", 2003.08.06]

In terms of the previously proposed taxonomies, the Location Construction once again meshes with the concept of Medial. For instance, Gerritsen (1990:51, 53-54) (1990: 51, 53-54) posits a subcategory labeled as the processual decausative Vsja. The label subsumes, for instance, the verb *сохраняться* 'preserve, stay,' as illustrated in 7.3-2.

7.3-2 […] *вс-я* *жизн-ь* *в* *Одесс-е* *в* *конц-е*
whole-NOM life-NOM PR NAME-PREP PR end-PREP
1920 *год-а* *и* *в* *1921* *год-у*
1920 year-GEN and PR 1921 year-PREP
сохрани-л-а-сь *в* *мо-ей* *памят-и* […].
stay-PST-F-RM PR my-PREP memory-PREP
the whole life in Odessa at the end of the of 1920 and in the year of 1921 stayed in my memory.
[Gerritsen (1990:54, Example (51, PR), (Паустовский 36))][135]

Although Gerritsen acknowledges that the interpretation of the oblique might be construed as a place (i.e., *в моей памяти* 'in my memory'), the difference lies in the derivational relation according to her derivational account. If the oblique argument can be construed as the Agent of the neighbor verb, the instantiation belongs to the category of Medial (Gerritsen, 1990:51, 53-54). Example 7.3-4 demonstrates construal of the spatial location in the strict sense.

7.3-3 [*ч*]*то* *вс-е* *эти* *вещ-и*
that all-NOM.PL this.NOM.PL thing-NOM.PL
сохрани-л-и-сь *в* *мир-е.*
preserve-PST-PL-RM PR world-PREP
That all these things were preserved in the world.
[676, RNC, Темные силы против Масленицы // "Народное творчество", 2004.02.16]

It is difficult to maintain the position that agentivity would be a crucial factor in this case, especially when the verb *сохраняться* is grouped with *находиться* 'be located, be' in her Medial category (cf. Gerritsen, 1990:37).

Additionally, certain verbs, which profile motion, can appear in the Location Construction if an inanimate entity is used in the subject slot, as in 7.3-4 and 7.3-5 with the reflexive verb *садиться* 'sit down' and *таиться* 'hide, lie,'

7.3-4 *На* *т-ом* *мест-е,* *где с* *ади-л-а-сь* *тарелк-а,*
PR that-PREP place-PREP where sit-PST-F-RM saucer-NOM
лежа-л *ровн-ый,* *нетронут-ый* *снег.*
lay-PST.M smooth-NOM untouched-NOM snow.NOM
On that place where the saucer was, laid a smooth and untouched snow.
[1620, RNC [Сергей Козлов. Новогодняя сказка // "Мурзилка", №1", 2003][136]

[135] The glossing and the translation were added to Example by the author.

[136] Example describes how a UFO landed on Earth.

7.3-5 *Однако в разнообрази-и та-ит-ся*
However PR diversity-PREP lie-3S.PRS-RM
и некотор-ая опасност-ь, […].
also certain-NOM danger-NOM
However, certain danger also lies in diversity.
[712, RNC, Вероника Стрельникова. Опять акробатика, милый? // "Даша", №10", 2004]

The distinction between the Location and the Existence Constructions is a widely discussed and debated topic. In most studies on existential constructions, the semantic weight is placed on the subject argument and its features. The subject of the Existence Construction is characterized as being indefinite and non-referential (Abbott, 1997; Paducheva, 2008:151). Additionally, Paducheva (2008:149) argues that the profiled location in the Existential Construction could be characterized as generalized being or coming into being. Typically, a negation test is proposed to be able to separate these two types. This is when the clausal negation is used in the existential construction, and the subject is marked with the genitive case. However, Paducheva (2008:148, 153) has pointed out that this structural distinction is not a necessary nor a sufficient criterion, and there are other semantically motivated constructions which under the scope of negation obligatorily undergo the genitive alternation such as verbs of perception, (e.g., *наблюдаться* 'observe' and *слышаться* 'notice'). If the genitive test cannot be used to identify spatial construction from existential ones, other semantic features need to be explored in order to facilitate the differences between these two types.

Paducheva (2008:152; 2004:433-436) proposes another semantic feature, namely availability (cf. Апресян, Ю. Д., 1986). According to her classification, this feature is explicitly expressed by the verb *иметься* 'be' in Russian. Thus, the function of availability brings the concepts of *to have* and *to exist* into a close conceptual proximity. Example 7.3-6 gives the canonical reflexive verb, *иметься* 'exist', supporting the Existence Construction.

7.3-6 *Нет, акропол-ь име-ет-ся во* мног-их
No, acropol-NOM exist-3S.PRS-RM PR many-PREP.PL
город-ах, […].
city-PREP.PL
No, there is an acropolis in many cities.
[194. RNC, Интеллектуальные игры "З-С": Ответы на викторину "Привычные заблуждения // "Знание — сила", 2003]

The reflexive verb *остаться* 'stay, remain' also supports the Existence Construction, given in 7.3-7.

7.3-7 *А* *что* *оста-л-о-сь* *от* греческ-ого
And what.NOM remain-PST-N-RM PR Greek-GEN
наследи-я *сейчас,*
heritage-GEN now
через 10 или 12 веков после того как Гомер сочинил "Илиаду" и "Одиссею"? […]
And what remained of the Greek heritage now, after 10 or 12 centuries Homer had written Iliad and Odyssey?
[274, RNC, Сергей Смирнов. Конец серебряного века. Anno Domini 180 // "Знание — сила", №9", 2003]

If availability is understood broadly to cover also discourse function, the notion of availability can be linked to Huumo's argumentation on existential being disconnected, (i.e., locatives and existential serve a different discourse function). The entity introduced in an Existential Construction severs the connection between elements implicated or mentioned in the preceding discourse. Thus, the primary function of Existential Construction is to introduce new referents into the discourse (Huumo, 1996; 2003).[137]

Example 7.3-8 illustrates this particular property where a new referent, (i.e., 'mentioning'), is introduced with the Existence Construction.

7.3-8 *Как* *о* *лекарственн-ом* *растени-и* *упоминани-е*
As PR medicinal-PREP plant-PREP mentioning-NOM
о *ней* *встреча-ет-ся* *ещё* *у* *античн-ых*
PR it.PREP occur-3S.PRS-RM already PR ancient-GEN.PL
автор-ов.
author-GEN.PL
As a medicinal plant, the mentioning of it occurs already with ancient authors.
[879, RNC; Юрий Комаров. Горько! // "Сад своими руками", 2003.09.15]

However, Paducheva also distinguishes between existence and appearance, although these two types are, according to her, highly interconnected. For example, the verb *оказаться* 'seem' or appear' is considered to be purely perceptual. (Paducheva, 2008). In 7.3-9, the subject argument can be interpreted as being viewed from a certain perspective contrasting 7.3-10 where the profiled abstract entity shifts the profile more firmly to the canonical existential semantics.

[137] In contrast, Arutyunova (Арутюнова, 1976:221-223) has argued that the introductory existential construction should to be distinguished from other potential existential types.

7.3-9 *В* *рук-ах* *у* *него* *оказа-л-и-сь*
PR hand-PREP.PL PR his appear-PST-PL-RM
подтяжк-и *председател-я.*
bracer-NOM.PL chairman.GEN
The bracers of the chairman appeared in his hands.

7.3-10 […] [*т*]*о* *в* *более* *молод-ой* *част-и*
that PR more young-PREP portion-PREP
окаж-ет-ся *примерн-о* *равн-ое*
seem-3S.FUT-RM approximate-ADV equal-NOM
количеств-о *мужчин* *и* *женщин.*
quantity-NOM man.GEN.PL and woman.GEN.PL
That there will be approximately an equal quantity of men and women in the younger portion.
[246, RNC. Янис Астафьев. Кто будет работать в России в 2015 году? // "Отечественные записки", 2003]

Depending on the exact profile, the existential semantics can become closer to spatial location, as illustrated in 7.3-11.

7.3-11 *А* *почему* *карт-а* *оказыва-ет-ся* именн-о
And why map-NOM appear-3S.PRS-RM exact-ADV
в *этой* *колод-е?*
PR this.PREP stack-PREP
And why is the map in this stack exactly?
[1876Андрей Геласимов. Фокс Малдер похож на свинью (2001)]

7.3-12 [*к*]*онтужен-ый* *раз* *в* *год*
wounded-NOM once.NOM PR year.ACC
заявля-л-ся *к* *сво-ему* *благодетел-ю* […].
appear-PST.M-RM PR own-DAT benefactor-DAT
The wounded used to appear once a year to his benefactor
[1963. RNC, Олег Павлов. Карагандинские девятины, или Повесть последних дней // ""Октябрь", №8", 2001]

The latter instantiation type connects other reflexive verbs to the Existence Construction, as in 7.3-12 that are more strongly intertwined with the spatial relation.

7.3.1 Location Construction

The Location Construction contains 44 instances in the RF model and covers 15 unique reflexive verbs. Generally, the construction type appears to be infrequent and supported by a small number of reflexive verbs displaying strong lexical connectivity. The canonical instances are *находиться* 'be located, be,' *сохраниться* 'preserve' and *располагаться* 'be situated.' Although the number of unique reflexive verbs is small in this type, all the reflexive verbs have a neighbor displaying a stable form-based cross-paradigmatic relation although the perceived semantic similarity varies across verbs, as expected. For instance, both *находиться* ‘be located, be’ and *располагаться* ‘be situated’ are intermediate, whereas *сохраниться* ‘preserve’ is similar. The definition of the Location Construction follows.

Function: Profiles a spatial relation between entities.

Form: Nominative subject and a construal-specific secondary slot.

Similar to the Motion Construction, the encoding of the secondary slot is primarily anchored to the whole construal of the event rather than being an inherent property of the reflexive verb. From a semantic point of view, it is commonly argued that the secondary slot anchors the subject argument of the construction (Babby & Comrie, 1980:100; Paducheva, 2008; Золотова, Г. А., 2005 [1973]:231-232). The encoding of the secondary slot is dependent on the spatial configuration. Thus, they are independent of each other, and only in the configuration they become intertwined (Langacker, 1987:298-302). Example 7.3-13 gives an instantiation and the encoding is illustrated in Figure 7.3-1.

7.3-13 *Вообще-то* *замок* *находи-л-ся* *далеко-о-о-о-о*
Generally castle.NOM locate-PST.M-RM far-r-r-r-r
от *избушк-и,*
PR hut-GEN
но глашатай кричал так громко!
[…]
Generally, the castle was located so far-r-r-r-r from the hut, but the herald shouted so loudly.
[1710, RNC, Сергей Седов. Доброе сердце Робина // "Мурзилка", №7", 2002]

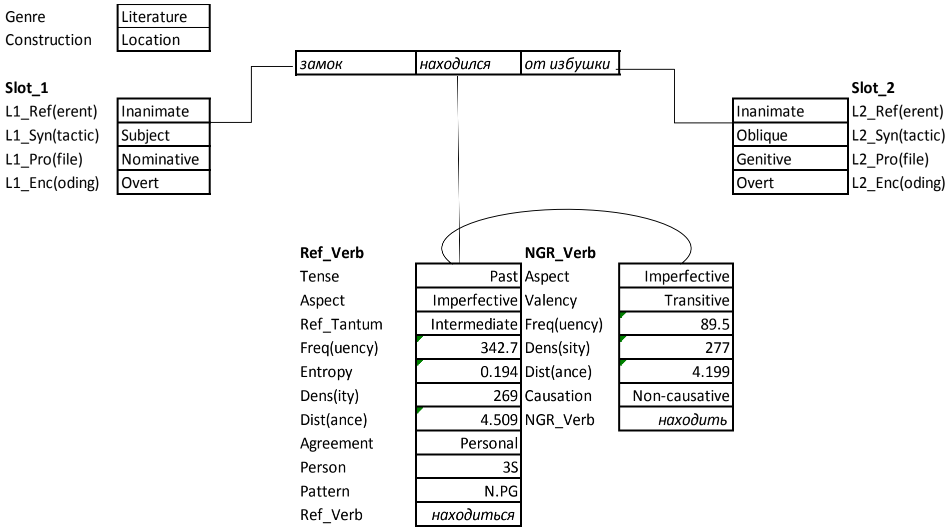


Figure 7.3-1 Layered structure of the canonical Location Construction.

The confusion matrix is given in Table 7.3-1.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	**0**	**0**	**0**	**0**	**29**	**9**	**0**	**0**	**0**	**0**	**0**	**6**	**0**
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 7.3-1 Confusion matrix of the predicted constructions. The Location Construction is given in bold.

The class-wise error of 0.341 shows that a fairly large portion of the instances is mislabeled by the model. Thus, this type follows the trend attested with such types as the Motion, Reciprocal, and Reflexive Engagement Constructions in terms of the class-wise error. Additionally, the recall of 0.659 indicates that the model has difficulties in identifying this particular construction type when compared globally to the stock of the instantiations in the data. In contrast, the precision value of 1 indicates that the model is able to classify these instances and the type does not attract other instantiations. The estimated class probabilities are given in Figure 7.3-2.

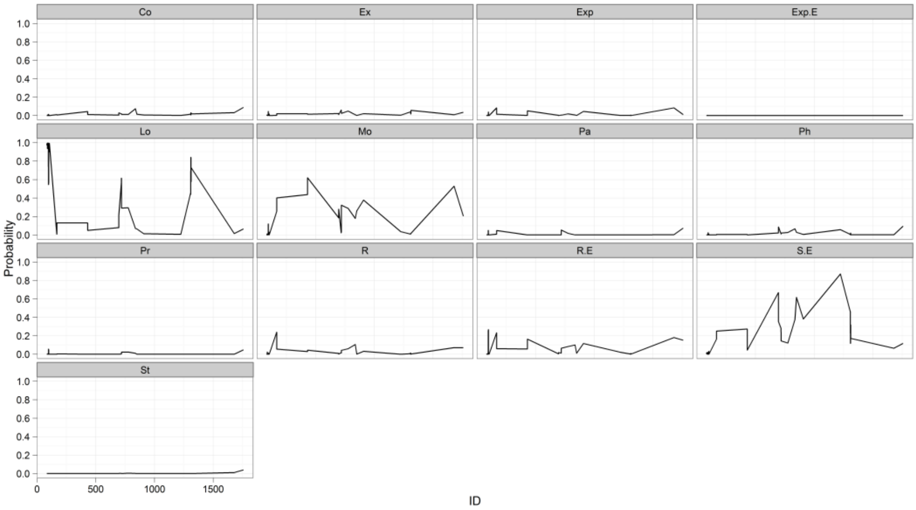


Figure 7.3-2 Faceted estimated class probability plot of the Location Construction (Lo). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The uncertainty in the classification is apparent based on Figure 7.3-2. One possible motivation to this is most likely the small number of data points considering the possible range of encoding the spatial relation, the input appears to be insufficient in covering it. Importantly, the Location Construction does not compete with the Existence although both of them share the basic configuration in terms of verb-specific properties. Instead, the competition appears to be with the Spontaneous Event Construction.

7.3.2 Existence Construction

The Existence Construction covers 85 instances, and contains 17 unique reflexive verbs. Similar to the Location Construction only a small number of reflexive verbs support this type displaying a strong lexical gravitation rather than a global generalization. The type is supported with such verbs as *иметься* ‘exist’, *остаться* ‘stay, remain,’ and *встречаться* ‘occur.’ Additionally, the canonical verbs of perception such as *оказаться* ‘seem, appear’ can also be merged with this construction type along with other reflexive verbs containing a perceptual component such as *появиться* ‘appear, emerge.’ The divergence of the reflexive verbs and, crucially, gravitation is present with the previously mentioned verbs. The cross-paradigmatic relation is supported with *иметься* ‘exist,’ and *встречаться* ‘occur,’ both having a similar neighbor verb contrasting *оказаться* ‘seem, appear,’ which is perceived dissimilar. Finally, the verbs *остаться* ‘stay, remain,’ and *появиться* ‘appear, emerge’ do not have neighbor verbs. Another layer of complexity occurs with the latter verb, where the prefixation creates a paradigmatic gap although the root, *-яв-*, supports cross-paradigmatic relations, (e.g., *являться* ‘be’ ~ *являть* ‘display, be’). The definition of the Existence Construction follows.

Function: Profiles an existential relation between entities.

Form: Nominative subject and a construal-specific secondary slot.

Similar to the Location Construction the encoding of the secondary slot is dependent on the profiled relation rather than a property of the reflexive verb. Example 7.3-14 illustrates the canonical instantiation and the encoding is given in Figure 7.3-3.

7.3-14 *На* *задн-ей* *стенк-е* *име-л-о-сь*
PR back-PREP wall-PREP exist-PST-N-RM
созвезди-е *загадочн-ых* *дырочек*
constellation-NOM enigmatic-GEN.PL hole.GEN.PL
таинственн-ого *происхождени-я,* […].
mysterious-GEN origin-GEN
A constellation of enigmatic holes of mysterious origin exists on the back wall.
[1817, RNC, Вячеслав Пьецух. Шкаф (1997)]

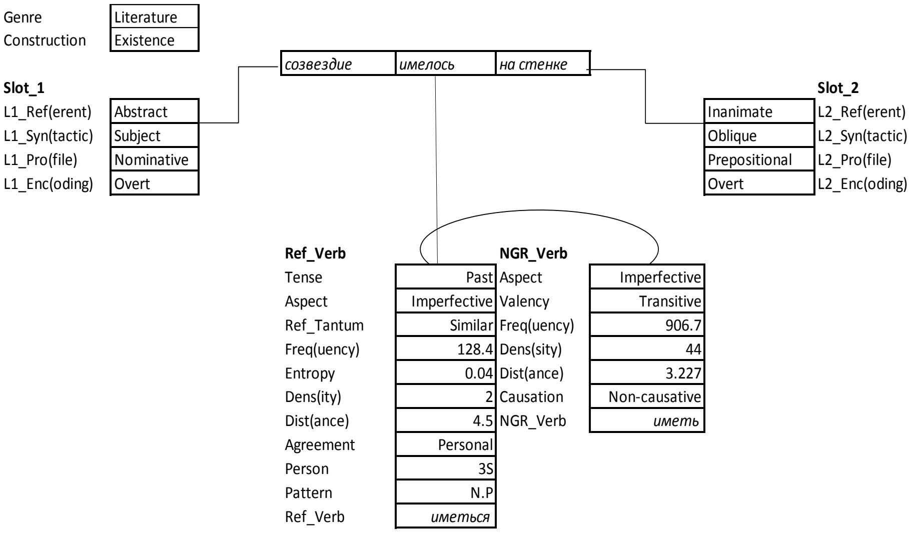


Figure 7.3-3 Layered structure of the canonical Existence Construction.[138]

The confusion matrix is given in Table 7.3-2.

[138] The syntactic role of the secondary slot is considered to be oblique as the roles are defined relative to the construction type rather than verb-specific properties. Thus, copula verbs such as *остаться* 'stay, remain' are not encoded with the role complement as is done, for instance in Paducheva (2008).

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	**0**	**77**	**0**	**0**	**0**	**4**	**0**	**0**	**0**	**0**	**1**	**3**	**0**
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	0	0	0	0	0	0	0	0	171	0	0	0	0
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 7.3-2 Confusion matrix of the predicted constructions. The Existense Construction is given in bold.

The class-wise error is low with the value of 0.094, considering that the structural properties of the Existence Construction are fairly similar to the Location and Spontaneous Event. Similarly, the recall value of 0.906 indicates that the type is identified through the structural properties of the verbs when contrasted globally. Furthermore, the precision value of 0.987 indicates that the model is able to accurately classify the instances and the type does not attract other instantiations in the model. One possible motivation behind the performance of the Existence Construction might also be related to the data. As the type covers only a small proportion of the reflexive verbs relative to the type frequency, the variable may contain sufficient input for the RF model to learn the patterns.

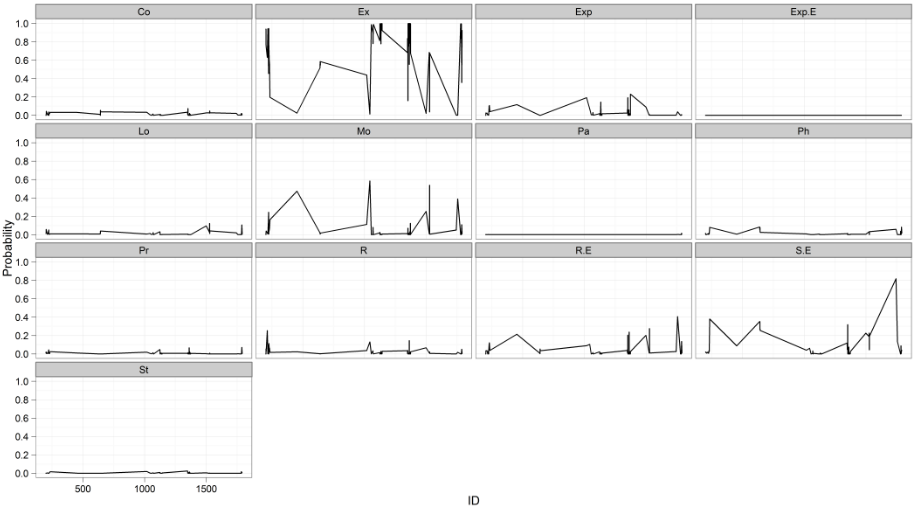


Figure 7.3-4 Faceted estimated class probability plot of the Existence Construction (Ex). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

The competition between different instantiations appears to be centered on the Reflexive Engagement (R.E) and the Spontaneous Event (S.E) in Figure 7.3-4. Example 7.3-15 illustrates this.

7.3-15 *Неожиданн-о перед ним появи-л-ся*
Sudden-ADV PR he.INS appear-PST.M-RM
человек и выстрели-л в лиц-о,
man.NOM and shoot-PST.M PR face-ACC
затем в ног-у.
then PR leg-ACC
Suddenly, a man appeared before him and shot him in the face and then to the leg.
[512, rnc, Военнослужащий отпускник ранен у кафе // "Московский комсомолец" в Саранске", 2004.12.23]

Based on the structural properties of the reflexive verbs, this particular instance appears to hover between the Existence, with the estimated class probability of 0.554, and the Reflexive Engagement, with the estimated class probability of 0.239. Generally, the canonical instances appear to be well-separated and the fluctuation concerns specific profiles rather than the construction type as such.

8 Property Constructions

This section introduces the basic argument construction type conventionalized in profiling the properties of a certain entity. The lexical basis of the generalization is connected to the reflexive verb *являться* 'be' either profiling a single property or a bundle of properties. The former type is typically realized by including adjective in the profile and the latter by nouns (cf. Jespersen, 1924:81; Langacker, 1987:183, 198; 1991:4-5; 2009:6-7; Wierzbicka, 1986). Generally, these instantiations are connected to copula constructions, one of the basic types of intransitive argument structures in language. At the same time, as a type these instances are typically excluded and rarely appear in any taxonomy proposed for the Russian Reflexive Marker. An exception to this is Gerritsen's (1990) study where these are briefly mentioned. Another important aspect is that the reflexive verbs supporting this type are among the most frequent reflexive verbs in Russian.

The three most frequent reflexive verbs are *оказаться* 'seem, appear', *являться* 'be,' and *остаться* 'stay, remain,' based on the Frequency dictionary, cf. Section 3.1.7. Consequently, they also form the lexical basis of the Property Construction. From a usage-based perspective, the incorporation of this type to the system of the Russian Reflexive Marker is of at most importance. Section 8.1 outlines the theoretical basis of the Property Construction connecting it to a larger body of studies on copula verbs. Additionally, the variation in the form pole is connected to verb-specific constructions. Finally, the Property Construction is modeled in Section 8.2.

8.1 Definitions of the Property Construction

Typological studies have shown that copulas are cross-linguistically a divergent category (cf. Eriksen, 2005; Hengeveld, 1992; Pustet, 2003; Stassen, 1997). [139] The term copula is encountered in most descriptive and theoretical works on Indo-European languages. However, the term seems to be self-evident and rigid definitions are lacking (cf. Krasovitsky et al., 2008). The traditional definition of a copula is related to the semantic emptiness assumption (cf. Hengeveld, 1992:32; Pustet, 2003:80-81; Stassen, 1997:65). The copula does not add any semantic content to the linguistic expression other than to carry the verbal morphology, (i.e., Tense, Aspect, and Mood). Thus, they require other elements to form the predicate core.

The category of copula is typically divided at least into two subgroups, (i.e., the copula proper, and semi- or quasi-copula). According to Hengeveld (1992:35), semi-copulas are differentiated from copula proper by two criteria. First, they add semantic content, and second, cannot be left unexpressed without changing the semantic content of the expression. According to Russian Grammar, the copula is defined as a segregated category from auxiliary verbs.

[139] Traditionally, the copulas are defined relative to truth condition, for example in Arutyunova (Арутюнова, 1999:449-450).

Moreover, the copula introduces an additional element to the clause which carries the semantic content. From the point of view of the Reflexive Marker, the following verbs are considered as the copula proper according to Russian Grammar: *являться* 'be,' *явиться* 'be, emerge,' and *называться* 'be called.' Russian Grammar also acknowledges the semi-copulas labeled as poly-meaningful verbs which are typically associated with such semantic content as existence, change of state, remaining in state, and discovery. Such verbs as *оказаться* 'seem, appear,' and *остаться* 'remain' exemplify the latter subgroup (Шведова & другие, 1982:120-121).

The verb *быть* 'be' is the canonical copula verb in Russian and the reflexive verb *являться* 'be,' is considered to be semantically synonymous, although associated with written language. In contrast, Shvedova (cf. Шведова, 2001 and references therein) considers that *быть* 'be' is a deictic verb that retains a meaningful function, an analysis akin to Cognitive Grammar, where copulas are considered to be meaningful (Langacker, 1991:64-65).[140] Examples 8.1-1 and 8.1-2 illustrate the canonical usage pattern, where either a noun (8.1-1) or adjective (8.1-2) is used with the copula verb to form the predicate core.

8.1-1 […] *предпринимател-ей* *котор-ые* *явля-ют-ся*
entrepreneur-GEN.PL which-NOM.PL be-3P.PRS-RM
банкрот-ами *или* *полу-банкрот-ами.*
bankrupt-INS.PL or half-bankrupt-INS.PL
Entrepreneurs who are bankrupt or nearly bankrupt.
[1096, RNC, Беседа на телевидении С. Шустера и С. Борисова, НТВ, "Герой дня" (2002)]

8.1-2 *Естественно, что в качестве базисного модельного аппарата*
[…]
эффективн-ым *явля-ет-ся* *использовани-е* *подход-ов*
effective-INS be-3S.PRS-RM use-NOM approach-GEN.PL
математичес-кого *и* *компьютерн-ого* *моделировани-я.*
mathematical-GEN and computer-GEN modeling-GEN
Naturally, the use of mathematical and computer modeling is effective as a basic modeling tool.
[281, RNC, Автоматизированная компьютерная система сопряженного геоэкологического мониторинга // "Геоинформатика", 2002.09.25]

There are, however, two caveats related to the semantic emptiness assumption. First, most of the reflexive copula verbs can appear in multiple argument

[140] Hengeveld (1992:39-42) also proposes a third label: the pseudo-copula. The label is reserved for such verbs as *seem*. They can combine with reduced complements with *to*-infinitive Construction in English, for instance *He seems ill* versus *He seemed to be ill*. This analysis resembles the raising predicates test used in formal approaches with *оказаться* 'seem, appear' (Perlmutter, David & Moore, 2002:628-629). The latter type has been analyzed as an instantiation of the Content construction in this study.

constructions. It is difficult to motive semantic expansions with empty categories (cf. Eriksen, 2005:16-17). Second, it is commonly acknowledged that certain full verbs can be used as copulas (cf. Hengeveld, 1992:39). It is unclear how this modification of the argument structure can be accounted for in a systematic way if the category of copula is semantically empty, or at least nearly empty.

The standard lexicalist position would ultimately lead to a description where the argument structure of a certain verb would be fully meaningful in one pattern, while in others it would be (semi)-meaningless. The reflexive verb *держаться* 'hold, contain' demonstrates the issue. It was shown that *держаться* can appear in the Reflexive Engagement Construction as in 8.1-3, repeated for convenience. The Property Construction is given in 8.1-4.

8.1-3 *Они* *держа-л-и-сь* *из* *последн-их* *сил* [...].
They.NOM keep-PST-PL-RM PR last-GEN.PL strength.GEN.PL
They hanged on with their last strength.
[1582, RNC, Александр Дорофеев. Эле-Фантик // "Мурзилка", №1 5", 2003]

8.1-4 [*м*]*о-и* *стих-и,* *когда* *хорош-и*,
My-NOM.PL poem-NOM.PL when good-NOM.PL
держ-ат-ся *мысл-ью,* [...].
contain-3P.PRS-RM meaning-INS
My poems, when they are good, contain meaning.
[478, RNC, С. Г. Бочаров. Из истории понимания Пушкина (1998)]

Example 8.1-4 aligns with the canonical copula verb *являться* 'be' in 8.1-5.

8.1-5 *Я* *напомн-ю* *что* *по* *федеральн-ому*
I.NOM remind-1S.FUT that PR federal-DAT
законодательств-у *мы* *явля-ем-ся*
legislation-DAT we.NOM be-1P.PRS-RM
орган-ом *власт-и* *законодательн-ой* [...].
body-INS administration-GEN legislative-GEN
I shall remind that according to the federal legislation we belong to the body of the legislative administration.
[1463, RNC, Заседание Московской городской думы 2004 // (2004.09.22)]

When the Property Construction is considered to be an argument construction type, the extensions can be motivated. The core of the Property Construction is established with the verb-specific construction type following the Nominative Instrumental pattern. Examples 8.1-6 and 8.1-7 illustrate the pattering with *оказаться* 'seem, appear' and *предполагаться* 'assume.'

8.1-6 *Дик-ие* *звер-и* *оказа-л-и-сь* *не*
Wild-NOM.PL beast-NOM.PL appear-PST-PL-RM NEG
менее *сообразительн-ыми,* *чем* *люди.*
less intelligent-INS.PL than people.NOM.PL
The wild animals did not appear less intelligent than people.
[21, RNC, [Пираньи // "Мурзилка", №8", 1999]

8.1-7 [*д*]*вижени-е* *жидкост-и* *предполага-ет-ся*
movement-NOM liquid-GEN assume-3S.PRS-RM
однонаправленн-ым [...].
unidirectional-INS
The movement of liquid is assumed to be unidirectional.
[132, RNC, Интерпретация результатов компьютерного моделирования фильтрации воды, нефти и оторочки меченой жидкости для зонально-неоднородного и слоисто-неоднородного нефтяного пласта-коллектора // "Геоинформатика", 2004.03.31]

Importantly, the reflexive verb *являться* 'be' contains 65 instances in the sample, concluding that the instrumental case is an obligatory element of the verb-specific construction, contrasting the verb *быть* 'be' that displays variation between the nominative, and the instrumental case (cf. Krasovitsky et al., 2008:102-103 for a detailed diccussion with references; Nichols, 1981; Timberlake, 1986; 2004; Булыгина & Шмелев, 1997:118-119).[141]

Recently, Krasovitsky et al. show, based on diachronic corpus data, that the usage of the instrumental case has shifted towards the grammatical end of the continuum. In the earlier periods, the case alternation between the nominative and instrumental is variable and displays an affinity towards multiple competing factors which determine the case marking.[142] However, the last period shows no overlap between the earlier ones and the occurrence of the instrumental case is more frequent indicating that the instrumental case marking might be undergoing grammaticalization in contemporary Russian (Krasovitsky et al., 2008).[143] This variation is illustrated in 8.1-8.

[141] Apresyan (Апресян, Ю. Д., 1995a) gives six different senses for the verb *быть* 'be:' copular, spatial, possessive, existential, modal-existential and auxiliary.

[142] Several semantically motivated criteria are proposed in the literature to account for the variation, examples include Janda (1993b), Arutyunova (Арутюнова, 1999), Bulygina and Shmelev (Булыгина & Шмелев, 1997), and Markman (2008). From a functional perspective, the semantics of the complement is often connected to time-stability (Givón, 1979; Pustet, 2003).

[143] The data set analyzed by Krasovitsky et al. consist of fifty-year periods between 1801 and 2000.

8.1-8 *А стенк-а оказа-л-а-сь тоненьк-ая и*
Ah wall-NOM appear-PST-F-RM thinnish-NOM and
от удар-ов слома-л-а-сь.
PR hit-GEN.PL break-PST-F-RM
Ah, the wall appeared to be thinnish and it broke when hit (lit. from hit).
[1985, RNC, Татьяна Рик. Про вредную Бабку-Ёжку // "Мурзилка" №6", 2001]

The variation in terms of case marking appears to be a verb-specific property. Additionally, another type of variation is present when adjectives are combined with the copula verb. Adjectives are divided into long and short forms in Russian. The latter type only appears in the predicative position in contemporary Russian.

The short forms are commonly analyzed as similar to the instrumental case, displaying some form of non-permanent property. Additionally, the short form adds a dynamic reading compared to the long form (Виноградов, 1972:213, 321-322).[144] Examples 8.1-9 and 8.1-10 demonstrate these patterns.

8.1-9 *Иначе, невзирая на ударные дозы удобрений,*
[…]
плодороди-е почв буд-ет остава-ть-ся
fertility-NOM soil.GEN.PL be-3S.FUT remain-INF-RM
низк-им.
low-INS
Otherwise, regardless of the shock dose of the fertilizer, the fertility of the soil will remain low.
[RNC, Сад на кислых почвах // "Сад своими руками", 2003.09.15]

8.1-10 *Слав-а бог-у жив-ы оста-л-и-сь.*
Thank-NOM God-DAT alive-NOM.PL remain-PST-PL-RM
Thank God they remained alive.
[1198, RNC, Монолог о прививках (2006.11)]

The reflexive verb *являться* 'be' can be considered to function as the gravitational center for the previously outlined instantiations connecting additional extension types exemplified by such verbs as *славиться* 'be famous,' *прикинуться* 'pretend,' and *называться* 'be called (names)', exemplified in 8.1-11 and 8.1-12.

[144] Additionally, the short forms can have lexicalized senses, for example *готовый* 'ready' versus *готов* 'drunk' (Виноградов, 1972:321). The usage of the short forms of the adjectives is declining steadily in Russian. Typically, the variation of the long and the short form is attributed to genre (Grannes, 1984).

8.1-11

Квентин	*слав-ит-ся*	*умение-м*	*находи-ть*
NAME.NOM	be.famous-3S.PRS-RM	ability-INS	find-INF
людей, […].			
people.ACC.PL			

Quentin is famous for the ability to find people

[608. RNC, Rendez-vous // "Экран и сцена", 2004.05.06]

8.1-12 *Опасней такие / как Явлинский /*

[…]

они	*прикидыва-ют-ся*	*интеллигенци-ей*
They.NOM	act-3P.PRS-RM	intelligentsia-INS
/ *демократией* […].		
/ democracy-INS		

Dangerous are those, like Yavlinskij. They act like intelligentsia and democracy.

[1025, RNC, В. Жириновский. Выступление В. Жириновского в программе "Свобода слова", НТВ (2004)]

Another alignment can be attributed to the verb *становиться* 'become.' It does not have a neighbor verb in Russian, but it is semantically in close proximity to such verbs as *получиться* 'turn out to be,' and *обернуться* 'turn into,' exemplified in 8.1-13 and 8.1-14. Similar instantiations are analyzed as part of the Medial type in Gerritsen (cf. 1990:36-37).

8.1-13

Короче роман	*получи-л-ся*	*бурн-ый*
Brieflyromance.NOM	turn.out.to.be-PST.M-RM	turbulent-NOM
и *коротк-ий.*		
and short.lived-NOM		

Briefly, the romance turned out to be turbulent and short-lived.

[1046, RNC, Разговор на улице между мужчиной и женщиной (2005.04.13)]

8.1-14

Паник-ой	*среди*	*сыктывкарц-ев*	*и*
Panic-INS	PR	NAME-GEN.PL	and
административн-ыми	*разборк-ами*	*оберну-л-о-сь*	
administrative-INS.PL	dissassembly-INS.PL	turn.into-PST-N-RM	
заявлени-е			
statement-NOM			

министра здравоохранения Коми Эльвиры Нечаевой, […].

[…]

The statement of the Minister of Public Health of Komi, Ehl'vira Nechaeva, turned into panic among the Syktyvkarians and into adminstrative dissassemblies.

[984, RNC, Ольга Муравская. Пойманное слово. // "Московский комсомолец в Сыктывкаре", 2003.08.06]

The previously outlined structural properties indicate that the Property Construction displays similar behavior to the canonical copula verb *быть* 'be,' with one critical difference, the variation in the profile of the argument construction is connected to verb-specific constructions, whereas the copula verb *быть* ‘be’ itself covers the whole range.

From a usage-based perspective, the Property Construction repeats the typical behavior of most linguistic categories. The core is well-established, but the edges are blurred (cf. Bybee, 2001:31-32; Lakoff, 1987:436-437). Considerable deviations occur when elements form idiomatic patterns, illustrated in 8.1-15.

8.1-15 *Кстати, и Эдвард Дженнер не оста-л-ся*
Besides and NAME.NOM NAME.NOM NEG remain-PST.M-RM
без наград-ы.
PR reward-GEN
Besides, Edward Jenner did not remain without a reward.
[26, RNC, Как родилась иммунология // "Знание — сила", №7", 2003]

The preposition *без*$_{\text{gen}}$ forms a lexicalized pattern with the verb *остаться* ‘stay, remain’ (Кузнецов, 2009 [1998]). Similar lexicalization can be demonstrated with the reflexive verb *годиться* ‘be useful,’ that is, the pattern *для*$_{\text{gen}}$ given in 8.1-16.

8.1-16 […], *и мног-ие из них*
and many-NOM.PL PR they.GEN.PL
год-ят-ся для изготовлени-я циновок.
be.useful-3P.PRS-RM PR manufacturing-GEN mat.GEN.PL
And many of them are useful for manufacturing mats.
[882, RNC, Елизавета Мельникова. Жатва на болоте // "Сад своими руками", 2003.09.15] [омонимия снята]

In terms of structural properties, Example 8.1-17 shows a strong departure from the canonical profile with the element *на руку* ‘beneficial, [lit. to the hand]’ that functions as the complement.

8.1-17 [..], *изменени-е регламент-а оказа-л-а-сь*
change-NOM regulation-GEN seem-PST-F-RM
нам на рук-у.
we.DAT PR beneficial-ACC
The change of regulation seemed beneficial to us.
[861, RNC, Наум Рашковский, Олег Стецко. Один за всех, все за одного // "64 - Шахматное обозрение", 2003.10.15]

Finally, most reflexive verbs of perception can combine with a dative argument when used in the Property Construction retaining their inherent verb-specific structure, as in 8.1-18 (cf. Падучева, 2004:231, 240-241, 253-255).

8.1-18 *Нам,* *однако,* *так-ая* *гипотез-а*
We.DAT however that-NOM hypothesis-NOM
представля-ет-ся *сомнительн-ой,* [...].
seem-3S.PRS-RM doubtful-INS
To us, however, that kind of hypothesis seems doubtful.
[191, RNC, Молекулярная эпидемиология вируса ECHO 30 на территории России и стран СНГ // "Вопросы вирусологии", 2002]

These examples profile an extension type where the entity perceiving the event is encoded with the dative case (Huumo, 2005:113-114, 125-126; Langacker, 2002:315-316; Noë, 2004:172-173). The verbs of perception show that the core of the Property Construction is present but the semantics of the argument construction is augmented by verb-specific constructions.

8.2 Property Construction

The Property Construction covers 171 data points, and contains 26 unique reflexive verbs in the RF model. The data suggest that, as a type, the Property Construction is fairly frequent, but lexically highly specific supported by three basic verb-specific constructions: *являться* 'be,' *становиться* 'become,' and *оказаться* 'seem, appear.' These reflexive verbs also form the basic semantic relations supporting the verb-specific constructions. A stative relation is profiled with *являться* 'be,' and a change of state is lexically grounded with the *становиться* 'become' (Апресян, Ю. Д., 1995b:37). In contrast, the *оказаться* 'seem, appear' type introduces a perceptual component to the relation. These verbs also capture the complex cross-paradigmatic structure of the reflexive verbs: *являться* 'be' has a semantically similar neighbor verb contrasting *оказаться* 'seem, appear,' which is semantically dissimilar to its neighbor.[145] Finally, *становиться* 'become' does not form a cross-paradigmatic relation. These reflexive verbs are also the most frequent reflexive verbs in Russian, cf. Section 3.1.7. Thus, a surface structure can be supported through multiple types leading to gravitation rather than rule-governed system. The definition of the Property Construction follows.

Function: Profiles a property of an entity.
Form: Nominative subject and a complement secondary slot.

The function of this construction type is to profile a certain property of an entity. The profiled property is introduced with a complement forming the predicate core with the reflexive verb. The canonical form pole is profiled with the Nominative-Instrumental pattern. The form pole can be further augmented by varying the encoding of the secondary slot. On one hand, the variation is localized to verb-specific constructions, such as the Nominative-Nominative pattern with the reflexive verbs *оказаться* 'seem, appear' and, *становиться* 'become' which is not attested with *являться* 'be.' On the other, this variation is

[145] In contrast, Kolomackij (Коломацкий, 2009:37) considers that the reflexive verb *являться* is dissimilar to its neighbor verb.

simultaneously a construal-specific related to the general variation attested with copula verbs in Russian. Example 8.2-1 illustrates the instantiation of the Nominative-Instrumental pattern and the encoding is given in Figure 8.2-1.

8.2-1 *В рамках общей политики снижения налогов*
[…]
приоритет-ом *бюджет-а* *явля-ет-ся*
priority-INS budget-GEN be-3S.PRS-RM
ненаращивани-е *госрасход-ов,* […].
decrease-NOM state.expenditure-GEN.PL
The decrease of state expenditure is the prority within the general policy of tax cuts.
[979, RNC, Михаил Классон, Алексей Полухин. Бюджет-2004 честный // "Время МН", 2003.08.06]

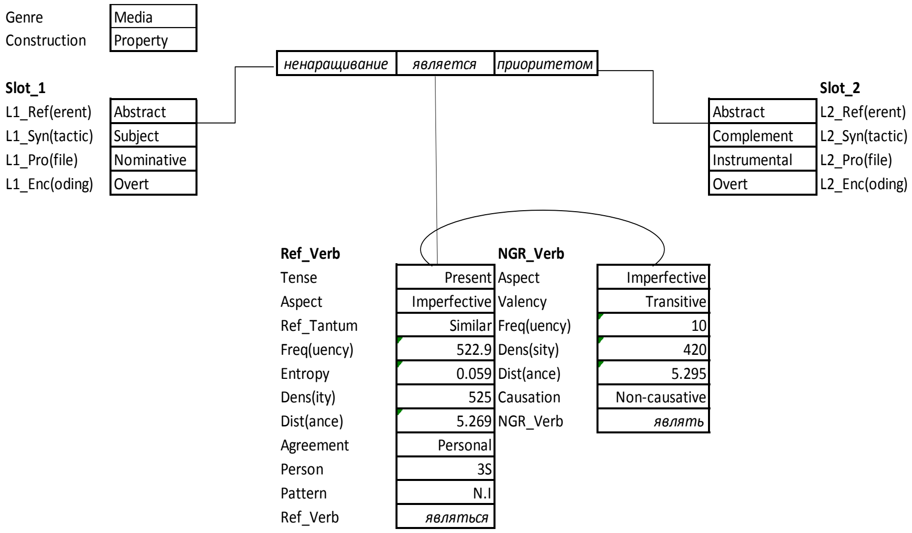


Figure 8.2-1 Layered structure of the canonical Property Construction.

The confusion matrix is given in Table 8.2-1. The structural properties of the reflexive verbs appear to capture exhaustively the generalization as both the recall and the precision are estimated to be 1. Similarly, the class-wise error is 0. These factors indicate that the construction does not enter into competition with other types in the sample.

	Predicted												
Observed	Co	Ex	Exp	Exp.E	Lo	Mo	Pa	Ph	Pr	R	R.E	S.E	St
Co(ntent)	93	0	5	0	0	3	0	0	0	0	4	5	0
Ex(istence)	0	77	0	0	0	4	0	0	0	0	1	3	0
Exp(eriencer)	1	0	93	0	0	12	2	1	0	4	21	2	1
Exp(eriencer) E(xtension)	0	0	0	54	0	0	0	0	0	0	0	0	0
Lo(cation)	0	0	0	0	29	9	0	0	0	0	0	6	0
Mo(tion)	0	1	6	0	0	141	0	0	0	9	23	32	0
Pa(ssive)	0	0	0	0	0	1	204	1	0	0	2	0	0
Ph(ase)	0	0	1	0	0	0	4	127	0	0	2	11	0
Pr(operty)	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**171**	**0**	**0**	**0**	**0**
R(eciprol)	0	0	2	0	1	13	0	0	0	70	13	4	0
R(eflexive) E(ngagement)	4	0	9	0	0	28	4	1	0	4	145	20	0
S(pontaneous) E(vent)	1	0	4	0	0	20	2	0	0	3	23	275	1
St(imulus)	0	0	0	0	0	0	0	0	0	0	0	0	65

Table 8.2-1 Confusion matrix of the predicted constructions. The Property Construction is given in bold.

Importantly, the estimated class probabilities display a degree of fluctuation in Figure 8.2-2 that is masked by the categorical labels.

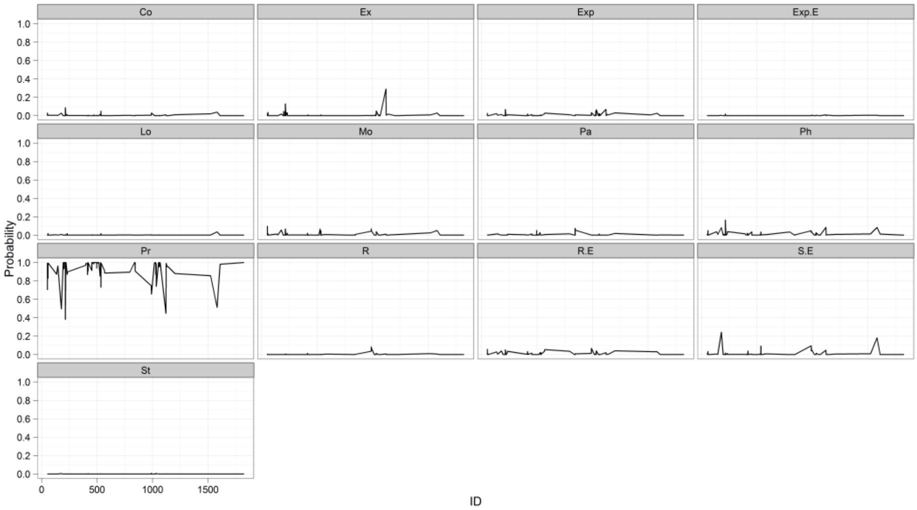


Figure 8.2-2 Faceted estimated class probability plot of the Property Construction (Pr). The y-axis gives the probability of the estimated class and the x-axis gives the number of the data point in the sample (ID). The estimated construction types are given in the facets.

Departures from the canonical instantiation type are reflected in the estimated class probabilities, for instance, such reflexive verbs as *годиться* 'be useful,' and *оказаться* 'seem, appear,' when profiled with the *на руку* 'beneficial, [lit. to the hand].' The latter instantiation has the lowest estimates class probability of 0.164, and is estimated to compete with the Existence Construction estimated class probability of 0.128. In this sense, the model captures the essential semantic structure of this particular verb in terms of fluctuation as *оказаться* 'seem, appear' also supports the Existence Construction. In sum, the Property Construction appears to be well separated when a global comparison is made across the argument constructions marked with the Reflexive Marker.

9 Minor Construction Types

The argument construction types deemed minor based on their type frequency in the sample are discussed in this chapter. The minor argument construction types come in two patterns, either forming semantically related types among themselves, or forming the extensions of the types already included in the RF model. The former pattern is related to the Semantic Reflexive and the Permissive Constructions covered in Sections 9.1, and 9.2. The latter pattern covers extensions from the Experiencer and the Stimulus Construction, discussed in Sections 9.4 and 9.3. The last two sections cover the extensions from the Property Construction.

9.1 Semantic Reflexive Construction

The Semantic Reflexive Construction covers 25 instances in the database, and 20 unique reflexive verbs. Traditionally, this type is defined in terms of the co-reference of the Agent and the Patient, exemplified in 9.1-1.[146]

9.1-1	*быстреньк-о*	*оде-л-а-сь*	*и*	*убежа-л-а*	*домой.*
	fast-ADV	dress-PST-F-RM	and	run-PST-F	to.home

She got dressed up fast and ran away to home.
[793, RNC, Я желанна. Разве это стыдно? // "Даша", №10", 2004]

The Construction type is supported by such verbs as *одеться* 'dress oneself,' *защищаться* 'defend oneself,' and *раздеться* 'undress oneself.' They have a semantically similar and a non-causative transitive neighbor verb. In functional accounts, the Semantic Reflexive Construction is considered to occupy the central position, and all other types of the Reflexive Marker are its extensions (Ahn, Hyug, 2005; Bostoen & Nzang-Bie, 2010; Kemmer, 1993). Nonetheless, the type frequency of the Semantic Reflexive appears to be fairly low. It seems that the low type frequency is not simply an artifact of the analysis. Knyazev (Князев, 2007:268-270) argues that this usage pattern is rather infrequent in contemporary Russian similar to Muchnik (Мучник, 1971:49).

According to Dankov (Данков, 1981:72, 76) the type frequency of the Semantic Reflexive has steadily declined in Russian. A similar distributional pattern is also estimated by Kalashnikova and Saj (Калашникова & Сай, 2006:3-4), 4% (*n* = 73) out of 1937 data points.[147] The definition of the Semantic Reflexive Construction follows.

Function: Profiles an action performed by an entity upon oneself.

Form: Nominative subject and construal-specific secondary slot.

The performed action in this construction type is primarily related to grooming and body care verbs discussed in Kemmer (1993:16-17). The encoding of the secondary slot reflects this. It is primarily connected to construal of the

[146] This category is labeled as the Agentive Reflexive in Gerritsen (1990:79).

[147] This type is labeled as Reflexive Proper in their study.

realization of the event, such as location or time. Israeli (1997:56) extends this classification by including verbs which are used to profile an activity acted upon oneself, for example, *защищаться* 'defend oneself.' This slightly broader definition enables to include such reflexive verbs as *описаться* 'wet oneself' to this construction type, as given in 9.1-2.

9.1-2 *Ирин-а описа-л-а-сь в тот самый момент,*
NAME-NOM wet-PST-F-RM PR that.ACC same-ACC moment.ACC
когда Володька её целовал.
[…]
Irina wet herself at the same moment when Volod'ka kissed her.
[1657, RNC, Токарева Виктория. Своя правда // ""Новый Мир", №9", 2002]

This extension can be used to include such problematic instances such as *спасаться* 'save oneself,' and *отогреваться* 'warm oneself,' as given in 9.1-3, and 9.1-4.

9.1-3 *Как спаса-ют-ся зайц-ы?*
How save-3P.PRS-RM rabbit-NOM.PL
How rabbits save themselves [from harm]?
[8, RNC, Как спасаются зайцы? // "Знание — сила", №7", 2003]

9.1-4 *Кафельн-ая плитк-а, на котор-ой мы*
Tile-NOM hot.plate-NOM PR which-PREP we.NOM
отогрева-л-и-сь в мороз-ы.
warm-PST-PL-RM PR cold-ACC
A tile hot plate, where we warmed in the cold.
[619, RNC, Иваново. Детство // "Экран и сцена", 2004.05.06]

These instantiations still have a bodily basis in the loose sense. They are: save oneself from bodily harm, and make oneself warm.

In derivational approaches, the Semantic Reflexive is traditionally considered to be synonymous or at least near-synonymous with the heavy reflexive marker. Thus, it is claimed that this constitutes a complementary distribution between the transitive verb, the reflexive verb the reflexive pronoun construction, such as, *мыть* 'wash' ~ *мыться* 'wash oneself' ~ *мыть себя* 'wash oneself' (cf. Geniušienė, 1987:10-11; Gerritsen, 1990:4-6, 76-77). In contrast, Israeli (1997:55-56) illustrates these patterns with the verbs *брить* 'shave' and ?*брить себя* 'shave oneself'.[148] The substitution with the reflexive pronoun is questionable at best. Similar objection to this substitution test is already expressed in Muchnik (Мучник, 1971:48) and Janko-Trinickaya (Янко-

[148] From a constructionist perspective, the dative reflexive pronoun aligns with the other functions associated with the dative case, namely the External Possessor in this case. (Podlesskaya & Rakhilina, 1999).

Триницкая, 1962:53-55).[149] However, the inclusion of the reflexive pronoun is possible but in the dative case as in 9.1-5.

9.1-5 *Брат* *бре-ет* *(себе)* *бород-у.*
Brother.NOM shave-3S.PRS -self.DAT beard-ACC
Brother is shaving his beard.
(Israeli, 1997:55)[150]

Additional distinction is made based on the affected entity, whether the verb profiles the body as a whole or a body part, such as, *зажмуриться* 'blink.' These types are typically classified as partitive reflexives verbs in the Russian Tradition (cf. Geniušienė, 1987:81). The position taken in Kemmer (1993) and Israel (1997) is taken here. Kemmer (1993:55) considers that the body as a whole constitutes semantically the basic type. Example 9.1-6 illustrates this pattern. Thus, they are considered as instantiations of the Semantic Reflexive Construction.

9.1-6 *Пчёлка Зоя сложила крылышки и*
[…]
зажмури-л-а-сь *от* *удовольстви-я.*
blink-PST-F-RM PR joy-GEN
Zoya the bee furled her wings and blinked from joy.
[1671, RNC, Виктор Кологрив. Медовый луг // "Мурзилка", №5", 2002]

In sum, the data suggest that the Semantic Reflexive forms a small cluster of semantically related verbs centered on profiling an event type grounded in bodily action.

9.2 Permissive Construction

The Permissive Construction covers nine instances and seven unique reflexive verbs in the sample. They are: *фотографироваться* 'get one's photograph taken,' *излечиться* 'get healthy, cured,' *поправляться* 'get well,' *лечиться* 'undergo treatment,' *обследоваться* 'get oneself examined,' and *выписаться* 'sign out.' It seems that this type appears to be related to the Semantic Reflexive Construction as the reflexive verbs appear to be centered on the bodily actions although the reflexive verbs *фотографироваться* 'get one's photograph taken,' and *выписаться* 'sign out' profile an action performed upon oneself. Examples 9.2-1 and 9.2-2 illustrate the Permissive Construction.

[149] Interestingly, similar observations are made in Geniušienė (1987:76-77 and referecense therein) but this difference is not considered to be significant enough to disclaim the near-synonymous position.

[150] The example was transcribed into Cyrillic and glossing was added by the author.

9.2-1 *Философ – тот, кто долже-н излечи-ть-ся*
Philosopher.NOM – that, who.NOM obligate-M cure-INF-RM
от мног-их недуг-ов рассудк-а, […].
PR many-GEN.PL ailment-GEN.PL mind-GEN
A philosopher is a person who is obligated to get oneself cured from the many ailments of the mind.
[374, RNC, Владимир Успенский. Виттенштейн и основания математики (2002)]

9.2-2 *Она фотографиру-ет-ся постоянн-о /*
She.NOM photograph-3S.PRS-RM constant-ADV
/ везде просто / я те говорю / она делает это целыми днями.
[…]
She is constantly getting her photographs taken. Simply everywhere. I mean that she is doing it all day, every day.
[1053, RNC, Разговор при выходе из дома, Москва (2005.04)]

The definition of the Permissive Construction follows.

Function: Profiles a caused action by an entity upon oneself.
Form: Nominative subject and construal-specific secondary slot.

In derivational approaches, these reflexive verbs are often labeled as Reflexive Causative by Knyazev (Князев, 2007) and Geniušienė (1987). The causativity of this type is related to a prior event that is caused by the subject referent. The unfolding of the prior event causes the profiled event in the argument construction (cf. Frajzyngier, 2000). The profiled events tend to be intentional, illustrated with the broader context in 9.2-2.

In sum, the data suggest that the Permissive Construction is an extremely small type, and is connected to the Semantic Reflexive Construction through the shared semantic basis of bodily action.

9.3 Stimulus Extension Construction

The Stimulus Extension Construction is covered by four instances, and three unique reflexive verbs in the total sample. They are: *ощущаться* 'be felt,' *забыться* 'forget,' and *почувствоваться* 'become felt.' All of them have a non-causative and semantically similar neighbor verb. Examples 9.3-1 and 9.3-2 illustrate the usage patterns.

9.3-1 *Вс-е это забуд-ет-ся.*
All-NOM this.NOM forget-3S.FUT-RM
All this will be forgotten.
[1078, RNC, Беседа с рок-музыкантами в ресторане "Японский городовой" (2003)]

9.3-2 *Ирони-я противопоставлени-я, заложе-нн-ая в ней,*
Irony-NOM contrast-GEN embed-PPP-NOM PR it.PREP
особенн-о ощуща-ет-ся в сам-их Луховиц-ах, […].
especial-ADV feel-3S.PRS-RM PR very-PREP.PL NAME-PREP.PL
The irony of contrast, embedded in it, is especially felt in Luhovicy.

[543, RNC, Луховицкие узоры // "Народное творчество", 2003.08.18]

These patterns are considered to be extensions of the Stimulus Construction. The definition of the Stimulus Extension Construction follows.

Function: Profiles an augmented mental relation with focused Entity and backgrounded Entity.
Form: Nominative subject and a construal-specific secondary slot.

The profiled mental relation is augmented with this extension type. The Stimulus occupies the subject slot but the Experiencer is fully backgrounded. Paducheva (2004:204) considers that the basic semantic structure of *ощущать* 'feel, sense,' the neighbor verb of *ощущаться* 'be felt,' is a sensation inside oneself excluding the outside world. Furthermore, Paducheva (Падучева, 2004:209-212) assumes a difference between semantic roles of Experiencer and Viewer/Perceiver (Наблюдатель).[151] The profile excludes the role of the Experiencer and introduces the Viewer/Perceiver which is incorporated as part of the semantic structure of the expression. Thus, there is a contrast between Examples 9.3-3 and 9.3-4 illustrated with the reflexive verb *слышаться* 'hear.'[152]

9.3-3 *[с]лыш-ит-ся шум.*
hear-3S.PRS-RM noise.NOM
A noise is heard.
(Падучева, 2004:212)

9.3-4 *[м]не слыш-ит-ся шум.*
I.DAT hear-3S.PRS-RM noise.NOM
I noticed [lit. hear] a noise.
(Падучева, 2004:212)

Thus, this distinction is captured as a difference between the Stimulus Construction (9.3-4) and the Stimulus Extension Construction (9.3-3) in terms of the proposed argument construction types. Paducheva's examples conveniently illustrate the multiple patterns of the verb *слышаться* 'hear'.

9.4 The Experiencer Perspectivization Construction

The Experiencer Perspectivization Construction contains one instantiation in the sample, *житься* 'live.' It has a non-causative and semantically similar neighbor verb. This particular construction is extremely infrequent but displays, at least, partial productivity as was already discussed in Section 2.2.2. Importantly, this argument construction breaks away from the general pattern of who-did-what-to-whom (cf. Goldberg, 2006:106; Tomasello, 2003:126). The generalization incorporates a manner component, typically a modifier. (Kyröläinen, 2008). Hence, the mental relation is augmented by a construal of

[151] It is worth pointing out that this kind of sublexical properties are analyzed in Langacker's (2002) Cognitive Grammar in relation to the concept of Active Zone.

[152] The glossing and the translations were added to Examples 9.3-3 and 9.3-4 by the author.

perspectivization, exemplified in 9.4-1.

9.4-1 *Что же делается,*
[…]
чтобы люд-ям жи-л-о-сь по-людск-и?
that people-DAT.PL live-PST-N-RM people-ADV
What would one do so that people would live in a fair way [lit. human]?
[936, RNC, Б. Варецкий. Стыдные уроки барства. Власть и бедность // "Советская Россия", 2003.08.21]

Overall, this type is one of the most prominent patterns among the various impersonal construction types in Russian. Thus, this type has received enormous attention in the literature (Geniušienė, 1987; Gerritsen, 1990; Israeli, 1997; Галкина-Федорук, 1958; Золотова, Г. А., 2000b; Недялков, 1978). The definition of the Experiencer Perspectivization Construction follows.

Function: Profiles an evaluation on a mental state of an entity.
Form: Dative subject and a modifier.

Generally, this type is one of the few argument constructions of the Russian Reflexive Marker that is not compositional. First, the verb slot is rendered as an expression of state. This property cannot be attributed to the dative argument or derived from the neighbor verb. Second, the modifier becomes an essential part of the argument construction. In contrast to previous taxonomies, Israeli (1997:136) considers such patterns as *думается что* 'think that' as instantiation of this type. Traditionally, this pattern would be analyzed as some subtype of the Passive Construction, (i.e., an impersonal passive). The pattern is not attested in the sample, but it follows the Stimulus Construction proposed in this study with the clausal subject, cf. Section 6.4. However, Gerritsen divides this pattern into subtypes depending on the degree of volition contrasting such verbs as *работаться* 'work' ~ *работать* 'work,' and *икаться* 'hiccup' ~*икать* 'hiccup' (Gerritsen, 1990:167-174). Certainly a possible analysis if one is determined to find the smallest possible partitioning. In terms of Construction Grammar, this would not change the analysis. The schematic argument construction, labeled here as the Experiencer Perspectivization, covers these instantiations and the degree of volition constitutes the semantic component of the verb-specific constructions.

9.5 Inclusion Construction

The Inclusion construction covers 38 instantiations, and 11 unique reflexive verbs in the sample. The construction type is supported with such reflexive verbs as *заключаться* 'consist,' *базироваться* 'be based,' *основываться* 'be based, found,' and *относиться* 'belong, pertain.' All the instantiations have a neighbor verb in Russian. The type can be considered as an extension of the Property Construction, cf. Sections 8.1 and 8.2. Examples 9.5-1 and 9.5-2 illustrate usage patterns.[153]

[153] Some of these verbs are labeled as Medial Proper in the taxonomy proposed by

9.5-1 *Данн-ая публикаци-я относ-ит-ся к перв-ой*
This-NOM publication-NOM pertain-3S.PRS-RM PR first-DAT
попытк-е системн-ой оценк-и так-их
attempt-DAT systematic-GEN evaluation-GEN such-GEN.PL
важн-ых природно-техногенн-ых объект-ов,
important-GEN.PL natural.technological-GEN.PL facility-GEN.PL
как ПХГ.
as UGSF
This publication pertains to the first attempt of a systematic evaluation of important natural-technological facilities such as UGSF.[154]
[101, RNC, Геоинформационное картографирование для оценки воздействия на окружающую среду объектов нефтегазовой промышленности // "Геоинформатика", 2001.03.14]

9.5-2 *Наш Новый год связан с архетипическими представлениями о календарном*
[…]
времени,
[…]
на котор-ых основыва-ю-тся традици-и
PR which-PREP.PL base-3P.PRS-RM tradition-NOM.PL
календарн-ой обрядност-и вс-ех народ-ов
calendar-GEN ritualism-GEN all-GEN.PL nation-GEN.PL
мир-а.
world-GEN
Our New Year is connected to the archetypical conceptualizations of calendar time upon which the traditions of the calendar ritualism of all the nations of the world is based.
[239, RNC, Олег Николаев. Новый год: праздник или ожидание праздника? // "Отечественные записки", 2003]

The definition of the Inclusion Construction follows.

Function: Profiles a part-whole relation between entities.

Form: Nominative subject and verb-specific secondary slot.

The function captures the semantics of the core instantiations given previously. The form pole follows the Nominative subject pattern, but the reflexive verbs display item-specificity in terms of the encoding of the secondary slot. In contrast, certain reflexive verbs are analyzed as pertaining to the class of converse reflexive verbs in derivational approaches, for example, содержаться 'contain.' Geniušienė considers this pattern to be related to the Passive from a functional perspective. The object of the base verb is assumed to be promoted to subject position (Geniušienė, 1987:271-273), exemplified in 9.5-3–9.5-5.

9.5-3 *Обычно они обходят стороной деревья,*
[…]

Israeli (1997:67).

[154] ПХГ = UGSF = underground gas storage facility.

в листь-ях котор-ых содерж-ат-ся алкалоид-ы.
PR leaf-PREP.PL which-PREP.PL contain-3P.PRS-RM alkaloid-NOM.PL
Typically, they [giraffes] bypass trees whose leaves contain alkaloids.
[74, RNC, Как и люди // "Знание — сила", №1", 2003]

9.5-4 *На экран-е отобража-ет-ся таблиц-а символ-ов*
PR screen-PREP reflect-3S.PRS-RM table-NOM symbol-GEN.PL
из пят-и строк и десят-и столбц-ов.
PR five-GEN row.GEN.PL and ten-GEN column-GEN.PL
The table of symbols consisting of five rows and ten columns is reflecting on the screen.
[574, RNC Клавиатура из одной клавиши // "Computerworld", 2004.07.30]

9.5-5 *[и] я не очень понимаю*
[…]
почему миров-ые цен-ы на нефт-ь
why world-NOM.PL price-NOM.PL PR oil-NOM
так жестк-о отража-ют-ся на наш-их
so close-ADV reflect-3P.RMS-RM PR our-PREP.PL
внутренн-их цен-ах.
internal-PREP.PL price-PREP.PL
And I do not quite understand why the world prices of oil are reflected so closely in our internal prices.
[1459, RNC, Первый канал, Москва // (2004.10.13)]

Similar positions are taken by Knyazev (Князев, 2007:287), and Dolinina (Долинина, 1991:329). However, once we move away from the traditional examples, such as *содержаться* 'contain,' and *отражаться* 'reflect,' the strict derivational account becomes more difficult to maintain. Such reflexive verbs as *вписываться* 'fit' ~ *вписывать* 'fill in, insert' and *приходиться* 'fit, suit' ~ *приходить* 'come, arrive,' and *относиться* 'belong, pertain' ~ *относить* 'take, carry' are perceived to be semantically intermediate. Importantly, the latter verb combines with multiple patterns. It is also attested in the Experiencer Extension Construction with the Dative-Infinitive pattern, (i.e., 'have to'), cf. Section 6.3. The Inclusion Construction is exemplified in 9.5-6.

9.5-6 *(ведь не секрет, что значительн-ая част-ь*
surely NEG secret.NOM that large-NOM part-NOM
продаж дорог-их алкогольн-ых брэнд-ов
sale.GEN.PL e xpensive-GEN.PL alcoholic-GEN.PL brand-GEN.PL
приход-ит-ся как.раз на оптов-ые закупк-и
fit-3S.PRS-RM precisely PR wholesale-ACC.PL purchase-ACC.PL
ночн-ых клуб-ов, дискотек и больш-их
night-GEN.PL club-GEN.PL disco.GEN.PL and large-GEN.PL
вечеринок).
reception.GEN.PL
Surely, it is not a secret, that the large part of sales of the expensive

alcoholic brands fits precisely the wholesale purchases
of night clubs, discos and large receptions.
[840, RNC, Владимир Ляпоров. Молодая гвардия. Искусство
быстрого завоевания новых рынков сбыта // "Бизнес журнал",
2003.10.23]

The reflexive verb *строиться* 'build' exemplifies another complicating factor for the derivational approach. Example 9.5-7 demonstrates another verbs-specific construction of this particular verb with the Nominative-Prepositional pattern.

9.5-7 *Тем.более, что исследовани-е стро-ит-ся* на
Moreover that research-NOM build-3S.PRS-RM PR
системно-картографическ-ом подход-е.
systematic.cartographic-PREP method-PREP
Moreover, the research builds on the systematic-cartographic method.
[400, RNC, Геоинформационное картографирование для оценки
воздействия на окружающую среду объектов
нефтегазовой промышленности // "Геоинформатика", 2001.03.14]

Although *строиться* 'build' is a prototypical verb in the literature, this particular instantiation is rarely mentioned. Kolomackij (Коломацкий, 2009:165-166) discusses this pattern, but considers it as an instantiation of the Passive Construction. If the *строиться* 'build' is simply derived from the neighbor verb *строить* 'build,' certain semantic aspect of this construction type is difficult to motivate, especially if it considered to be an instantiations of the Passive Construction. As also noted in Kolomackij (Коломацкий, 2009:165-166), this pattern is stative. Thus, the derivational directionality becomes difficult to maintain considering that the (Reflexive) Passive is not associated with stative semantics. The same derivational rule cannot be applied to form both the (Reflexive) Passive and this particular type. From a constructionist perspective, this particular extension aligns with other verbs-specific constructions, such as *базироваться* 'be based', *основываться* 'be based, found,' and *заключаться* 'be confined' within the verb-specific constructions of the Inclusion Construction. Thus, the stative semantics is the contribution of this particular subtype.

The Inclusion Construction appears to be supported by a small number of reflexive verbs but its function is well separated within the system of the Russian Reflexive Marker. Importantly, the function can be viewed as an extension from the Property Construction which is primarily supported by the most frequent reflexive verbs in Russian.

9.6 Relation Construction

The Relation Construction covers 26 instantiations in the sample, and six unique reflexive verbs. They are: *относиться* 'regard,' *отличаться* 'differ,' *приклеиваться* 'adhere,' *равняться* 'correspond,' *чередоваться* 'alternate,' and *соотноситься* 'correlate.' Examples 9.6-1 and 9.6-2 illustrate usage patterns.

9.6-1 *как я отнош-у-сь к этим люд-ям*
How I.NOM regard-1S.PRS-RM PR this.DAT.PL people-DAT.PL
или как я отнош-у-сь к этой професси-и?
or how I.NOM regard-1S.PRS-RM PR this.DAT profession-DAT
How do I regard these people or how do I regard this profession?
[1260, RNC, Интервью с кинорежиссером на телеканале НТВ (2006.04)]

9.6-2 *Даже счастлив-ая семь-я, о котор-ой* реч-ь,
Even happy-NOM family-NOM PR which-PREP speech-NOM
пережива-л-а неприятн-ые минут-ы,
go.through-PST-F unpleasant-NOM.PL moment-NOM.PL
поскольку очень отличал-а-сь от остальн-ых
because considerably differ-PST-F-RM PR other-GEN.PL
семей этого город-а.
family.GEN.PL this.GEN city-GEN
Even a happy family, like the one being discussed, has gone through unpleasant moments because they differed considerably from the other families in the city.
[1572, RNC, Александр Дорофеев. Эле-Фантик // "Мурзилка", №1 5", 2003]

The definition of the Relation Construction follows.

Function: Profiles a comparative relation between entities.

Form: Nominative subject and verb-specific secondary lost.

The verb-specificity is apparent in Examples. The reflexive verbs, nonetheless, appear to form small verb-specific constructions, such as *относиться* ‘regard’ and *приклеиваться* ‘adhere’ combine with the Nominative-Prepositional Dative contrasting *равняться* ‘correspond,’ which combines with the Nominative-Dative pattern.

In terms of function, this construction type profiles a comparison between entities. Langacker posits three basic functional components to capture an act of comparisons. Like most cognitive operations, the act of comparison is taken to be directional consisting of the Standard (first component), and the Target (second component) yielding an asymmetrical schematic structure of S → T. The S functions as the baseline for the comparison against which the T is evaluated. The third component is the asymmetrical operation (scanning) holding between the S and T (Langacker, 1987:102-103). Croft and Cruse (2004:54-55) go even further by associating an act of comparison with categorization.

In sum, this argument construction appears to be supported by a small number of reflexive verbs similar to the Inclusion Construction, but the basic function pertains to a basic function, an act of comparison.

10 Towards the Reflexive Space

This chapter goes beyond the descriptive devices and zooms into the category of the Russian Reflexive Marker. Section 10.1 introduces the concept of variable importance estimated based on the RF model. The relative importance of the predictors can be understood as an explicit ranking of the slots of the argument constructions when the slots are understood broadly covering all the variables in the model. Following the tenets of the usage-based model, the variables of the linguistic model do not necessarily have an equal status. This inequality then becomes the difference in importance in predicting the proposed set of argument constructions. Lastly, Section 10.2 deals with the fundamental property of Construction Grammar, namely the network structure of argument constructions.

Although the network structure is the primary means to state generalizations in Construction Grammar, the theoretical basis of establishing the actual network has received relatively little discussion in usage-based models. Several different networks have been proposed in the literature for the Russian Reflexive Marker (Ahn, Hyug, 2005; Ahn, Hyung, 2012; Enger & Nesset, 1999; Janda, 1993a; Williams, 1993). In contrast, a novel method is introduced to form networks in a data-driven manner that builds on Goldberg's (1995:70-72) formulation on motivating linguistic structures in a network. Linguistic structures tend to be partially overlapping, amounting to redundancy. A similar issue was already discussed in Section 4 in terms of correlation between predictors. The theoretical basis of the method rests on this assumption that redundancy is a fundamental property of a language, and by exploiting it a network structure can be derived in a data-driven manner. The RF model contains a distance measure and this is used to form generalizations over the argument constructions.

In sum, this chapter explores the slots of the proposed layered structure of the argument constructions. In addition, the utilization of the clustering technique examines the possibility of forming a smaller number of similar units from the RF model. Section 10.1 deals with the concept of variable importance and the global ranking of predictors is established. The principles of establishing the network structure are discussed in Section 10.2.

10.1 Variable Importance: Global Ranking of Predictors

Variable selection is one of the most difficult tasks related to modeling. It depends on how the concept of importance is operationalized, for example by using *p*-values as an indicator of importance. The caveat, however, is a statistically significant predictor in a model offers evidence for its importance in modeling the response variable. However, the reverse does not hold. Variable selection becomes increasingly difficult when a high number of predictors are used. Guyon and Elisseeff (2003:1158) discuss another related issue in variable selection, namely the difference between relevant and useful variables. Predictors may be relevant for the phenomenon but if they are redundant, the model may be less than optimal for predicting the response variable. In contrast, a set of useful predictors may improve prediction accuracy but exclude a number of relevant predictors. In linguistics, variables that tap into the same functional region may be partly redundant, (e.g., the case marking of the subject and the syntactic role of the subject). From a usage-based perspective, variable importance can be considered to be an important property. Large-scale corpora and databases have become available in recent years, increasing the importance of models that can handle a substantial number of predictors.

Although random forests are still under active research, they are used in genetics and bioinformatics for variable selection to identify potentially important variables among a set of hundreds, or thousands of predictors. Lunetta et al. (2004) have shown that the variable importance measure estimated, based on random forests, outperforms standard univariate methods such as Fisher's exact test. In a linguistic setting, Arppe (2008:216-217) discusses similar connections between univariate and multivariate methods. The attractive property of the variable importance measure is that interactions are included in the estimation.

Although this is an explanatory study, the ranking of the predictors extracted from the model may offer some guidelines for future studies. The ranking of the variables is based on the estimated importance of the predictors in modeling the response variable. Section 10.1.1 introduces the variable importance measures available with the random forests algorithm. Resampling was used to establish the ranking of the predictors based on 1,000 samples to factor in uncertainty. Section 10.1.1 covers the technical aspects of estimating variable importance with random forests for the purposes of application. The random forest algorithm contains three variable importance measures.

The ranking of the predictors in relation to usage-based models is discussed in Section 10.1.2. The interpretation is not easily related to the existing literature on the Russian Reflexive Marker. On one hand, the diathesis tradition does not operate on degrees or importance. On the other, the lack of sample-based studies means that the results cannot be compared to previous taxonomies. Nonetheless, a connection can be established with functional studies, specifically with the Relevance Hypothesis proposed by Bybee. The Relevance Hypothesis was proposed for testing whether a particular morphological marker of the verb is likely to pertain to the continuum between inflection and derivation. Additionally Bybee proposed that the more relevant a certain marker is, the greater the impact on the semantic structure of the verb (Bybee, 1985:20-23). Because the model is used to predict the abstract semantics of the argument constructions, the variables that are proposed impact on the semantics should correspond to important predictors estimated with the RF model. The results show that the raking of the predictors can be partly attributed to the Relevance Hypothesis.

10.1.1 Variable Importance Measures with Random Forests

The simplest variable importance measure would be to count how many times a certain predictor is selected in the ensemble. However, this selection frequency tells very little about the predictor. There are three variable importance measures available with the Random Forests algorithm. Their interpretation requires a slight reorientation compared to regression analysis, because the importance is not associated with a *p*-value. The commonality of these variable importance measures is that they are averaged across the ensemble. This alleviates some of the potential ordering effects that might occur when multiple predictors are used in the model (Breiman, 2001a:cf. ; Díaz-Uriarte & Alvarez de Andrés, 2006; Strobl et al., 2009b).

The simplest of the available variable importance measures is reduction in node impurity, also referred to as the Gini importance. This measure is related to how the ensemble is constructed. Recall that a binary split is performed on the data, and the daughter nodes are always more pure compared to the mother node. The Gini importance reflects this property of the random forests, an average improvement in splitting the response variable into more homogenous groups across all the trees in the ensemble. However, recent simulation studies have shown that this measure is sensitive to the scale of the predictors or the number of levels of categorical predictors. Specifically, the Gini importance is biased towards variables with multiple cut-off points. (cf. Boulesteix et al., 2012; Strobl et al., 2007).

With large samples and a binary response variable, the variable selection of CART follows χ^2 distribution. When the recursive partitioning of the data continues, the categorical variables can become exhausted earlier compared to numeric variables or categorical variables with multiple levels. This bias appears to be related to the depth of the trees in the forest (Boulesteix et al., 2012). This bias may be especially prominent in linguistic applications, because typical studies contain both categorical and frequency-based variables. To make this more palatable, the selection frequency of the predictors based on stumps was recorded using resampling. Stump refers to a tree model. The first split in the tree is two terminal nodes. Conceptually, this kind of model is related to bivariate models. The simulated models had the same parameters as the RF mode used in the analysis of the argument construction with the exception that the number of splits was restricted to one, that is, the number of trees = 2,000, and the number of randomly selected variables at each split = 8. In this vein, the stumps had eight randomized variables to choose from the total number of predictors (25). Because the random forests algorithm already contains a random selection of data, a basic resampling is straightforward. Thus, 1,000 models were built without specifying the random seed to create divergent data sets.

Figure 10.1-1 gives the results obtained using stumps. The y-axis gives the label of the predictors ordered based on the selection frequency and the x-axis gives the selection frequency. The boxplots visualize the variability.

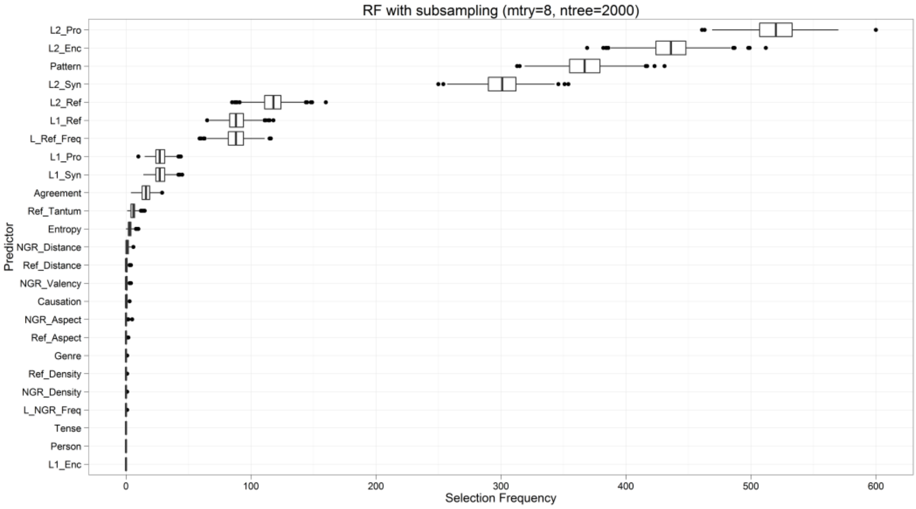


Figure 10.1-1 Boxplots of the selection frequencies of the predictors in the RF model based on resampling (1000) stumps.

The results show that four predictors have the highest selection frequency. They are: the Profile of the Secondary Slot (L2_Pro), the Encoding of the Secondary Slot (L2_Enc), the Pattern, and the Syntactic Role of the Secondary Slot (L2_Syn). A second resampling was conducted with the RF model but the maximum number of terminal nodes was now increased to 13, reflecting the number of the levels of the response variable. On one hand, this procedure incorporates the possible interaction between the predictors. On the other, this illustrates the change in the selection frequencies when the depth of the trees becomes greater. The results are given in Figure 10.1-2.

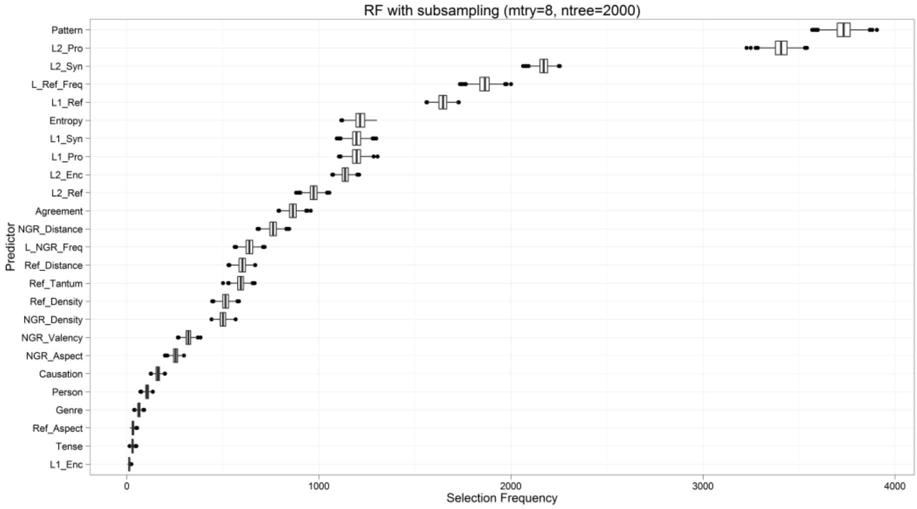


Figure 10.1-2 Boxplots of the selection frequencies of the predictors in the RF model restricted to 13 terminal nodes.

The Pattern and the Profile of the Secondary Slot (L2_Pro) have the highest selection frequency. The selection frequency of the Encoding of the Secondary Slot (L2_Enc) has decreased and the log Frequency of the Reflexive Verb (L_Ref_Freq) is selected more often now compared to the stumps. A final resampling was conducted with the RF model by increasing the maximum number of terminal nodes to 155. The mean value of the terminal nodes in the RF model was 309 (min. = 262 and max. = 362), cf. Section 4.4. Thus, this resampling reflects where the depth of the trees has become greater compared to the second one. The results are given in Figure 10.1-3.

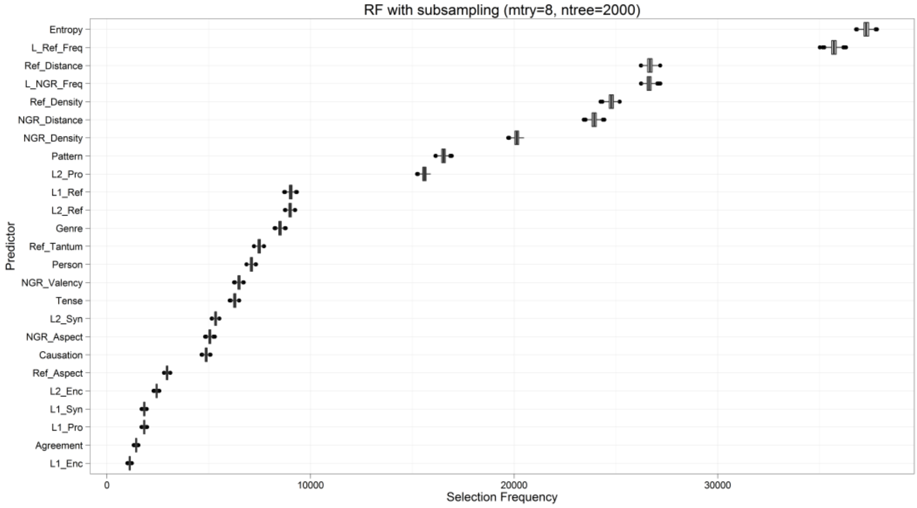


Figure 10.1-3 Boxplots of the selection frequencies of the predictors in the RF model restricted to 155 terminal nodes.

The bias towards predictors with multiple cut-off points is clearly visible. All the numeric predictors were selected most often. In sum, these resampling showed the effect of the depth of the trees at various stages when a single tree is constructed. The variable split criterion in the RF model is biased with the randomForest algorithm, and is reflected in the Gini importance (cf. Boulesteix et al., 2012; Nicodemus et al., 2010; Strobl et al., 2007). Furthermore, the Gini importance does not have an easily interpretable relation to the theoretical basis of Construction Grammar, (e.g., the conceptual counter-part of the node impurity in Construction Grammar).[155]

In addition to Gini importance, two advanced importance measures are available. Both of them, referred to as the Permutation variable importance, are conceptually related but differ whether the measure is scaled or not. Recall that the data are divided into two sets during the model building: in-bag and the out-of-the-bag (OOB). A single tree in the ensemble has never seen the OOB-data. This procedure also forms the basis of the estimated prediction accuracy of the model, cf. Section 4.3. For the advanced variable importance measure this property is exploited. Breiman proposes that a reasonable estimation of the importance of the predictors can be obtained by randomly permuting a predictor, and then using this permuted predictor together with the original predictors in predicting the response variable. The estimation of the importance of the predictor is computed by comparing the difference in accuracy before and after permutation, averaged over the trees in the ensemble. (Breiman, 2001a:23-24).

The rationale of this procedure is that if the predictor is associated with the response variable, this association is broken when the predictor is permuted and this impacts the prediction accuracy of the model. If the referent type of the subject argument of the Passive Construction is associated with the levels, Inanimate and Abstract, by randomly permuting the levels into Animate, this association is broken, and the prediction accuracy decreases. In this vein, the Permutation importance mimics the influence of the predictor when it is absent from the model. The Permutation importance is a non-parametric approach, but the same principle can also be extended to other statistical models in evaluating variable importance as is done in Baayen (2011:308-310).

Based on Strobl et al., the Permutation importance is computed. The Permutation importance of the predictor X_j is:

[155] It might be conceivable to connect the reduction in node impurity to schematicity in terms of probabilities. Initially, the unpartitioned data represents the most schematic representation, and the partitioning of the data gives the probabilities of arriving at the more homogenous groups. For example, the Transitive Construction is a schematic representation but can be partitioned into verb-specific constructions such as mental verbs, speech act verbs and verbs of motion. Given a set of predictors, the partitioning would represent the path to these homogenous groups.

the out-of-bag sample $\overline{\mathfrak{B}}^{(t)}$ for a tree t, with $t \in \{1, \ldots ntree\}$:

$$VI^{(t)}(X_j) = \frac{\sum_{i\in\overline{\mathfrak{B}}^{(t)}} I\left(y_i = \hat{y}_i^{(t)}\right)}{\left|\overline{\mathfrak{B}}^{(t)}\right|} - \frac{\sum_{i\in\overline{\mathfrak{B}}^{(t)}} I\,(y_i = \hat{y}_{i,\psi_j}^{(t)}}{\left|\overline{\mathfrak{B}}^{(t)}\right|}$$

The predicted class for observation i is $\hat{y}_i^{(t)} = f^{(t)}(X_i)$ and $\hat{y}_{i,\psi_j}^{(t)} = f^{(t)}\left(X_{i,\psi_j}\right)$ after permuting the predictor. The unscaled Permutation importance is calculated from the average importance over all trees:[156]

$$VI(X_j) = \frac{\sum_{t=1}^{ntree} VI^{(t)}(X_j)}{ntree}$$

The second Permutation variable importance is a scaled version, the variable importance divided by standard error. (Strobl et al., 2009b:21-22). However, recent studies have indicated that the scaled Permutation importance is sensitive to the number of trees (ntree) used in the ensemble. The parameter ntree is user-defined, and can be changed fairly arbitrarily indicating that the scaled version is not a reliable measure of importance (Boulesteix et al., 2012; Díaz-Uriarte & Alvarez de Andrés, 2006; Strobl et al., 2007; Strobl & Zeiles, 2008). Thus, the unscaled version is used and we concentrate on its properties, referred to henceforth as the Permutation importance.

A number of simulation studies have shown that the Permutation importance is unbiased under the null hypothesis, (i.e., there is no association between the predictor and the response variable). The bias of the base learners does not carry over to the OOB-data because they have never seen it, but the Permutation importance has a higher variance due to the bias of the base learners. (Boulesteix et al., 2012; Díaz-Uriarte & Alvarez de Andrés, 2006; Nicodemus et al., 2010). However, simulation studies indicate that the Permutation variable importance shows a slight bias towards correlated predictors (Strobl et al., 2008). Nicodemus et al. (2010) report a slight bias for correlated predictors under null hypothesis, but the difference between the median values of the correlated and the uncorrelated predictor was less than 0.014.

The conditional random forests algorithm contains a new, conditional permutation importance measure that factors in correlation between the predictors but it is computationally demanding. For example, Nicodemus et al. used a smaller proportion (*n* = 500) of their data (*N* = 2,000) in calculating the conditional permutation importance (Nicodemus et al., 2010:11). In this study, the data would require a matrix with 2e+12 columns; the decimal is moved 12 times to the right. This demonstrates that even a fairly small linguistic data set can become extremely complex when the whole structure is attempted to be analyzed simultaneously. Thus, the conditional permutation importance can become computationally infeasible, as is the case here. As discussed in Nicodemus et al. (Nicodemus et al., 2010), the slight bias towards correlated

[156] The unscaled Permutation importance can yield negative values because an irrelevant predictor may appear as relevant when permutated.

predictors also depends on the research question. For example, if the goal is estimate a set of variables that govern the same function, the correlation may be beneficial. On the other hand, correlated variables may also lead to spurious results.

In terms of a comparison between different methods, Nicodemus et al. (2010:11) offer evidence that the Permutation importance yields results that can be considered to be intermediate to estimations obtained from multivariate linear regression. Additionally, Boulesteix et al. (2012) indicate that the predictors that were estimated to be highly important received a higher estimated importance. In comparison, Baayen (2011:308-310) has shown that different models can yield different estimations for predictors in modeling the Dative alternation in English, although the models had a fairly similar prediction accuracy. Thus, method can yield different results because the underlying assumptions vary.

In sum, this section introduced the Permutation variable importance. The current evidence indicates that the measure has the following properties: 1) it is unbiased under null hypothesis, 2) has slight bias towards correlated predictors, and 3) has a higher variance. Resampling is generally recommended to test the stability of the ranking of the predictors with random forests because another random component is involved in the estimation, the permutation of the predictors.

10.1.2 Variable Importance: The Relevance Principle Revisited

The Permutation variable importance is an estimation of the importance of the predictors in predicting the response variable. In regard to a polytomous response variable, the ranking of the predictors is estimated to apply across the levels of the response variable. The estimation does not evaluate the difference between the Passive and the Spontaneous Event. Instead, it offers a perspective on separating globally the proposed set of the argument constructions. This is an important distinction because traditionally categories are evaluated in a binary form, but there are 13 labels to choose from in the model given the input. A final distinction is also relevant to keep in mind. It is the difference between prediction, and description. Certain variables may be important for describing a phenomenon in linguistics, but the descriptive device need not be a good predictor. Thus, the ranking of the predictors gives estimation in terms of prediction given the proposed set of the argument constructions and the input.

In terms of interpretation, the global ranking of the predictors can be interpreted in comparison to Bybee's Relevance Hypothesis for morphological markers of verbs. The hierarchy is semantically motivated. The categories are ordered in the hierarchy, from left to right, in terms of the semantic impact on the verb. For example, valency change is assumed to have the greatest semantic impact, and the person/number agreement the lowest (Bybee, 1985:20-23). It is described below.

valency change < voice < aspect < tense < mood < person/number agreement

Bybee defines the valency as the number or the role of the argument the verb contrasting voice that is defined relative to perspective, that is, a change in the relation of the surface subject differentiating such voice types as the passive and reciprocal (Bybee, 1985:20, 28). For these reasons, the Relevance Hypothesis can only be used to motivate certain facets of the predictors. The valency change can be viewed as an indicator of the case marking of the slots and the variable Pattern. Similarly, the voice can be understood relative to the type of the subject roles.

As outlined in the previous section, the ranking is estimated with the unscaled Permutation importance in this study. Resampling was used and 1,000 models were grown with the same tuning parameters as the original RF model: the number of trees = 2,000, the number of randomly selected variables available in a split = 8, and subsampling of the data. The estimated variable importance appears to be able to be divided into three parts. However, the magnitude of the variable importance measure should not be interpreted exactly in terms of percentages, but only in comparison to the relative ranking of the other predictors (Strobl et al., 2009b:30). The six most important variables appear to be fairly well separated from the remaining ones, and are given in Figure 10.1-4. The variable importance plot is zoomed into the region of the high ranking variables to preserve the scales.

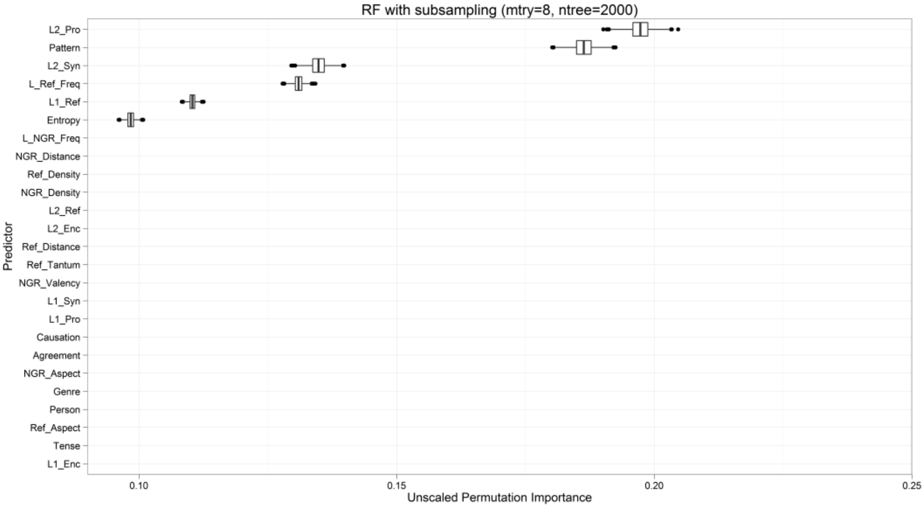


Figure 10.1-4 Zoomed Boxplots of the high ranking predictors in the RF model estimated with the Permutation importance based on resampling (1000).

The ranking of the predictors is evident, but the higher variance of the Permutation importance is visible in the plot. For example, the outliers of the Profile of the Secondary Slot (L2_Pro), and the Pattern partly overlap, but the median values of the predictors are well separated, indicating that the results are stable. Thus, the resampling offers a way to tackle the uncertainty in estimations.

The two most important predictors in the model are estimated to be the

Profile of the Secondary Slot (L2_Pro), and the Pattern. They are followed by the Syntactic Role of the Secondary Slot (L2_Syn). Although non-subject roles tend to have a marginal role in Construction Grammar (cf. Croft, 2001:272-275; Goldberg, 2006:42-43), similar results are reported in Hwang et al. in modeling the variation between the Caused Motion Non-Caused Motion (*N* = 1,880) in English. According to their results, the lexical preposition (76 types) such as *through* was the most important predictor compared to such variables as the semantic type of the preposition (27 types) extracted from the VerbNet and the semantic classes of the verb (123 types) extracted from the VerbNet[157] (Hwang, Nielsen & Palmer, 2010). Additionally, the log Frequency of the Reflexive Verb (L_Ref_Freq), the Referent Type of the Subject (L1_Ref), and the Entropy are estimated to be important predictors in the model. The relative ranking of the log Frequency and the Entropy follow the tenet of the usage-based models, where frequency-based variables are significant contributors to semantic structures (Bybee, 2010; Goldberg, 2006; Hay & Baayen, 2005) Additionally, the importance of the Referent Type of the Subject aligns with the descriptive practice in the diathesis tradition (cf. Geniušienė, 1987). The third set of predictors is given in Figure 10.1-5.

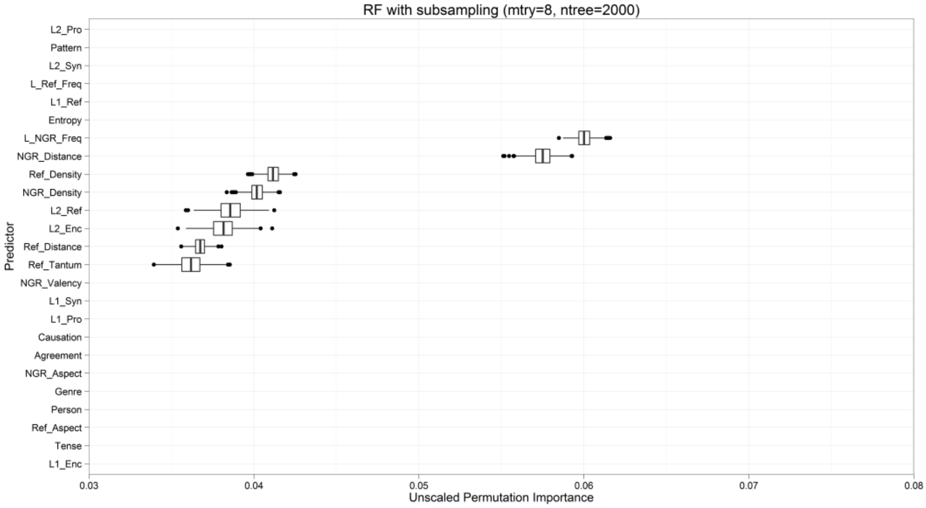


Figure 10.1-5 Zoomed Boxplots of the middle ranking predictors in the RF model estimated with the unscaled permutation variable importance based on resampling (1000).

There are eight predictors in this set labeled as the middle ranking ones, but the log Frequency of the Neighbor Verb (L_NGR_Freq), and the Neighborhood Distance of the Neighbor Verb (NGR_Distance) form their own set in terms of the estimated importance. The joint contribution of these

[157] VerbNet builds on Levin's (1993) classification of verb alternation types in English. The database is freely available online at the following address http://verbs.colorado.edu/~mpalmer/projects/verbnet.html. (Kipper-Schuler, 2005).

predictors indicates the relevance of the cross-paradigmatic relation, but both of these variables adhere to the degree of connectivity between items, and not to the discrete derivation.

The remaining predictors among the middle ranking ones form a fairly tight cluster. However, the Neighborhood Density of the Reflexive (Ref_Density), and the Neighbor (NGR_Density) Verb are estimated to be more important than such traditional variables as the Referent Type of the Secondary Slot (L2_Ref), and the Encoding of the Secondary Slot (L2_Enc). From a theoretical point of view, the fairly low importance of the Reflexiva Tantum is surprising, considering that it was indented to capture the degree of perceived semantic similarity between cross-paradigmatic items. The results, however, are expected once we factor in the distribution of the variable. The vast majority of the unique neighbor verbs (*n* = 565) were perceived to be semantically similar to the reflexive verbs. Very little cue-validity originates from a variable with a fairly homogenous distribution. Another surprising result is the weaker performance of the Neighborhood Distance of the Reflexive Verb compared to the Neighborhood Distance of the Neighbor Verb. This result might indicate that the operationalization of this degree relation might be more prone to target cross-paradigmatic relations rather than paradigmatic ones. The final set of the predictors is given in Figure 10.1-6.

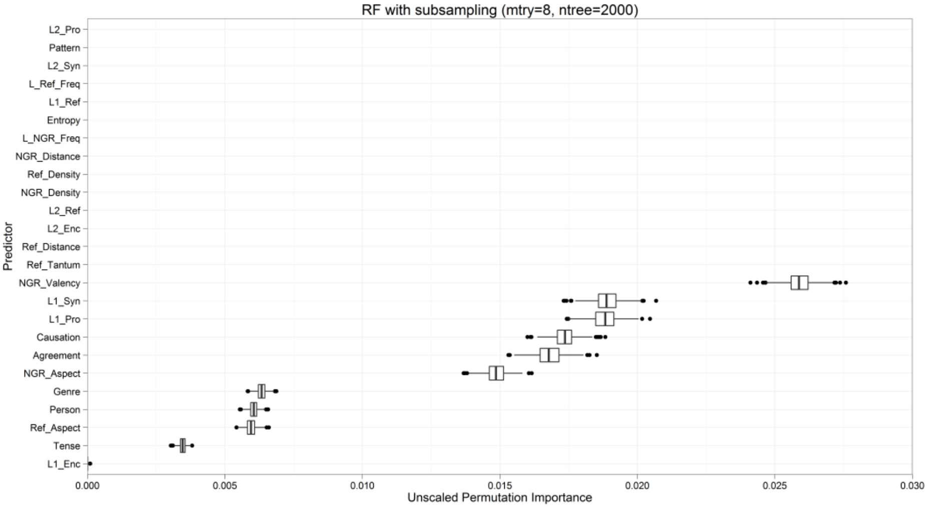


Figure 10.1-6 Zoomed Boxplots of the low ranking predictors in the RF model estimated with the unscaled permutation variable importance based on resampling (1000).

The final set of the ranking covers predictors with minimal contribution, (low ranking ones). The Encoding of the Subject Slot (L1_Enc) appears to be a truly irrelevant variable in the model. The variable is primarily concerned with information structure, a change in the encoding type offers very little cue-validity about the semantics of the argument construction. Intuitively, the ranking of the predictor is excepted. The Valency of the Neighbor Verb (NGR_Valency) is traditionally one of the primary descriptive tools in linguistic theories, but the ranking of the predictor is estimated to be low similar to the Reflexiva Tantum. A similar explanation might be used to motivate its low importance. The vast majority of the unique neighbor verbs (*n* = 533) were estimated to be transitive, excluding the bivalent ones. Simply knowing that the neighbor verb is transitive offers very little cue-validity in itself. This also applies to the Syntactic Role of the Subject (L1_Syn), and the Profile of the Subject (L1_Pro). In the latter case, the majority of the data points, once again, were attested with the Nominative subject in the RF model: 1,763 instances out of 1,878. Intuitively, this estimation is accurate considering the distribution. If the Nominative subject, and the fact that the verb is marked with the Reflexive Marker are the only variables known about the clause, predicting the clausal meaning is a fairly hopeless endeavor.

The Causation is another weak predictor. A direct comparison to the diathesis tradition cannot be made because this study is concerned with surface structures and not with derivational rules. At the same time, the Decausative alternation does not constitute a clausal meaning. A difference between the Motion Construction such as *кинуться* 'fling, flop' ~ *кинуть* 'throw, fling,' and the Experiencer Construction such as *успокоиться* 'calm down' ~ *успокоить* 'calm' would have to be achieved in some other manner, perhaps linking through the semantic classes of the verbs (cf. Падучева, 2001; Падучева, 2004). For a rule-based account, the question about the contribution of the causative component is different.

The Aspect of the Reflexive Verb (Ref_Aspect), and the Neighbor verb (NGR_Aspect) are also estimated to be weak predictors. As the ranking is estimated based on a global comparison, the results are expected. Certainly, aspect is important for the description of certain argument construction types, the Passive being perhaps the most prominent candidate, cf. Section 5.1. But once a global comparison is made, the cue-validity is fairly minimal similar to the Syntactic Role, and the Profile of the Subject Slot.

The low ranking of the Agreement, the Tense, and the Person follow Bybee's Relevance Hypothesis. The ranking of the Person, nonetheless, warrants a slight elaboration. The data suggest that the argument constructions are heavily skewed towards the third person singular marking (*n* = 1,255). For future studies on the formation of the cross-paradigmatic relation, this skewed distribution should be taken into consideration. Finally, the low ranking of the Genre may be due to the level of granularity. The semantics of the proposed argument constructions occupy a fairly coarse-grain level. Thus, the genre-related importance of a specific reflexive verb diminishes.

Generally, the contribution of the six most important predictors suggests that a fairly large portion of the argument constructions can be captured, based on item-specific information (cf. Goldberg, 2006:49-54) such as the log Frequency of the Reflexive Verb, the Entropy, and the Profile of the Secondary Slot. The Pattern and the Referent Type of Subject may be considered to be less item-specific, because they generalize over larger patterns. The importance of the latter predictor also mimics the concept of subject- and object-orientated verbs proposed by Janko-Trinickaya (Янко-Триницкая, 1962), with an emphasis on the mimic, as the distinction is not defined in categorical terms. Instead, these variables present the categorical distinction in probabilistic terms as subjects tend to be human, or animate, and objects tend to be abstract, or inanimate.

Although the high ranking variables are strictly established through generalizations over surface structures, the importance of the log Frequency of the Neighbor Verb and the Neighborhood Distance of the Neighbor Verb support the view that larger network structures influence semantics, specifically, the cross-paradigmatic relation. The diathesis tradition has underscored this from early on (Князев, 2007; Падучева, 2002; Храковский, 1974), but the ranking of these predictors indicates that there is more to this relation than the dichotomy aligning with the concept of gradient structures (Bybee, 2010; Hay & Baayen, 2005). Furthermore, the density- and the distance-based variables of the reflexive and the neighbor verbs emphasize the more fine-grained structure that originates from the verb-specific constructions (cf. Barðdal, 2008; Iwata, 2008). By introducing degrees through these variables, the contribution of this property can be systematically modeled as an influence on the semantics of the schematic argument constructions.

In sum, the variables evaluated in this study are portable and scale well with larger samples. The ranking of the predictors can be evaluated straightforwardly in future studies, when, or if gold-standard data becomes publicly available in Russian. The high ranking variables also offer evidence that a fairly decent performance in disambiguating reflexive verbs in Russian might be achieved with a set of six predictors that are based on surface structures.

10.2 The Constructional Network

The network structure is the primary means to establish generalizations in Construction Grammar. The basic principle of network models in constructionist approaches is to establish systematic relations between constructions through inherited properties. Another important property is that the connections, or links themselves, are considered to be a special kind of construction types (Croft, 2001; Fried & Östman, 2004; Goldberg, 1995; Kay & Fillmore, 1999). These will be referred to as linking constructions, and are assumed to be the abstract structure of the Russian Reflexive Marker. Section 10.2.1 addresses the principles of the network model in Construction Grammar. Importantly, this enables us to maintain the traditional distinction between the Reflexive Verb, and the Reflexive Marker in the diathesis tradition (Князев, 2007).

At the same time, the sum of the linking constructions is not postulated to be an invariant meaning not are the linking constructions some unique functions of the Reflexive Marker. Instead, the linking constructions follow the same principle as was argued to be present for the lexical densities. Lexical densities, and the verb-specific constructions, support the abstract argument construction, whereas the linking constructions are assumed to be another consequence of a density-based category, a layer that supports the argument constructions. Thus, the principle of redundancy is exploited fully. Section 10.2.2 describes how the random forests can be utilized to form the network in conjunction with clustering. The implementation and the interpretation of the results are discussed in Sections 10.2.3 and 10.2.4.

10.2.1 Links between Argument Constructions

Several competing descriptions are available to establish a constructional network. An excellent survey of these is given in Goldberg (1995 chapter 3). The discussion in this study is delimited to the so called full-entry model which is employed in the Cognitive Construction Grammar and Radical Construction Grammar. The full-entry models are redundant in a sense that the same information is shared between all construction types, and the inheritance links between constructions are the shared information between them. The full-entry model is more suitable, in my view, for variable-based models, as the links can be potentially induced from the data rather than committing to a specific model of abstraction. In a full-entry model, the same set of variables is present for every construction type. However, there is a tacit assumption that full-entry models are inferior to other modes. Contrary to this claim, the aim is to demonstrate that linguistically meaningful generalizations can be obtained, even in full-entry models. Finally, a full-entry model can be used to form a data-driven approach to establish linking constructions, because the same information is available at all levels, there is no need for a priori reduction of variables.

Four types of links between argument constructions are elaborated in Goldberg (1995). The polysemy link is used to establish a primary sense of a construction that connects extensions through it. Goldberg illustrates this with the Ditransitive Constructions: 1) central sense – X causes Y to receive Z (*give*), 2) conditions of satisfaction imply X causes Y to receive Z (*promise*), 3) X enables Y to receive Z (*permit*), 4) X causes Y not to receive Z (*refuse*), 5) X intends to cause Y to receive Z (*bake*), and 6) X acts to cause Y to receive Z at some future time point (*bequeath*). A subpart construction is defined as link type when a construction is a proper subpart of another construction, but exists independently. This relationship is illustrated with the Caused-Motion and Intransitive Motion Construction. (Goldberg, 1995:75-78).

In my view, the relation between the Caused-Motion and the Intransitive Motion would differ in Radical Construction Grammar, as more generic constructions are assumed to link to more specific ones. Caused motion is certainly more specific than a highly generic Intransitive Motion (Croft, 2001:53-58 cf. Figure 1.15). The third link is Instance Link, defined as a special case of another construction illustrated with the verb *drive*. When the verb is used in the Resultative Construction, it conveys the sense of 'crazy.' The fourth and final type is the metaphorical extension. An example is, from motion to change, and from location to state (Goldberg, 1995:78-83).

In comparison, Radical Construction Grammar effectively posits only a single link type, meronomic. A part-whole relation is an essential building block in linguistic theory. As the name implies, Croft reduces all connections between constructions to this single type. Resembling this position, Leino and Östman introduce the concept of meta-construction, templates which can be used to expand the constructional inventory of a language. The definition of a meta-construction is as follows. "They are not generalizations which only capture the similarities of a given group of constructions. Rather, they capture systematic similarities and differences which occur between several pairs of constructions." (Leino & Östman, 2005:207). Booij takes a similar position. He uses the term meta-construction to refer to a generalization that covers a wide range construction types (Booij, 2010:28-29). Another important aspect of the proposal by Leino and Östman is that a meta-construction is not simply considered to consist of a single variable which connects constructions. Instead, Leino and Östman postulate that a meta-construction may contain a number of variables which are readily available for extension (Leino & Östman, 2005:209). In this sense, a meta-construction is true construction type, not just a single value.[158]

Nonetheless, the definition of the meta-construction is established through pairs, methodologically akin to finding minimal pairs which may vary. In a situation, where a considerable stock of constructions is established, and their relations are explored, the assumption of establishing meta-constructions in a systematic manner may become infeasible as the interactions may become multifaceted instead of confining to pairs. Thus, the concept of Linking Construction is used which incorporates the basic principles of taxonomic relations among different constructions, and the nature of meta-constructions. A linking construction may possess multiple inherited variables (Goldberg, 1995:73). Effectively, we are giving up on the four linking types proposed by Goldberg. This is not to deny their existence but the types are far too fine-grained to be established in a data-driven manner without utilizing some specifically designed encoding schema in an attempt to capture them. Section 10.2.3 introduces a methodological solution in establishing links among argument constructions by utilizing the framework of classification in combination with Random Forests, and clustering.

Goldberg proposes two relevant principles for language organization: the principle of maximized expressive power, and the principle of maximized economy. These factors operate in different directions. The former would increase the number of objects to the point where every instance would consist of a unique label. The latter would posit only a single label to capture generalizations (Goldberg, 1995:67-69). The interaction of these principles can

[158] This statement is a simplification of their account as the proposal is grounded in the framework of CxG. A meta-construction defined relative to the Attribute Value Matrix.

be viewed as shaping language (cf. Bybee, 2010:18-19). Certainly, we do not know what an optimal solution between these principles might be. But, if generalizations can be captured with fewer objects the solution would be simpler. Furthermore, linking constructions are by definition abstractions, generalizations applicable to a wider range. They may serve to characterize particular instances in a more precise manner, because fewer relations are required to connect structures making the between-connections shorter, and tighter in a network. (cf. Croft & Cruse, 2004:74-75, 287-288).

A final principle is motivation. Goldberg builds on Lakoff's (1987) characterization for motivated structures that maximizes redundancy. For example, if two argument constructions share a number of properties the more redundant these become but the system is increasingly motivated. Consequently, forming extensions from known patterns becomes easier and faster as patterns are overlapped in a motivated system. Similarly, the incorporation of new information displays the same properties as the probability of a partial overlap between the old, and the new is increased (Goldberg, 1995:70-72). These observations blend with discussion on the correlation between variables in linguistics in Section 4. Thus, the method to form linking constructions meshes with the tenets of usage-based model by maximizing redundancy.

It is critical to emphasize that in the usage-based approaches, methodological issues are intertwined with theoretical ones. The applications of clustering cover a wide array of tasks, and have been recently utilized also in cognitive approaches, cf. Section 10.2.2. It is a standard practice in clustering that items are aggregated, in that all specific verb forms are fused into one. For example, all instances of *оказаться* 'seem, appear' would constitute a single object. In usage-based models, the gradient membership implies that all the usage patterns of *оказаться* 'seem, appear' are not necessarily equally good members. As was shown in the analysis of the argument constructions through Chapters 5–8, and even in the case where the model captures the argument constructions exhaustively, certain data points may still display slight fluctuation. This was illustrated with the estimated class probability plots, (e.g., the fluctuation with the Experiencer Extension Construction), cf. Section 6.3. Once the data are aggregated, this property of the usage-based model is lost. Geeraerts (2011) considers that this kind of lack of sensitivity is one of the major methodological issues in applying clustering to usage-based models. The employed methodology in this study is one possible way to increase sensitivity.

10.2.2 Unsupervised Learning: Clustering Algorithms

Dividing a data set into homogenous groups is perhaps the most basic task in linguistics, but with larger data sets, and a large number of variables, the process of finding groups in the data may become infeasible. Cluster analysis corresponds to techniques that are used to "discover" homogenous groups in data. For this reason, cluster analysis is considered to be unsupervised learning, because the data points are not labeled before the analysis. In essence, the main objective of cluster analysis is to find similar objects and group them together by minimizing the intra-cluster distance and maximizing the inter-cluster distance (Everitt, Landau, Leese & Stahl, 2011; Kaufman & Rousseeuw, 2005 [1990]).

Cluster analysis is widely used. For example, Schulte im Walde (2006) utilizes partitioning methods in studying the semantic structure of German verbs, and Gries and Stefanowitsch apply clustering to Ditransitive alternation, *into*-causative, *way*-construction in English. Dunn et al. examine the classification and historical development of 15 Papuan and 16 Austronesian languages through 125 binary grammatical features. Their cluster solution resembles the geographical map of the region (Dunn, Terrill, Reesink, Foley & Levinson, 2005). Cluster analysis has been applied to a certain extent to Russian data. For example, Divjak and Gries (2006) explore the near-synonymous verbs depicting *intending* and *trying*.[159]

One reason for the popularity of clustering methods is that they offer a way to visualize high-dimensional (cf. Hastie et al., 2009; Kaufman & Rousseeuw, 2005 [1990]). Due to the general applicability of clustering, a large number of different algorithms are available. They are typically divided into two large families: hierarchical, and partitioning methods. The former method builds a hierarchy of clusters. Initially, all the individual objects from separate clusters, and gradually are merged into a single cluster typically visualized as a tree. In contrast, the latter breaks the data into groups (Everitt et al., 2011; Hastie et al., 2009; Parashar & Vinay, 2008). Both of these methods assume that the investigated objects, constructions in this case, are encoded with a set of variables which can be used to determine the (dis)similarity between them. In essence, we are making a tacit assumption that distributional properties of constructions are comparable to semantic similarity. This view is in line with lexical semantics in Construction Grammar (cf. Bybee, 1985; Divjak & Gries, 2006; Goldberg, 2006), with Radical Construction Grammar (Croft, 2001), and generally, with certain branches of computational linguistics (cf. Schulte im Walde, 2006).

However, cluster analysis crucially depends on the used distance measure between objects which needs to be preselected. Perhaps the most popular quantification between two objects is the Euclidian distance, which is a true geometrical distance. To qualify as a distance, the quantification has to satisfy the following four criteria: 1) the distances are nonnegative numbers, 2) the distance of an object to itself is zero, 3) the distance is symmetrical, and 4) the direct distance between two objects is shorter than going through a third object (cf. Kaufman & Rousseeuw, 2005 [1990]:11-13). However, the chosen distance measure can radically alter the results (cf. Huang, 2008; Strehl & Ghosh, 2000). Instead of utilizing some well-known distance measure, dissimilarities can also

[159] Standard clustering procedures operate on a metric distance which is symmetrical. Tversky and Gati have pointed out that although clustering may offer a useful description of complex data its cognitive plausibility is not necessarily adequate as it does not incorporate the possible asymmetry of (dis)similarity (Tversky & Gati, 1978:81-82, 84, 95-98).

be used in clustering.

As Kaufman and Rousseeuw note, it is often assumed that dissimilarity measures fail to satisfy the fourth condition required for a distance measure. However, none of these properties are considered essential for a successful application of clustering. Dissimilarity measures that are nonnegative and small, indicate that the objects in question are similar, while larger values indicate dissimilarity (Kaufman & Rousseeuw, 2005 [1990]:16-17). However, the choice of selecting an appropriate distance or dissimilarity measure is less of an issue when the full potential of RF model is exploited. Thus, the proximity measure of the random forests is utilized in this study.

Shi and Horvart show that the attractiveness of the RF proximity measure originates from its data-driven property as it is based on the underlying tree predictors. An especially important property is its capability of handling mixed variables. This data set contains both categorical and numeric variables (Shi & Horvart, 2006:119, 134-135). Effectively, the binary splits of the trees discretize the numeric predictors but this is achieved in a data-driven manner.

The similarity measure for the labeled data is constructed in the following manner. First, each data point goes through the forests. Second, if data points i and j end up in the same terminal node, the similarity between them is increased by one. After the forest reaches its maximum number of trees, the similarities are summarized and divided by the number of the trees. Finally, the similarity of a data point to itself is set to one. Thus, the similarity measure is symmetric, positive, and ranges between the values of zero and one (Liaw & Wiener, 2002; Shi & Horvart, 2006:123). The proximity measure can be conceptually related to Construction Grammar. Instances that are located in the same terminal node are more similar to each other given the input.

The proximity measure is a similarity between objects as is indicated by the fact that the remoteness of an object to itself is one not zero as it is the case with standard dissimilarity measures. The RF similarity measure is extracted with the proximity() function and it can be converted to a dissimilarity measure which is $D = \sqrt{(1 - similarity)}$ (Shi & Horvart, 2006:123). As most distance measures are squared, it is also applied to forming the RF dissimilarity measure (cf. Kaufman & Rousseeuw, 2005 [1990]; Parashar & Vinay, 2008). A dissimilarity measure is obtained in a data-driven manner, and any clustering algorithm capable of functioning on dissimilarity measures can be applied.

Hierarchical methods are commonly employed in cognitive corpus linguistics (cf. Divjak & Gries, 2008; Gries & Stefanowitsch, 2010), and also in usage-based models (cf. Arppe, 2008).[160] When a large number of objects are clustered as is the case in this study (*n* = 1,878), the clustering may become difficult to interpret (Kaufman & Rousseeuw, 2005 [1990]:199-206). For this reason the

[160] Similarly to measuring the remoteness between objects a fairly large number of techniques are available to link objects in hierarchical clustering. These are discussed and their properties are described, for example, in Kaufman and Rousseeuw (2005 [1990]:225-242).

partitioning method is applied. The basic principle in partitioning methods is to divide a data set into *k* of clusters by starting with an initial clustering, and then iteratively reallocating the data points into the defined *k* clusters. An algorithm called partitioning around medoids (PAM), described in detail in Kaufman and Rousseeuw (2005 [1990]), is used in this study. In order to determine the data clusters, PAM tries to find representative objects in the data, called medoids. After that, the remaining objects are evaluated based on their (dis)similarity to these representative objects. The assignment of the remaining objects yields clusters.

Certainly, not every object qualifies as a representative object. A representative object minimizes the average dissimilarity of the other objects within the same cluster (Kaufman & Rousseeuw, 2005 [1990]:40-41). From a linguistics perspective, this can be regarded analogous to establishing exemplars and then grouping other objects based on their dissimilarity to them. Similar to hierarchical methods, partitioning methods have their caveats. First, the employed dissimilarity measure considerably affects the results. Second, the number *k* needs to be predefined. In the case of PAM, the number of *k* determines the number of representative objects and, at the same time, the number of clusters. PAM is implemented in R in the cluster package.

The results crucially depend on the previously outlined principles, (e.g., on the implement (dis)similarity measure). Thus, validation procedures to examine the solution are a vast research area even in linguistics applications (Moisl & Jones, 2005; Parashar & Vinay, 2008). The strongest evidence to support a certain clustering solution is external data (converging evidence). For example, Divjak and Gries (2008) validate their clustering results of nine Russian near-synonymous verbs of *trying* with experimental data. Another set of evaluation procedures consists of methods assessing the structure of the proposed clustering solution. Nonetheless, it should be remembered that even if a validation procedure is used, it does not automatically imply that the best possible solution has been obtained. Clustering typically finds some kind of structure in the data. Thus, a crucial component of evaluating a cluster solution is its meaningfulness. If results can be interpreted in a linguistically meaningful way, the cluster solution may be considered plausible.

10.2.3 Linking Constructions through RF Clustering

As outlined in the previous Section, the proximity measure was extracted from the RF model, and the similarity measure was transformed into the dissimilarity measure. The dissimilarity matrix consisted of all the data points in the model, a 1,878 by 1,878 matrix. Hence, the matrix is not shown. The question becomes whether it is possible to group the data points into smaller number of units, namely clusters. It is important to emphasize that the clustering used labeled data, (the argument constructions). In this sense, the clustering technique employed here can be considered to be semi-supervised.[161]

[161] The random forests algorithm can also be used for unsupervised learning. Shi and

A critical part in using PAM, or any other partitioning methods in general, is the preselection of *k*, the number of representative objects which also determine the number of clusters in PAM. Additionally, a cluster solution should have good properties. That is, it needs to be compact, well-separated, connected, and stable (cf. Brock, Pihur, Datta & Datta, 2008 and references therein). Most importantly, the cluster solution must be interpretable.

A standard method of evaluating a cluster solution is the silhouette, which is a measure of object's dissimilarity in relation to other objects in the solution. First, an average dissimilarity is calculated for an object within its cluster. Then, a neighbor for the object is defined by calculating the minimum average dissimilarity. A silhouette of an object is calculated in combination of these two values. The measure ranges between -1 and 1. Large positive silhouette values indicate a good clustering. The neighbor object establishs within, and between dissimilarities. A large silhouette value indicates that the within dissimilarity is much smaller compared to the between dissimilarity. The basic function of any classification is to group objects which are as similar as possible. The within dissimilarity between these objects is small, while the dissimilarity to other objects is larger, (i.e., the between dissimilarity).

The silhouette is quantification of the width. As each object has a silhouette value, an average can be calculated, and used to estimate the cluster solution. The higher the average silhouette width for the data set, the better the solution in terms of its structure according to this metrics (Kaufman & Rousseeuw, 2005 [1990]:84-85). The silhouette width is one possible way to initially evaluate a cluster solution. The silhouette is defined in relation to another object located in a different cluster. The minimum number of clusters is two, and the maximum is the number of objects. In this vein, the silhouette cannot be used to determine whether a clustering should be performed on the data or not. Figure 10.2-1 gives the average silhouette width for the data set based on *k* numbers varying between 2 and 17.

Horvart used this technique to improve the classification accuracy of the model on tumour samples by implementing the clusters as rules. (Shi & Horvart, 2006). The technique used here follows the redundancy principle instead of trying to discover unobserved structure in the data.

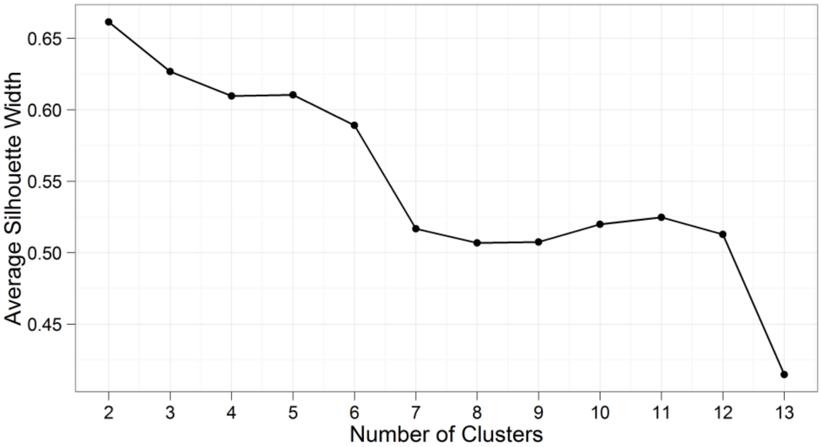


Figure 10.2-1 Average silhouette width for different cluster solutions (*k*=2-13) with the PAM algorithm using the RF dissimilarity matrix. Y-axis gives the average silhouette width for the cluster solutions and the x-axis represents the number of the clusters (medoids).

Exhaustive search through the dissimilarity matrix would consist of evaluating the possible values of k, ranging from 2 to 1,878. However, only a limited number of cluster solutions were considered (ranging from 2 to 13) covering the maximum number of the argument construction types in the RF model. The silhouette width shows a drastic drop after preselected *k* value five, indicating that a good cluster solution might be found between *k* values two, and five. The solution with five clusters was chosen. Although based on the silhouette width alone, the cluster solution with the *k* value of two is the best. Nonetheless, there is no drastic difference between the models containing either two or five clusters. Additionally, it will be shown that the latter solution possesses strong, and meaningful linguistically properties. Finally, the subsequent analysis will demonstrate that the variables underlying the five cluster solution yield, in practice, a perfect predictive accuracy, pointing to the conclusion that the solution appears to be appropriate. The silhouette plot illustrates the cluster solution, as given in Figure 10.2-2.

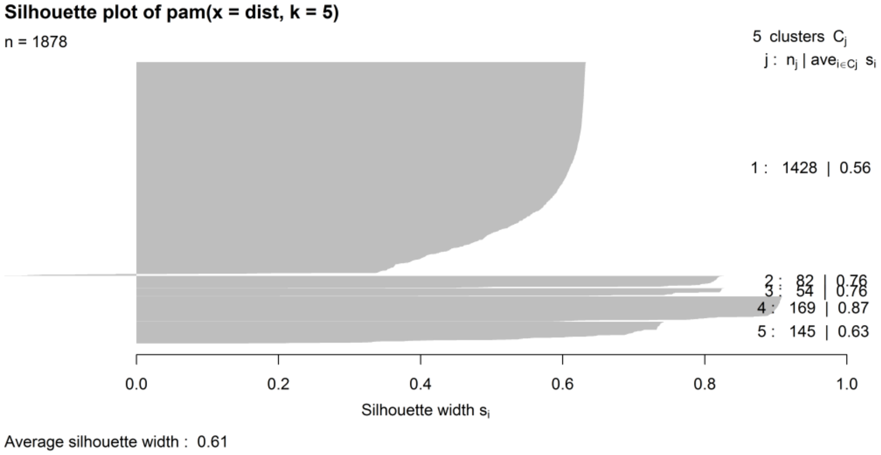


Figure 10.2-2 Silhouette plot of the PAM clustering based on five clusters (medoids) with the RF dissimilarity measure.

The data were partitioned into five clusters indicated with the number next to the silhouettes (right side of the plot). Additionally, the number of data points contained in the cluster, as well as the silhouette width of the cluster are displayed. The width of a cluster is the silhouette value, and the height is the number of objects in the cluster. This information is given in the plot with the string next to the silhouettes, for example, 1: 1,428 | 0.57. The string is the first cluster containing 1,428 instances with the average silhouette width of 0.56. The cluster solution shows that the majority of the data points are actually located in cluster 1. As a whole, cluster 1 appears to have a moderate structure estimated, based on the silhouette, because a small number of data points appear to be less prominently attached to this cluster as indicated by the silhouette curvature. In contrast, the remaining clusters appear to have a fairly strong structure. Overall, the solution based on the silhouette width indicates that a reasonable structure has been found.

It should be remembered that clustering methods are explanatory and additionally they do have their own underlying principles. Pam and most partitioning methods attempt to find spherical clusters (cf. Kaufman & Rousseeuw, 2005 [1990]; Schulte im Walde, 2006). A violation of this assumption may cause some data points to appear in a less optimal position. Due to the mechanics of how the random forests proximity measure is established, the structure tends to have a star shape (Hastie et al., 2009:595). Nonetheless, on average, the clusters are fairly pronounced. Keeping these caveats in mind, we can conclude that five linking constructions have been established. In terms of Construction Grammar, this has very little value in itself, as these clusters cannot be easily interpreted without connecting the clusters to the underlying variables in the model and the argument constructions types.

One possible solution to relate the variables in the model for the cluster is to use classification and regression trees (Shi & Horvart, 2006).[162] Because the variables are of mixed type, utilizing some univariate method to explore which cluster is associated with a particular variable becomes challenging. Additionally, we retain all the excellent properties of CART, as was outlined in Section 1.4.2. The conditional inferences trees were used to model the clusters, a polytomous categorical response variable, as the number of the cluster does not imply any inherent order. The variables used in the RF model were also used as predictors in this model.

In order to utilize the tools already used in this study, the method employed here demonstrates that the same principle is applicable to linguistics. Methodologically, this is a simple and elegant solution offering the same powerful visualization tools as the CART models. Furthermore, all required steps are committed within a single framework, namely classification. Additionally, as the CART builds the trees in a data-driven manner, possible interactions are automatically included. Thus, this is a superior method compared to utilizing a univariate method to evaluate the importance of the variables, as linguistic variables tend to form a relation of interconnectedness rather than a strong unique contribution.

Calculating *p*-values for the variables is perhaps not required, as clustering is an explanatory method. However, the *p*-values are used as a stopping criterion with the conditional inference trees. Importantly, the solution obtains virtually perfect classification accuracy (99.6%). Assuming that maximizing redundancy is the fundamental property of the network model, we can conclude that the implemented method is fully redundant. It fully captures the proposed set of argument constructions. The tree solution is given in Figure 10.2-3. The clusters are now referred to as linking constructions.

[162] Divjak and Gries (2006) explore the clustered structure of nine near-synonymous Russian verbs denoting trying by calculating *t*-values and *z*-scores.

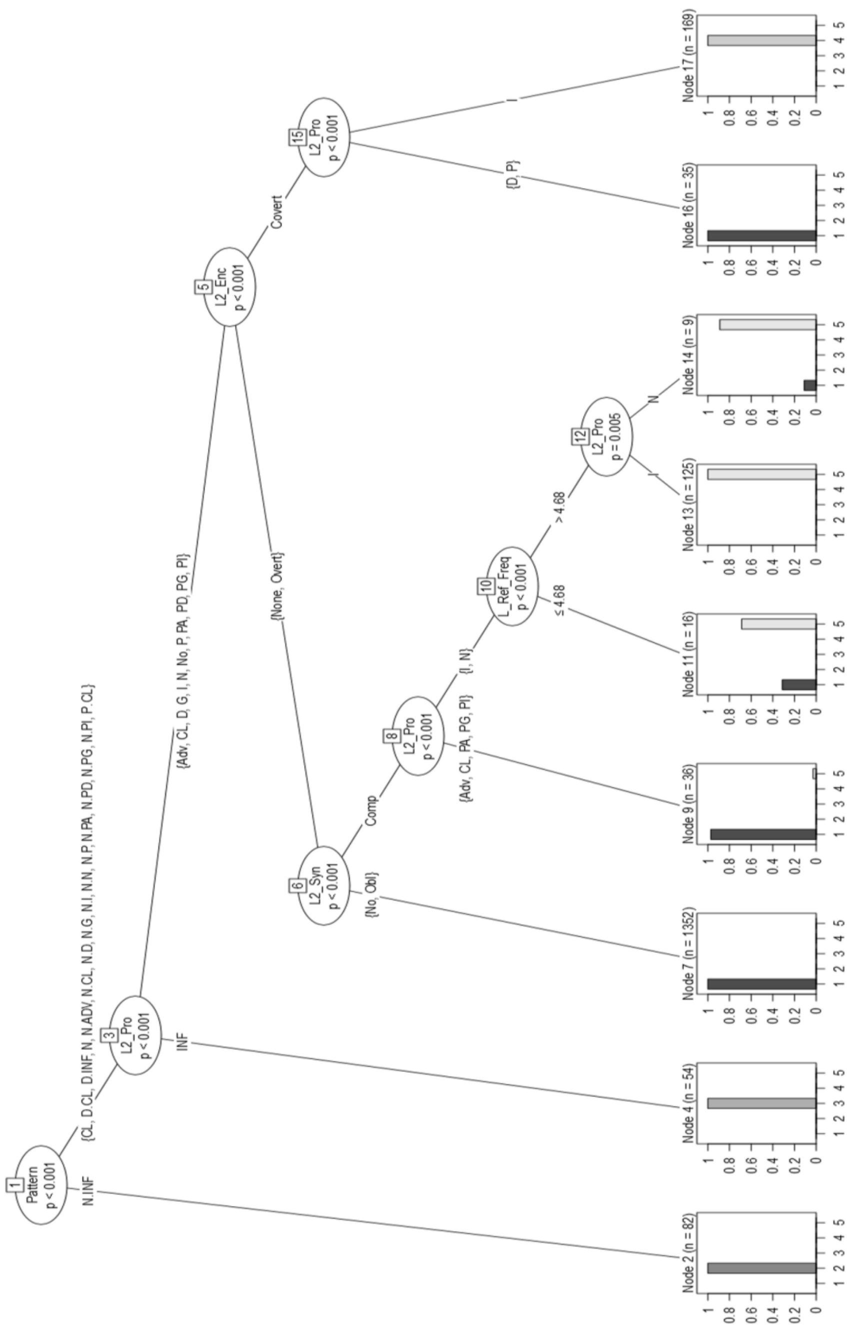


Figure 10.2-3 Classification tree of the clusters as a function of the predictors.

The linking constructions are located in the terminal nodes (TN) along with the number of data points. The prediction accuracy of the tree solution is visible in the terminal nodes, as they mostly consist of pure nodes that contain only a single linking construction type. The first split, node (N) 1, was established with the variable Pattern, which was divided into the Nominative-Infinitive pattern versus all other patterns. This subgroup was not split further. A terminal node (TN 2) represents the Linking Construction 2, consisting simply of the pattern Nominative-Infinitive exemplified with such reflexive verbs as *попытаться* 'try,' *приняться* 'proceed,' and *стремиться* 'strive, aim.' The partitioning continued in N 3 and the Pattern interacted with the Profile of the Secondary Slot. A split was established between the Infinitive versus all others. Terminal node (TN 4) is given that covers the instantiations of the Linking Construction 3. This Linking Construction is in interaction with the Dative-Infinitive pattern and the level Infinitive of the variable Profile of the Secondary Slot. Examples are reflexive verbs such as *приходиться* 'happen, have to,' *хотеться* 'want,' and *оставаться* 'stay, remain.'

All the remaining data points interacted with the variable Encoding of the Secondary Slot in N 5 partitioned into "None" and Overt versus Covert. The latter subgroup was further partitioned again with the Instrumental of the Profile of the Secondary Slot. In this vein, the terminal node (TN 17) represents a single type, Linking Construction 4, consisting of the Nominative-Instrumental pattern that interacts with covertly encoded Instrumental in the Secondary Slot. This type is supported with instantiations, such as *писаться* 'write,' *воспитываться* 'be raised,' and *добываться* 'obtain, procure.'

The terminal node (TN 26) gives a small subset of the instantiations of the Linking Construction 1. This illustrates that the method picks up small quirks in the data due to the sample size, such as the covertly encoded dative arguments (with such reflexive verbs as *нравиться* 'please, like' and *понадобиться* 'necessary') when they appear in the Nominative-Dative pattern.

The subgroup consisting of the None and Overt partitioned at N 5 was in interaction with the Syntactic Role of the Secondary Slot (L2_Syn) (N 6), dividing these data points into None and Oblique versus Complement. The former subgroup is given in Terminal node (TN 7), covering the vast majority of the instantiations associated with the Linking Construction 1. The latter group was in interaction in N 8 with the Profile of the Secondary Slot, and the subgroup was further partitioned into the subgroup consisting of the Instrumental and Nominative versus the remaining types. The latter subgroup was not partitioned further, and the terminal node (TN 9) is given containing a small subtype within the Linking Construction 1. These instantiations appear to be highly dissimilar to the other attested type in the sample, (e.g., *извиняться* 'apologize' with *that*-complement similar to *сомневаться* 'doubt,' *надеяться* 'hope,' and *бояться* 'be afraid.' Another set of reflexive verbs such as *становиться* 'become' and *оказаться* 'seem, appear,' seem to be dissimilar when the Complement is profiled with an Adverb.

The final subgroup was in interaction with the log Frequency of the Reflexive at N 10, and once again, with the Profile of the Secondary Slot at N 12. However, these types primarily pertain to the Linking Construction 5, (e.g., *оказаться* 'seem, appear,' *являться* 'be,' *становиться* 'become,' and *оставаться* ‘stay, remain’). For the more frequent verbs, those greater than 4.68 on log scale, the split in N 12 concerns the rare Nominative-Nominative pattern. At the same, time few instances of the Linking Construction 1 were located in terminal nodes 11 and 14: *получаться* ‘result, turn out,’ *прикидываться* ‘pretend to be,’ *славиться* ‘be famous,’ *представляться* ‘appear, seem,’ *прикинуться* ‘pretend,’ and *встречаться* ‘meet, occur, be.’

In sum, this section introduced the data-driven approach to derive linking construction, assuming that the network structure is grounded in maximizing redundancy. At the same this approach does not state the inheritance links in the following form:

$$\text{Argument Construction}_{\text{A}\overrightarrow{variable(s)}}\ \text{Argument Construction}_{\text{B}}$$

Instead the method yields links in the form:

$$\text{Argument Construction}_{\text{A}} \underset{variable(s)}{\Longleftrightarrow} \text{Argument Construction}_{\text{B}}$$

The link is shared between the argument constructions, but does not imply directionality. Additionally, the method can be readily applied to any variable-based model to form generalizations if the principle of maximizing redundancy is the target function. In sum, the five linking constructions were obtained from the data, offering a glimpse to the structure of the Russian Reflexive Marker. The next section anchors them to the proposed set of argument constructions enabling one to state generalizations and situate their location in the network.

10.2.4 Generalizations: Argument and Linking Constructions

The results suggested that there are five distinct regions within the network of the Russian Reflexive Marker. To situate the five Linking Constructions relative to the argument constructions, they are tabulated over the data points in the RF model given in Table 10.2-1. The argument constructions are located in the rows and the linking constructions in the columns.

	Linking Construction					
Argument Construction	**1**	**2**	**3**	**4**	**5**	Sum
Content	110	0	0	0	0	110
Existence	85	0	0	0	0	85
Experiencer	137	0	0	0	0	137
Experiencer Extension	0	0	54	0	0	54
Location	44	0	0	0	0	44
Motion	212	0	0	0	0	212
Passive	39	0	0	169	0	208
Phase	63	82	0	0	0	145
Property	26	0	0	0	145	171
Reciprocal	103	0	0	0	0	103
Reflexive Engagement	215	0	0	0	0	215
Spontaneous Event	329	0	0	0	0	329
Stimulus	65	0	0	0	0	65
Sum	1428	82	54	169	145	1878

Table 10.2-1 The linking constructions across the argument constructions. The number of the Linking Construction is given in bold.

Once the proposed set of the argument constructions are factored in, the connectivity through the Linking Constructions becomes apparent. The Linking Construction 2 connects to the Phase Construction through the Nominative-Infinitive pattern, whereas the remaining instantiations pertain to the Linking Construction 1. These results indicate that the Nominative-Infinitive pattern is dissimilar to all the other attested types in the sample.

Similar observation holds for the Linking Construction 3. The Dative-Infinitive pattern establishes a distinctive niche with the structure of the Russian Reflexive Marker, covering such verbs as *хотеться* 'want' and *доводиться* 'happen.' The Linking Construction 4 connects the majority of the instantiations of the Passive Construction, but a small group is attached to the Linking Construction 1. This result highlights the special status of the Passive Construction. The overtly encoded Secondary Slot is dissimilar to the covertly encoded subtype to the point that the former merges with the Linking Construction 1. A split is also present with the Linking Construction 4. As outlined in the previous section, a small group of verbs based on log frequency appear to be similar to the instantiations of the Linking Construction 1, (e.g., *прикидываться* 'pretend to be.' Finally, the Linking Construction 1 covers the remaining data points in the model.

The following labels can be attached to the linking constructions constituting the distinctive regions of the Russian Reflexive Marker in the network:

1) The Reflexive Intransitive Linking Construction covering $X_{Subject}\ Verb_{ся}\ Y_{Oblique}$ instantiations.
2) The canonical Phase Linking Construction covering $X_{Nominative}\ Verb_{ся}\ Y_{Infinitive}$ instantiations.
3) The Non-Canonical Subject Linking Construction covering $X_{Dative}\ Verb_{ся}\ Y_{Infinitive}$ instantiations.
4) The Canonical Passive Linking Construction covering

$X_{Nominative}$ $Verb_{ся}$ $Y_{Instrumental_{Covert}}$ instantiations.

5) The canonical Reflexive Property Linking Construction covering two main verb-specific construction types: $X_{Nominative}$ $V_{ся}$ $Y_{Instrumental}$ and $X_{Nominative}$ $Verb_{ся}$ $Y_{Nominative}$ instantiations.

The concept of the Linking Construction is inherently connected to the verb-specific, and the argument constructions occupying a relative position between them. Based on this, the network structure of the Russian Reflexive Marker can be depicted relative to these three levels of schematicity. Figure 10.2-4 illustrates where the reflexive verbs are connected to the argument constructions types. The connections between them are colored with the linking constructions, bringing forth the inherent interaction between them.[163]

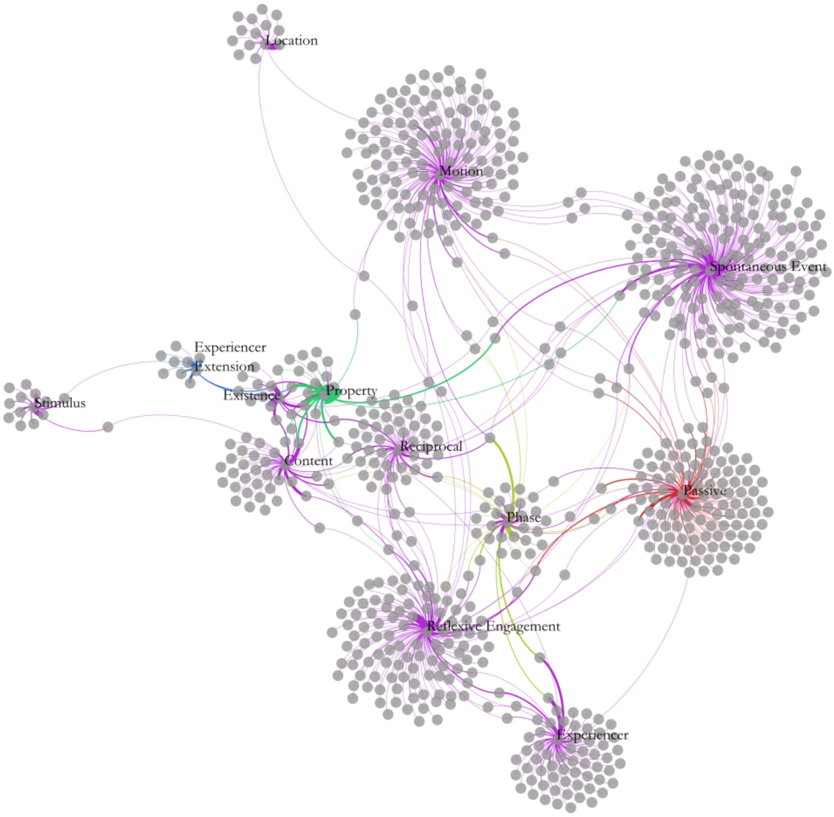


Figure 10.2-4 The network structure of the Russian Reflexive Marker through three levels of schematicity: the verbs (gray nodes) are connected to the argument constructions (labeled nodes) and the connections between them are colored with the linking constructions: Reflexive Intransitive (violet), Canonical Passive (red), Canonical Phase (yellow), Canonical Reflexive Property (green), and Non-Canonical Subject (blue).

[163] Gephi version 0.82 was used to build the graph. The software is freely available at https://gephi.org/.

In Figure 10.2-4, the reflexive verbs (gray nodes) are connected to the argument constructions (labeled nodes), and the connection between them is colored by the linking constructions. The interpretation of the model takes a turn back to the early Russian tradition, where intransitivity had a prominent role in describing the Russian Reflexive Marker. Examples include Vinogradov (Виноградов, 1972), Isachenko (Исаченко, 1960), and Shahmatov (Шахматов, 1925). In contrast, the intransitivity has been strongly opposed as a function of the Reflexive Marker. Examples include Kemmer (1993), Geniušienė (1987), and Knyazev (2007). Before turning to the generalizations that emerge from the proposed Linking Constructions, the Linking Construction 1, labeled as the Reflexive Intransitive, is not a function of the Reflexive Marker in a sense that it has been applied in previous studies. The Linking Constructions are distinctive regions in the network, a center that can attract new instatiations and support the existing types.

Related to the function of the Reflexive Marker, Kalashnikova and Saj (Калашникова & Сай, 2006) argue that the function of the Reflexive Marker is to profile the situation from a different perspective compared to the non-reflexive verb. A related position is also taken by Knyazev. He posits that the invariant function of the Reflexive Marker is to signal a change in semantic roles (Князев, 2007). In my view, both of these positions highlight important facets of the function of the Reflexive Marker as long as one small caveat is kept in mind. Knyazev (Князев, 2007) excluded, a priori,all the reflexive verbs that do not form pairs in the traditional sense to derive the invariant function. Similarly, Kalashnikova and Saj (Калашникова & Сай, 2006) removed 54% of the data to arrive at the function.

As the Liking Constructions are distinct regions in the network, they can be used to motivate distributional differences. As was outlined in Section 1.2.2, Krys'ko offers diachronic evidence on the interaction between transitivity and intransitivity in Russian. Once transitivity started to grammaticalize in Russian, certain groups of reflexive verbs, such as those of motion, gradually became intransitive (cf. Крысько, 2006:348-414). We know that grammaticalization evolves gradually and through different subgroups rather than globally (cf. Bybee, 2010; Levinson & Evans, 2010). This process fits the predictions of the model. Different regions in the network display sensitivity towards a specific property.

Another important diachronic property concerns the status of the dative subject connected to the Non-Canonical Linking Construction. Meyer (2010) argues that the dative subjects should be represented in the taxonomy of the Russian Reflexive Marker, as they are an integral part of the diachronic deveploment of it. This type appears to be distinctive among the data points. The estimated class probabilities were at ceiling, cf. Section 6.3. The data also suggest that the type frequency of the construction is fairly low and only a small proportion of reflexive verbs support it (cf. Divjak & Janda, 2008; Janda & Divjak, submitted). From a usage-based perspective, one would assume that the construction type would have undergone a change towards the canonical nominative subject. Instead, they have retained their distinctive properties. The estimated class probabilities point to the correct state of affairs. The structural properties of the argument construction type are salient and the usage patterns do not compete with other argument construction types, enabling them to retain their inherent position in the network (Geeraerts et al., 1994).

Finally, the results bring forth some of the "forgotten" types supported by the Canonical Phase Linking Construction with the Nominative-Infinitive pattern and the Canonical Reflexive Property Linking Construction. Both of them have received relatively little attention in the literature (although cf. Gerritsen, 1990). They are also among the most frequent reflexive verbs as was outlined in Section 3.1.7. The label Phase was taken from Langacker (2009), as he proposed it as an overarching term to cover the Nominative-Infinitive pattern. The importance of this linking construction is that it directly establishes another important pattern that crosses paradigms in Russian, in addition to the traditional relation between the Transitive and the Reflexive Constructions. Consequently, the Phase Linking Construction can be used to motivate such cross-paradigmatic mismatches as s *готовить* 'prepare' ~ *готовиться* 'prepare' and *собрать* 'gather' ~ *собраться* 'gather, intend.' The neighbor verbs do not combine with an infitinive. Thus, the infinitive pattern with the reflexive verbs is motivated through the network structure of the Reflexive Marker as an inherited property.

The Canonical Reflexive Copula Linking Construction can be used to bring forth the connectivity within the domain of copula constructions in Russian primarily associated with the verb *быть* 'be.' Additionally, the importance of this particular type can be connected to such a reflexive verb as *становиться*$_{\text{imp}}$ 'become.' According to Fasmer, the perfective verb is etymologically related to the Ancient Greek middle verb γίγνομαι 'become.'[164] Additionally, the cross-paradigmatic verbs *явить* 'be, emerge' ~ *явиться* 'be, emerge' are attested in Ancient Russian (Фасмер, 1986). The root form *-яв-* is in return connected to the canonical copula verb *являться* 'be,' forming the traditional aspectual pair *являться*$_{\text{imp}}$ ~ *явиться*$_{\text{perf}}$. Both of these reflexive verbs are also estimated to be more frequent, given on the original scale, than their neighbors *являться* 'be' (freq 522.9) ~ *являть* (freq 10) 'display, be' and *явиться* 'be, emerge' (freq 79.8) ~ *явить* 'be, emerge' (freq 6.9) in contemporary Russian.

Additionally, the Canonical Property Linking Construction can be used to motivate semantic drifts away from the cross-paradigmatic relations, such as *оказаться* 'seem, appear,' and *оказать* 'render.' According to Vinogradov, the change of the neighbor verb is connected to the formation of the compound predication type or phraseological unit consisting of an auxiliary-like predicate and a noun in the 18th century. Examples inlcude *иметь желание* 'wish, [lit. have wish]' ~ *желать* 'wish' and *принимать участие* 'participate, [lit. take participation]' ~ *участвовать* 'participate.' The neighbor verb appears primarily

[164] The verb form γίγνσθαι is given in Fasmer (Фасмер, 1986), (i.e., the middle/passive infinitive future).

in fixed expressions in contemporary Russian. These include *оказать услугу* 'serve [lit. render service]' ~ *услужить* 'serve' and *оказать помощь* 'help, [lit. render help]' ~ *помочь* 'help' (Виноградов, 1994). Thus, the diachronic pathways serve to demonstrate that this category is an integral part of the Russian Reflexive Marker and anchors the basic semantic relations established with the Property Construction (stative property 'be,' change of state 'become,' and perceived property 'seem, appear').

In sum, this section demonstrated the possible pathways that the proposed methodology can yield when applied to a variable-based model. The five linking constructions appear to be highly prominent properties of the Russian Reflexive Marker, and they can used to motivate deviations from the traditional pair account. Importantly, the connections are inherently dynamic, as they are based on the distributional properties of the lower level instantiations (cf. Bybee, 2010).

11 Discussion

This study has offered an implementation of usage-based theory to the Russian Reflexive Marker. Furthermore, the model builds on the assumption that constructions are the primary unit of language located on different levels of granularity. As this study concentrated on verbs, two levels of granularity figure prominently, namely the Verb-Specific and the Argument Constructions.

The following sections summarize the results, and pathways to future research are offered. In Section 11.1, the role of the verb in the model is discussed in conjunction with the variables proposed to model the gradient structure of categories. The variables of the model are discussed in Section 11.2 in relation to the Verb-Specific and the Argument Constructions. Finally, the three proposed domain-general principles are discussed in Section 11.3, connecting the implications that arose from the results to possible future diachronic, synchronic, and experimental studies.

11.1 Russian Reflexive Verbs and the Russian Reflexive Marker

The gradient structure of the reflexive and neighbor that was obtained from the data challenges the global directionality assumption postulated in the pair accounts between the cross-paradigmatic verbs, from the non-reflexive verb to the reflexive verb. Importantly, the cross-paradigmatic relation was operationalized with rigorously quantified variables. These include the Neighborhood Density, the Neighborhood Distance, the perceived semantic similarity (Reflexiva Tantum), and the log Frequency. First, the distributional difference in terms of the Neighborhood Density was not statistically significant, cf. Section 3.1.2. Similar results were observed for the log Frequency, cf. Section 3.1.7. The same distributional property appears to hold for the smaller, but semantically motivated partitions of the cross-paradigmatic relation, specifically the relation of Causation, cf. Section 3.1.10. Thus, data strongly suggest that the global directionality cannot be applied to the cross-paradigmatic relation. Instead, the cross-paradigmatic relation may pertain to subtle locality effects (cf. Bybee, 1985; Hay, 2001; 2002). For example, the following directionality should hold in terms of perceived basicness, as *пугать* ‘frighten’ → *пугаться* ‘become frightened’ contrasts *волноваться* ‘worry’ → *волновать* ‘agitate.’

In contrast, the results support the view that the reflexive verbs have a higher degree of specificity. The distributional difference of the Neighborhood Distance was statistically significant, cf. Section 3.1.3. This opens a perspective to the peculiar nexus of the Russian Reflexive Marker and the Reflexive Verb. On one hand, the Reflexive Marker is productive. On the other, the reflexive verbs tend to contain idiosyncratic properties. Observation that has been repeated throughout numerous studies (cf. Gerritsen, 1990; Israeli, 1997; Храковский, 1978a; Янко-Триницкая, 1962). These observations naturally follow from a density-based category. Due to the density of the Reflexive Marker, it can be readily applied to cover new instances, but the greater distances hinder the information flow across items, leading to the greater item-specificity. Consequently, specific argument construction types should be more prone to display partial productivity.

The results, however, need to be validated with separate studies and on a larger number of verbs, as the cross-paradigmatic relation covered 717 unique verbs in the sample. Nonetheless, the framework set forth here is readily applicable, as the theoretical basis is well-established for future studies on investigating cross-paradigmatic relations.

11.2 Verb-Specific and Argument Constructions

This study introduced the concept of the lexical network model in which lexical words are directly connected to each other, forming the Neighborhood in the lexicon. Consequently, the lexicon is assumed to be structured and this structuring influences abstractions. Additionally, the lexical network model is the lowest level of schematicity and abstractions arise over it. Specifically, three types of abstractions were investigated in this study: Verb-Specific Construction, Argument Construction, and Linking Construction. The idealized relations in the proposed model are given in Figure 11.2-1.

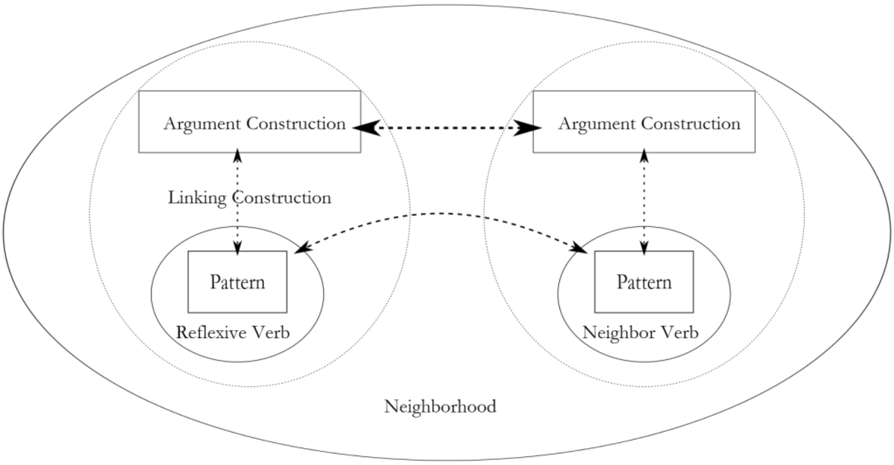


Figure 11.2-1 Idealized relations in the model.

The Linking Construction depicted with the lighter circle is assumed to be another form of abstraction, a distinct region in the network anchoring specific instantiations at various levels of granularity. However, the method employed here does not describe the exact inheritance of the variables. Instead, the method yielded links in the form as follows.

$$\text{Argument Construction}_{\text{A}} \underset{variable(s)}{\Longleftrightarrow} \text{Argument Construction}_{\text{B}}$$

Figure 11.2-1 also conveniently captures the direction for future studies, namely the structure of the Neighbor Verb. As the Construction is assumed to be the primary unit in language, the connections of the Neighbor Verb are certain to occupy an instrumental position in the formation of the various types marked with the Reflexive Marker. This is alluded to, for instance, in Vinogradov's characterization on the possible motivation for the drift between the cross-paradigmatic relation with the verbs *оказаться* 'seem, appear,' and *оказать* 'render.' The formation of the compound predication type or phraseological unit was discussed in Russian in Section 10.2.4.

The RF model suggested a fairly decent global performance of the model with the estimated classification accuracy of 82.2% to unseen data. The estimated classification accuracy, however, does not factor in the argument construction types that are "easier" to learn from the input. The performance of the model indicated that certain argument constructions did not readily emerge from the data. Two of them are especially important as they figure prominently in the literature: 1) the Experiencer Construction with the recall of 0.679 and precision of 0.753 and 2) the Motion Construction with the recall of 0.665 and precision of 0.61 (cf. Barðdal, 2008; Kemmer, 1993; Levin & Rappaport Hovav, 2005; Manney, 2000). It was argued that the semantics of these argument constructions might be more connected to the semantics of the individual verbs than to the set of the structural properties of the verb. Before turning to fine-grained semantic variables to model these verbs, the possibility to incorporate larger contextual factors is worth pursuing first, at least in my view. For example, the type of the modifier might appear as an important cue for specific argument constructions and the inclusion of ontological variables for the slots (cf. Arppe, 2008). Although WordNet is not publicly available for Russian, the manually tagged subcorpus of the Russian National Corpus contains a fairly fine-grained set of distinctions that could be exploited in future studies. This can be done effectively by combining the method of the behavioral profile analysis used in Cognitive Linguistics (cf. Gries, 2010a; Gries & Divjak, 2009) with the set of structural variables given in this study. The encoding of the verbs is publicly available and in machine readable format, as they can be straightforwardly implemented in a separate study.

As an exploratory study, the number of variables (p = 25) was fairly high in the RF model related to the argument constructions. Additionally, the potential importance of the variables could not be estimated based on previous studies because, to the best of my knowledge, statistical models have not been implemented for the Russian Reflexive Marker. Related to this, the ranking of the predictors in the RF model implemented another critical component of Construction Grammar, namely the relative importance of specific properties of the Argument Construction (Goldberg, 2006). Additionally, the ranking of the predictors can be partly motivated through the Relevance Hypothesis proposed by Bybee (1985). For example, the variable Person had low importance following the Relevance Hypothesis. This is excepted, as the reflexive verbs gravitate towards the third person singular.

In terms of surface generalizations, the variable Person has very little cue-validity because of the gravitation, but for the modulation of the cross-paradigmatic relation, this variable may offer interesting results in future studies.

Additionally, the six high ranked variables aligned with the results on other studies on predicting verb semantics or followed the tenets of previous studies, such as the Referent Type of the Subject Slot (L1_Ref). Consequently, the importance of the predictors is amiable for validation in separated studies.

11.3 Domain-General Principles

The two domain-general principles provide a basis for comparison between different studies, as they are not specifically related to any particular language or phenomenon. They come, however, with the assumption that the network structure is, at least partly, a property of the lexicon and high-order generalizations are mediated through the lexicon (Bybee, 2010). However, based on this assumption they are readily applicable to diachronic, synchronic, or experimental studies.

From a diachronic perspective, studies on grammaticalization or lexicalization are perhaps the most natural setting for the application of the principles (Bybee, 2010). The domain-general principles enable to quantify the gradient structure between categories as distances and number of connections between items. As grammaticalization tends to evolve in gradual changes through subsystems, the domain-general principles may offer a quantified method of modeling the directionality of grammaticalization (Bybee, 2010; Levinson & Evans, 2010). As frequency of use is one of the factors influencing grammaticalization, we would expect higher densities and shorter distances to facilitate the process even further. These predictions are grounded on the degree of connectivity, as was outlined in Section 3.1.3.

Another topic that was not addressed in this study is the formation of gaps between paradigms. The data suggest that the formation of gaps might be sensitive to differences in the structure of the lexicon in comparison to the cross-paradigmatic verbs that displays surprisingly similar properties. They are lexical density, frequency of use, and semantic similarity. In contrast, derivational accounts often formulate lexicalization is in opposition to grammaticalizatio. That leads to isolation of meaning due to the loss of the cross-paradigmatic relation, exemplified with such a reflexive verb as *бороться* 'fight,' which does not have a neighbor verb in contemporary Russian. The *бороть* is considered to be obsolete (cf. Вимер, 2001). Although it is reasonable to assume that the relative frequency is an important factor in the formation of gaps (cf. Hay, 2001), the domain-general principles might offer new possibilities to model this process in addition to the relative frequency.

From a synchronic perspective, the principles offer several new options to explore the connectivity in the lexicon. As the principles are grounded in the density-based perspective (cf. Baayen & Moscoso del Prado Martín, 2005), they might offer new possibilities to identify different verbs-specific constructions within an argument construction in addition to such well-established corpus-based methods as the analysis of co-occurrence (cf. Gries & Stefanowitsch, 2010). A second possible option is to model the senses of particular items. For example, testing whether the number of senses is sensitive to the principles in addition to frequency that is well established based on existing studies, cf. Section 3.1.5. In addition, their implementation also offers new methodological opportunities, specifically the utilization of graph-theory (cf. Steyvers & Tenenbaum, 2005; Vitevitch, 2008). Lexical densities, such as the rhyme neighbors, can be used to build a graph from the data, and graph-theory comes along with a set of well-establish metrics to operationalize the degree of connectivity even further. For example, Geeraert and Kyröläinen model eye-tracking data with graph-theoretical metrics built from the phonological neighbors of irregular past tense verbs in English (Geeraert & Kyröläinen, in prep.).

The most obvious choices for future experimental studies would be lexical decision tasks and similarity ratings. To the best of my knowledge, Russian reflexive verbs have not been studied from these perspectives. Lexical decision tasks would allow positioning of the domain-general principles in terms of early lexical access. Predictions for the lexical decision task are less than clear-cut based on the existing body of studies, cf. Section 3.1.2 (cf. Yarkoni & Balota, 2008). But, a higher lexical density should facilitate processing whereas longer distances should display an inhibitory effect. On the other hand, similarity ratings would offer a new perspective for the formation of "pairs" and how the degree relation is modulated. Based on previous studies, one would expect higher similarity ratings for densely populated items and a dissimilar effect should be present with the longer distance.

12 Conclusions

This study offered a new model of the Russian Reflexive situated in the probabilistic and usage-based framework, moving away from the "pair" model towards lexical structures and networks. Extensive discussion was given in fleshing out the role of the verb in theoretical linguistics covering usage-based theory, Construction Grammar and the diathesis tradition along with the influential derivational account by Rappaport Hovav and Levin (1998). As the proposed model builds primarily on usage-based theory and Construction Grammar, the concept of the Argument Construction was defined relative to the Verb-Specific Construction. This interconnection enabled the utilization of usage patterns and was extended to cover lexical networks. Because the latter type has not been incorporated as part of Construction Grammar but figures prominently in the studies of morphology (Bybee, 1985; Hay & Baayen, 2005; Plag & Baayen, 2010), a detailed discussion of its theoretical basis was offered along with the method of implementation.

The concept of the lexical network was defined as the Neighborhood and operationalized with rhyme verbs. The concept of Neighborhood pertains to network models that have become well-established in the studies of morphology and semantics. Once larger lexical structures are factored in, the concept of degree of connectivity between items can be operationalized. The concept of Neighbor Verb was offered to incorporate the inherent variability of language and the gradient structure that exists between paradigms. There are four important benefits in moving away from the "pair" model and positing the concept of Neighbor Verb. First, it comes with minimal a priori decisions required to establish them and prune the data. Second, it incorporates the traditional pair concept. Third, the "non-pairs" are maintained as part the system without the need to postulate a different mechanism specifically for them. Thus, a dialogue can be maintained with the diathesis tradition that has greatly influenced the formation of argument structures in linguistics (Geniušienė, 1987; Мельчук & Холодович, 1970; Падучева, 2004; Храковский, 1981). Fourth, the concept can be readily connected to a larger body of studies, thus enabling us to test convergence of evidence.

In terms of lexical items, the concept of Neighbor Verb opened a way to model hitherto underexplored structures of Russian verbs. The results suggest that the cross-paradigmatic verbs have strikingly similar densities in Russian. Three domain-general principles were formulated to motivate the resulting structure: the Hypothesis of Connectivity, Distance, and Gravity. Distributional differences in the network were used to support them, cf. Sections 3.1.1.–3.1.4, 3.1.5, 3.1.10 and 3.1.12. The results offer evidence that cross-paradigmatic relations are the preferred choice in the formation of complex categories. When cross-paradigmatic relations are used to connect different categories, the Connectivity between items is increased, whereas the Distance between items is decreased. These factors lead to more connected and tighter networks, which appear to be essential properties of complex networks in language (Steyvers &

Tenenbaum, 2005; Vitevitch, 2008). Thus, the domain-general principles can be used to answer the first question outlined in the Introduction. How are complex categories formed and maintained?

This study introduced random forest to model a polytomous response variable consisting of more than two outcomes. The process of implementing the random forests algorithm was covered in detail along with its tuning parameters. The current research on random forests suggests that they can be viewed as pertaining to weak and memory-based learnability. This aligns with the theoretical basis of Construction Grammar (Goldberg, 2006). There are typically multiple and even competing cues available in the input that guide the interpretation and the generalization processes. In contrast, a single cue in isolation is typically a weak predictor and interpretation requires support from additional cues, (i.e., weak learnability) (Arppe, 2008; Bresnan et al., 2007; Bresnan & Ford, 2010). Although random forests have a disadvantage in interpretability, they offer certain benefits. First, they can model a large number of predictors and the ranking of the predictors can be used for variable selection to reduce the set of potentially important variables. This reduced set can be used with other methods if more precise interpretation is required. Generally, the performance of the model indicates that the structural properties of the verb are utilized and partly involved in the process of forming abstractions. Thus, this provides the means for responding to the second question given in the Introduction. How can surface structures be used to form abstractions?

Another important feature of random forests is the proximity measure. This study demonstrated its application to forming the concept of the Linking Construction from the input. The obtained structure suggests that the method could also be used to investigate gradient category boundaries. For example, the Passive Construction was supported by two Linking Constructions based on the Encoding of the Secondary Slot as either covert or overt. The network structure of the argument constructions is the basic and fundamental descriptive tool in Construction Grammar, but the full-entry version has received very little attention. The formation of the network with the random forests was theoretically motivated using Goldberg's (1995) concept of maximizing redundancy. Thus, five Linking Constructions were established for the Russian Reflexive Marker forming distinctive regions in the network. This addresses the third question given in the Introduction. How are constructions, form-meaning pairings, interconnected in a network?

This study has been both theoretical and programmatic, but the components of the model were grounded in the usage-based theory of language. While cumulative and converging evidence is required to further anchor the results, the pathway is clear.

Bibliography

Abbott, B. (1997). Definiteness and existentials. *Language* 73(1): 103–108.

Ackerman, F. & Moore, J. (2009). Proto-properties and obliqueness. *'Case in and across Languages'*. Helsinki 27.–29.8.

Adamec, P. (1973). *Очерк функционально-трансформационного синтаксиса современного русского языка.* Praha: Statni pedagogicke nakladatelstvi.

Afonso, S. (2008). Existentials as impersonalising devices: The case of European Portuguese. *Transactions of the Philological Society* 106(2): 180–215.

Agresti, A. (2002). *Categorical data analysis.* New Jersey: John Wiley & Sons.

Ahn, H. (2005). *The semantics of SJA in Russian: Focus on the action*. Doctoral Dissertation, University of North Carolina at Chapel Hill.

Ahn, H. (2012). Semantic defocusing: semantically motivated syntax of Russian SJA constructions. *Russian Linguistics* 36(2): 193–211.

Altieri, N., Gruenenfelder, T. & Pisoni, D. B. (2010). Clustering coefficients of lexical neighborhoods: Does neighborhood structure matter in spoken word recognition? *Mental Lexicon* 5(1): 1–21.

Apresian, J., Boguslavsky, I., Leonid Iomdin, Lazursky, A., Sannikov, V., Sizov, V. & Tsinman, L. (2003). ETAP-3 linguistic processor: A full-fledged NLP 326 implementation of the MTT. *MTT 2003, First International Conference on Meaning – Text Theory.* Paris, École Normale Supérieure. 279–288.

Apresjan, J. D. (1974). Regular polysemy. *Linguistics* 12(142): 5–32.

Arbesman, S., Strogatz, S. H. & Vitevitch, M. S. (2010). Comparative analysis of networks of phonologically similar words in English and Spanish. *Entropy* 12(3): 327–337.

Arppe, A. (2008). *Univariate, bivariate, and multivariate methods in corpus-based lexicography: A study of synonymy.* Helsingin yliopisto: Yleisen kielitieteen laitoksen julkaisuja 44.

Arppe, A. & Baayen, R. H. (2012). Statistical classification and principles of human learning. In Zeldes, A. & Lüdeling, A. (Eds.), *4th conference on quantitative investigations in theoretical linguistics.* Humboldt-Universität zu Berlin.

Arppe, A., Gilquin, G., Glynn, D., Hilpert, M. & Zeschel, A. (2010). Cognitive corpus linguistics: five points of debate on current theory and methodology. *Corpora* 5(1): 1–27.

Arppe, A. & Järvikivi, J. (2007). Take empiricism seriously! In support of methodological diversity in linguistics. *Corpus Linguistics and Linguistic Theory* 3(1): 99–109.

Artstein, R. & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics* 34(4): 555–596.

Baayen, R. H. (2010a). Demythologizing the word frequency effect: A discriminative learning perspective. *Mental Lexicon* 5(3): 436–461.

Baayen, R. H. (2010b). A real experiment is a factorial experiment? *Mental Lexicon* 5(1): 149–157.

Baayen, R. H. (2011). Corpus linguistics and naive discriminative learning. *Revista Brasileira de Linguística Aplicada* 11(2): 295–328.

Baayen, R. H., Feldman, L. B. & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language* 55(2): 290–313.

Baayen, R. H. & Moscoso del Prado Martín, F. (2005). Semantic density and past-tense formation in three Germanic languages. *Language* 81(3): 666–698.

Babby, L. H. (1975). A transformational analysis of transitive *-sja* verbs in Russian. *Lingua* 35(3-4): 297–332.

Babby, L. H. (1983). The relation between causatives and voice: Russian vs. Turkish. *Wiener slawistischer Almanach* 11: 61–68.

Babby, L. H. & Comrie, B. (1980). *Existential sentences and negation in Russian.* Ann Arbor: Karoma.

Baerman, M. (2007). Morphological typology of deponency. In Baerman, M., Corbett, G. G., Brown, D. & Hippisley, A. (Eds.), *Deponency and morphological mismatches.* Oxford: Oxford University Press. 1–19.

Baker, C., Fillmore, C. J. & Lowe, J. B. (1998). The Berkeley Framenet Project. *Proceedings of the 17th International Conference on Computational Linguistics and the 36th Annual Meeting of the Association for Computational Linguistics*: 86–90.

Bakker, E. J. (1994). Voice, aspect and aktionsart: Middle and passive in Ancient Greek. In Hopper, P. J. & Fox, B. A. (Eds.), *Voice: Form and function.* Amsterdam and Philadelphia: Benjamins.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H. & Yap, M. J. (2004). Visual word recognition for single-syllable words. *Journal of Experimental Psychology: General* 133(2): 283–316.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B. & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods* 39(3): 445–459.

Barabási, A.-L. & Albert, R. (1999). Emergence of scaling in random networks. *Science* 286(509): 509–512.

Barðdal, J. (2004). The semantics of the impersonal construction in Icelandic, German and Faroese: Beyond thematic roles. In Abraham, W. (Ed.), *Focus on Germanic typology.* Berlin: Akademie Verlag. 105–137.

Barðdal, J. (2006). Construction-specific properties of syntactic subjects in Icelandic and German. *Cognitive Linguistics* 17(1): 39–106.

Barðdal, J. (2008). *Productivity: Evidence from case and argument structure in Icelandic.* Amsterdam: John Benjamins.

Barðdal, J. (2011). Lexical vs. structural case: A false dichotomy. *Morphology* 21.

Barðdal, J., Cennamo, M. & Eythórsson, T. (submitted). The rise and fall of anticausative constructions in Indo-European: The context of Latin and Germanic. In Kulikov, L. & Lavidas, N. (Eds.), *Typology of labile verbs: Focus on diachrony.*

Barðdal, J. & Eythórsson, T. (2003). Icelandic vs. German: Oblique subjects, agreement and expletives. *Chicago Linguistic Society* 39(1): 755–773.

Beavers, J. B., Levin, B., Rappaport Hovav, M. & Tham, S. W. (2010). A morphosyntactic basis for variation in the encoding of motion events. *Journal of Linguistics* 46(3): 331–377.

Bergen, B. & Wheeler, K. (2010). Grammatical aspect and mental simulation. *Brain and Language* 112(3): 150–158.

Bergen, B. K. & Chang, N. (2005). Embodied Construction Grammar in simulation-based language understanding. In Östman, J.-O. & Fried, M. (Eds.), *Construction Grammars: Cognitive grounding and theoretical extensions.* Amsterdam: John Benjamins. 147–190.

Berk, R. (2005). An introduction to ensemble methods for data analysis. *Department of Statistics Papers, Department of Statistics, UCLA*: 1–37.

Berk, R. (2008). *Statistical learning from a regression perspective.* New York: Springer.

Biau, G. (2012). Analysis of a random forests model. *Journal of Machine Learning Research* 13: 1063–1095.

Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing* 8(4): 219–241.

Blake, F. R. (Ed.) (1930). *A semantic analysis of case.* Baltimore: Linguistic Society of America.

Blevins, J. P. (2003). Passives and impersonals. *Journal of Linguistics* 39(3): 473–520.

Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics* 42(3): 531–573.

Bloomfield, L. (1939). Linguistic aspect of science. *International Encyclopedia of Unified Science* 1 and 2(4): 1–59.

Boas, H. C. (2011). Coercion and leaking argument structures in Construction Grammar. *Linguistics* 49(6): 1271–1303.

Bock, K., Loebell, H. & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review* 99(1): 150–171.

Bod, R. (2006). Exemplar-based syntax: How to get productivity from examples. *The Linguistic Review* 23(3): 275–290.

Bod, R. (2009). Constructions at work or at rest. *Cognitive Linguistics* 20(1): 129–134.

Booij, G. (2010). *Construction morphology.* Oxford: Oxford University Press.

Bostoen, K. & Nzang-Bie, Y. (2010). On how "middle" plus "associative/reciprocal" became "passive" in the Bantu A70 languages. *Linguistics* 48(6): 1255–1307.

Boulesteix, A.-L., Bender, A., Lorenzo Bermejo, J. & Strobl, C. (2012). Random forest Gini importance favours SNPs with large minor allele frequency: Impact, sources and recommendations. *Briefings in Bioinformatics* 13(3): 292–304.

Boyd, J. K. & Goldberg, A. E. (2011). Learning what not to say: the role of statistical preemption and categorization in "a"-adjective production. *Language* 87(1): 55–83.

Braginsky, P. & Rothstein, S. (2008). Vendlerian classes and the Russian aspectual system. *Journal of Slavic Linguistics* 16(1): 3–55.

Breiman, L. (1996a). Bagging predictors. *Machine Learning* 24(2): 123–140.

Breiman, L. (1996b). Out-of-bag estimation. Berkeley: Statistics Department, University of California Berkeley.

Breiman, L. (2001a). Random forests. *Machine Learning* 45(1): 5–32.

Breiman, L. (2001b). Statistical modeling: The two cultures. *Statistical Science* 16(3): 199–231.

Breiman, L. & Cutler, A. (2006). Random forests: Manual version 3.1.

Breiman, L., Friedman, J., Stone, J. V. & Olshen, R. A. (1984). *Classification and regression trees*. New York: Chapman and Hall.

Bresnan, J. (2007). Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Featherston, S. & Sternefeld, W. (Eds.), *Roots: Linguistics in search of its evidential base*. Berlin: Mouton de Gruyter. 77–96.

Bresnan, J., Cueni, A., Nikitina, T. & Baayen, R. H. (2007). Predicting the dative alternation. In Bouma, G., Krämer, I. & Zwarts, J. (Eds.), *Cognitive foundations of interpretations*. Royal Netherlands Academy of Arts and Sciences. 69–94.

Bresnan, J. & Ford, M. (2010). Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1): 168–213.

Brock, G., Pihur, V., Datta, S. & Datta, S. (2008). clValid: An R package for cluster validation. *Journal of Statistical Software* 25(4): 1–22.

Brown, D. & Hippisley, A. (Eds.) (2012). *Network morphology: A defaults-based theory of word structure*. New York: Cambridge University Press.

Brysbaert, M. & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods* 41(4): 977–990.

Budanitsky, A. & Hirst, G. (2006). Evaluating wordnet-based measures of lexical semantic relatedness. *Computational Linguistics* 32(1): 13–47.

Bühlmann, P. & Yu, B. (2002). Analyzing bagging. *The Annals of Statistics* 30(4): 927–961.

Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5): 425–455.

Bybee, J. L. (2001). *Phonology and language usage*. Cambridge: Cambridge University Press.

Bybee, J. L. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.

Bybee, J. L. & Moder, C. L. (1983). Morphological classes as natural categories. *Language*: 251–270.

Cardini, F.-E. (2008). Manner of motion saliency: An inquiry into Italian. *Cognitive Linguistics* 19(4): 533–569.

Chafe, W. (1992). The importance of corpus linguistics to understanding the nature of language. In Svartvik, J. (Ed.), *Directions in corpus linguistics: Proceedings of the Nobel Symposium 82, 4-8.8.1991*. Berlin: Mouton de Gruyter. 79–997.

Chan, K. Y. & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance* 35(6): 1934–1949.

Chesley, P. & Baayen, R. H. (2010). Predicting new words from newer words: Lexical borrowings in French. *Linguistics* 48(6): 1343–1374.

Chomsky, N. (1957). *Syntactic structures.* The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the theory of syntax.* Cambridge: MIT Pres.

Chomsky, N. (1970). Remarks on nominalizations. In Roderick, J. A. & Rosenbaum, P. S. (Eds.), *Readings in English Transformational Grammar.* Waltham, Massachusetts: Ginn and Company. 294–221.

Chomsky, N. (1981). *Lectures on goverment and binding.* Dordrecht: Foris.

Chomsky, N. (1986). The formal inadequacy of behavioristic theory. In Baars, B. J. (Ed.), *The cognitive revolution in psychology.* New york: Guilford Press. 338–350.

Chomsky, N. (2002). *On nature and language.* Cambridge: CUP.

Cohen, J., Cohen, P., West, S. G. & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences.* Mahwah: Lawrence Erlbaum Associates.

Colthearth, M., Davelaar, E., Jonasson, J. T. & Besner, D. (1977). Access to the internal lexicon. In Dornic, S. (Ed.), *Attention and performance VI.* Hillsdale and New Jersey: Lawrence Erlbaum Associates. 535–555.

Comrie, B. (1989). *Language universals and linguistic typology.* Chicago: University of Chicago Press

Corbett, G. G. (2011). The penumbra of morphosyntactic feature systems. *Morphology* 21(2): 445–480.

Cristofaro, S. (2005). *Subordination.* Oxford: Oxford University Press.

Croft, W. (1991). Syntactic categories and grammatical relations. The cognitive organization of information. Chicago and London: The University of Chicago Press.

Croft, W. (1998). Event structure in argument linking. In Geuder, M. B. W. (Ed.), *The Projection of Arguments.* Stanford: CSLI Publications. 21–63.

Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective.* Oxford: Oxford University Press.

Croft, W. (2003). Lexical rules vs. constructions: A false dichotomy. In Cuyckens, H., Berg, T., Dirven, R. & Panther, K.-U. (Eds.), *Motivation in language: Studies in honour of Günter Radden.* Amsterdam: John Benjamins. 1–43.

Croft, W. & Cruse, A. D. (2004). *Cognitive linguistics.* Cambridge: Cambridge University Press.

Cutler, R. D., Edwards, T. C. J., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology* 88(11): 2783–2792.

Dąbrowska, E. (2004). *Language, mind and brain: Some psychological and neurological constraints on theories of grammar.* Edinburgh: Edinburgh University Press.

Dąbrowska, E. (2008). The effects of frequency and neighbourhood density on adult speakers' productivity with Polish case inflections: An empirical test of usage-based approaches to morphology. *Journal of Memory and Language* 58(4): 931–951.

Daum, E. & Schenk, W. (1968). Die Russischen Verben. Leipzig: VEB Verlag Enzyklopäedie.

Davidse, K. & Heyvaert, L. (2007). On the middle voice: An interpersonal analysis of the English middle *Linguistics* 45(1): 37–83.

Davidson, A. C. & Hinkley, D. V. (1997). *Bootstrap methods and their application.* Cambridge: Cambridge University Press.

De Beule, J. & Steels, L. (2005). Hierarchy in fluid construction grammars. In Furbach, U. (Ed.), *Advances in Artificial Intelligence. Proceedings of the 28th German Conference on AI.* Springer. 1–15.

de Saussure, F. (1916). *Course in general linguistics.* New York, Toronto and London: McGraw-Hill.

de Vaan, L., Schreuder, R. & Baayen, R. H. (2007). Regular morphologically complex neologisms leave detectable traces in the mental lexicon. *The Mental Lexicon* 2(1): 1–24.

Díaz-Uriarte, R. & Alvarez de Andrés, S. (2006). Variable selection from random forests: application to gene expression data. *Bioinformatics* 7(3): 1–13.

Dickey, S. M. & Janda, L. A. (2009). Хохотнул, схитрил: the relationship between semelfactives formed with *-nu-* and *s-* in Russian. *Russian Linguistics* 33(3): 229–248.

Divjak, D. (2004). *Degrees of verb integration. Conceptualizing and categorizing events in Russian.* Doctoral Dissertation, Katholieke Universiteit Leuven.

Divjak, D. (2009). Mapping between domains. The aspect-modality interaction in Russian. *Russian Linguistics* 33(3): 249–269.

Divjak, D. & Gries, S. T. (2006). Ways of trying in Russian: clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2(1): 23–60.

Divjak, D. & Gries, S. T. (2008). Clusters in the mind? Converging evidence from near synonymy in Russian. *The Mental Lexicon* 3(2): 188–213.

Divjak, D. & Janda, L. A. (2008). Ways of attenuating agency in Russian. *Transactions of the Philological Society* 106(2): 138–179.

Dixon, R. M. W. (1991). *A semantic approach to English grammar.* New York: Oxford University Press.

Dowty, D. (1979). *Word meaning and Montague Grammar.* Reidel: Dordrecht.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language* 67(3): 547–619.

Dunn, M., Terrill, A., Reesink, G., Foley, R. A. & Levinson, S. C. (2005). Structural phylogenetics and the reconstruction of ancient language history. *Science* 309: 2072–2075.

Engdahl, E. (2006). Semantic and syntactic patterns in Swedish passives. In Solstad, T. & Lyngfelt, B. (Eds.), *Demoting the agent: Passive, middle and other voice phenomena.* Amsterdam and Philadelphia: John Benjamins. 21–45.

Enger, H.-O. & Nesset, T. (1999). The value of cognitive grammar in typological studies: The case of Norwegian and Russian passive, middle and reflexive. *Nordic Journal of Linguistics* 22: 27–60.

Eriksen, P. K. (2005). *On the typology of the semantics of non-verbal predication.* Oslo: Unipub AS.

Evans, V. & Green, M. C. (2006). *Cognitive linguistics: Introduction.* Edinburgh: Edinburgh University Press.

Everitt, B. S., Landau, S., Leese, M. & Stahl, D. (2011). *Cluster analysis.* London: Wiley.

Eythórsson, T. & Barðdal, J. (2005). Oblique subjects: A common Germanic inheritance. *Language* 81(4): 824–881.

Fagan, S. M. B. (1992). *The Syntax and semantics of middle constructions: A study with special reference to German.* Cambridge: cambridge University Press.

Faulhaber, S. (2011). *Verb valency patterns: A challenge for semantics-based accounts.* Berlin and New York: Mouton de Gruyter.

Fehrmann, D., Junghanns, U. & Lenertová, D. (2010). Two reflexive markers in Slavic. *Russian Linguistics* 34(3): 203–238.

Fillmore, C. J. (1968). The case for case. In Bach, E. & Harms, R. T. (Eds.), *Universals in linguistic theory.* New York: Holt. 1–90.

Fillmore, C. J. (1970). The grammar of 'hitting' and 'breaking'. In Jacobs, R. A. & S., R. P. (Eds.), *Readings in English Transformational Grammar.* Waltham: Ginn. 120–133.

Fillmore, C. J. (1982). Frame Semantics. In Korea, T. L. S. o. (Ed.), *Linguistics in the Morning Calm.* Seoul: Hanshin Publishing Company. 111–137.

Fillmore, C. J. (1983). How to know whether you're coming or going. In Rauh, G. (Ed.), *Essays on deixis.* Tübingen: Narr. 219–227.

Fillmore, C. J. (1999). Inversion and constructional inheritance. In Webelhuth, G., Koenig, J.-P. & Kathol, A. (Eds.), *Lexical and Constructional Aspects of Linguistics Explanation.* Standford: CSLI Publications. 113–128.

Fillmore, C. J. (2007). Valency issues in FrameNet. In Herbst, T. & Götz-Votteler, K. (Eds.), *Valency: theoretical, descriptive and cognitive issues.* Berlin and New York: Mouton De Gruyter. 129–160.

Fillmore, C. J., Kay, P. & O'Connor, C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64(3): 501–538.

Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. In Firth, J. R. (Ed.), *Studies in Linguistic Analysis.* Oxford: Blackwell. 1–32.

Flank, S. (1987). Phase subdivisions and Russian inceptives. *Die Welt der Slaven* 32(2): 310–316.

Forster, K. I. (2004). Category size effects revisited: Frequency and masked priming effects in semantic categorization. *Brain and Language* 90(1-3): 276–286.

Fox, J. & Weisberg, S. (2011). *An R companion to applied regression.* Los Angeles: SAGE.

Frajzyngier, Z. (2000). Domains of point of view and coreferentiality: System interaction approach to the study of reflexives. In Frajzyngier, Z. & Curl, T. S. (Eds.), *Reflexives, forms and functions.* Amsterdam and Philadelphia: John Benjamins. 125–152.

Freeze, R. (1992). Existentials and other locatives. *Language* 68(3): 553–595.

Freeze, R. (2001). Existential constructions. In Haspelmath, M., König, E., Oesterreicher, W. & Raible, W. (Eds.), *Language typology and language universals: An international handbook*. Berlin: Mouton de Gruyter. 941–953.

Fried, M. (2006). Agent back-grounding as a functional domain: Reflexivization and passivization in Czech and Russian. In Lyngfelt, B. & Solstad, T. (Eds.), *Demoting the agent: Passive, middle and other voice phenomena.* Amsterdam and Philadelphia: John Benjamins. 83–109.

Fried, M. & Östman, J.-O. (Eds.) (2004). *Construction grammar in a cross-language perspective.* Amsterdam and Philadelphia: John Benjamins.

Friedman, J. H. (2006). Comment: Classifier technology and the illusion of progress. *Statistical Science* 21(1): 15–18.

Garretson, G., O'Connor, M. C., Skarabela, B. & Hogan, M. (2004). Coding practices used in the project optimal typology of determiner phrases. Boston University.

Geeraert, K. & Kyröläinen, A.-J. (in prep.). Paradigmatic Levelling in English: The Effect of Phonological Neighbours.

Geeraerts, D. (1988). Where does prototypicality come from? In Rudzka-Ostyn, B. (Ed.), *Topics in cognitive linguistics.* Amsterdam and Philadelphia: John Benjamins. 207–229.

Geeraerts, D. (1992). Polysemy and prototypicality. *Cognitive Linguistics* 3(2): 219–231.

Geeraerts, D. (2010). *Theories of lexical semantics.* New york: Oxford University Press.

Geeraerts, D. (2011). There are many colourful parasols in China, or the virtues of quantitative corpus semantics. *The Third Finnish-Estonian Cognitive Linguistics Conference.* Saka.

Geeraerts, D. & Cuyckens, H. (Eds.) (2007). *The Oxford handbook of cognitive linguistics.* Oxford and New York: Oxford University Press.

Geeraerts, D., Grondelaers, S. & Bakema, P. (1994). *The Structure of Lexical Variation: Meaning, naming and context.* Berlin and New York: Mouton de Gruyter.

Geeraerts, D., Kristiansen, G. & Peirsman, Y. (Eds.) (2010). *Advances in cognitive sociolinguistics.* Berlin and New York: Mouton de Gruyter.

Gelman, A. & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

Geniušienė, E. (1987). *The typology of reflexives.* Berlin: Mouton de Gruyter.

Genzel, D. & Charniak, E. (2002). Entropy rate constancy in text. *Proceedings of the 40th annual meeting of the Association for Computational Linguistics (ACL).* Philadelphia. 199–206.

Genzel, D. & Charniak, E. (2003). Variation of entropy and parse trees of sentences as a function of the sentence number. In Collins, M. & Steedman, M. (Eds.), *Proceedings of the Conference on Empirical Methods in Natural Language Processing.* Sapporo. 65–72.

Gerritsen, N. (1990). *Russian reflexive verbs. In search of unity in diversity.* Amsterdam: Rodopi.

Givón, T. (1979). On understanding grammar. New York: Academic Press.

Givón, T. (1994). The pragmatics of de-transitive voice: Functional and typological aspects of inversion. In Givón, T. (Ed.), *Voice and inversion.*

Amsterdam: John Benjamins. 3–44.

Glazkova, A. (2011). *A construction grammar approach to Russian reflexives*. University of Washington.

Glynn, D. (2010). Corpus-driven cognitive semantics. An introduction to the field. In Glynn, D. & Fisher, K. (Eds.), *Quantitative Cognitive Semantics. Corpus-driven approaches.*

Goddard, C. & Wierzbicka, A. (Eds.) (2002). *Meaning and universal grammar: Theory and empirical findings.* Amsterdam/Philadelphia: John Benjamins.

Goldberg, A. E. (1995). *Constructions. A construction grammar approach to argument structure.* Chicago: The University of Chicago Press.

Goldberg, A. E. (2002). Surface generalizations: An alternative to alteration. *Cognitive Linguistics* 13(4): 327–356.

Goldberg, A. E. (2004). But do we need universal grammar? Comment on Lidz et al. *Cognition* 94(1): 77–84.

Goldberg, A. E. (2006). *Constructions at work. The nature of generalization in language.* Oxford: Oxford University Press.

Goldberg, A. E. (2009a). Constructions work. *Cognitive Linguistics* 20(1): 201–224.

Goldberg, A. E. (2009b). The nature of generalization in language. *Cognitive Linguistics* 20(1): 93–127.

Goldberg, A. E. (2010). Verbs, constructions and semantic frames. In Rappaport Hovav, M., Doron, E. & Sichel, I. (Eds.), *Syntax, lexical semantics and event structure.* Oxford: Oxford University Press. 39–58.

Goldberg, A. E. (2011). Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics* 22(1): 131–153.

Goldsmith, J. (1998). On information theory, entropy, and phonology in the 20th century. *Folia Linguistica* 1(2): 85–100.

Gooskens, C. & Heeringa, W. (2004). Perceptive evaluation of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change* 16(3): 189–207.

Grannes, A. (1984). Short form adjectives and participles as facultative objective predicatives in 18th century Russian. *Russian Linguistics* 8(1): 17–25.

Greenberg, G. R. & Franks, S. (1991). A parametric approach to dative subjects and the second dative in Slavic. *Slavic and East European Journal* 35(1): 71–97.

Gries, S. T. (2002). Evidence in linguistics: three approaches to genitives in English. In Brend, R. M., Sullivan, W. J. & Lommel, A. R. (Eds.), *LACUS Forum XXVIII: what constitutes evidence in linguistics?* Fullerton: CA: LACUS. 17–31.

Gries, S. T. (2006). Corpus-based methods and the cognitive semantics: The many senses of to run. In Gries, S. T. & Stefanowitsch, A. (Eds.), *Corpora in Cognitive Linguistics.* Berlin and New York: Mouton de Gruyter. 57–99.

Gries, S. T. (2007). New perspectives on old alternations. In Cihlar, J. E., Franklin, A. L. & Kaiser, D. W. (Eds.), *Papers from the 39th Regional Meeting of the Chicago Linguistics Society: Vol. II. The Panels.* Chicago: Chicago Linguistics Society. 274–292.

Gries, S. T. (2010a). Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *Mental Lexicon* 5(3): 323–346.

Gries, S. T. (2010b). Corpus linguistics and theoretical linguistics: A love-hate relationship? Not necessarily…. *International Journal of Corpus Linguistics* 15(3): 327–343.

Gries, S. T. & Divjak, D. (2009). Behavioral profiles: A corpus-based approach to cognitive semantic analysis. In Evans, V. & Pourcel, S. S. (Eds.), *New directions in cognitive linguistics*. Amsterdam and Philadelphia: John Benjamins. 57–75.

Gries, S. T. & Stefanowitsch, A. (2010). Cluster analysis and the identification of collexeme classes. In Newman, J. & Rice, S. (Eds.), *Empirical and experimental methods in cognitive/functional research*. Stanford: CSLI. 73–90.

Grimshaw, J. B. (1990). *Argument structure*. Cambridge: MIT Press.

Grondelaers, S., Speelman, D. & Geeraerts, D. (2007). Lexical variation and change. In Geeraerts, D. & Cuyckens, H. (Eds.), *The Oxford handbook of cognitive linguistics* Oxford and New York: Oxford University Press. 988–1011.

Gruber, J. S. (1965). *Studies in lexical relations.Ph.D. dissertation.* MIT.

Guhl, M. (2010). Towards a syntactic analysis of Russian *-sja*. *Russian Linguistics* 3: 261–283.

Guyon, I. & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research* 3: 1157–1182.

Haiman, J. (1983). Iconic and economic motivation. *Language* 59(4): 781–819.

Haiman, M. H. (1991). *Grammatical voice*. Cambridge: Cambridge University Press.

Hand, D. J. (2006). Classifier technology and the illusion of progress. *Statistical Science* 21(1): 1–14.

Harrell, F. E. (2001). *Regression modeling strategies*. Berlin: Springer.

Harwood, F. W. & Wright, A. M. (1956). Statistical study of English word formation. *Language* 32(2): 260–273.

Haspelmath, M. (1994). Passive participles across languages. In Fox, B. A. & Hopper, P. J. (Eds.), *Voice: Form and Function*. Amsterdam and Philadelphia: Benjamins. 151–177.

Haspelmath, M. (2004). Does linguistic explanation presuppose linguistic description? *Studies in Language* 28(3): 554–579.

Haspelmath, M. & Müller-Bardey, T. (2005). Valency change. In Booij, G., Lehmann, C. & Mugdan, J. (Eds.), *Morphology: A handbook on inflection and word formation*. Berlin: Mouton de Gruyter. 1130–1145.

Hastie, T., Tibshirani, R. & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. New York: Springer.

Hay, J. B. (2001). Lexical frequency in morphology: Is everything relative? *Linguistics* 39(6): 1041–1070.

Hay, J. B. (2002). From speech perception to morphology: Affix ordering revisited. *Language* 78(3): 527–555.

Hay, J. B. & Baayen, R. H. (2005). Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences* 9(7): 3432–3348.

Hay, J. B. & Bresnan, J. (2006). Spoken syntax: The phonetics of *giving a hand* in

New Zealand English. *The Linguistic Review* 23(3): 321–349.

Hay, J. B., Nolan, A. & Drager, K. (2006). From *fush* to *feesh*: Exemplar priming in speech perception. *The Linguistic Review* 23(3): 351–379.

Helasvuo, M.-L. & Kyröläinen, A.-J. (2010). A Random Forest Model for Contextually Induced Person Marking Strategies. *Conference: Contexts of Language: How to Analyze Context?* Helsinki.

Helasvuo, M.-L. & Kyröläinen, A.-J. (2011). Ilmisubjekti vai nolla? Syntaktisen variaation kontekstuaaliset piirteet tarkastelussa. *Symposium: Käyttö-pohjaisia näkökulmia kielioppiin.* Tartto.

Hengeveld, K. (1992). *Non-verbal predication: Theory, typology, diachrony.* Berlin and New York: Mouton de Gruyter.

Herbert, S. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society* 106(6): 467–482.

Herbst, T. (2010). Valency constructions and clause constructions or how, if at all, valency grammarians might sneeze the foam off the cappuccino. In Schmid, H.-J. & Handl, S. (Eds.), *Cognitive foundations of linguistic usage patterns: Empirical studies.* Berlin and New York: Mouton de Gruyter.

Hino, Y., J, L. S. & Pexman, P. M. (2002). Ambiguity and synonymy effects in lexical decision, naming, and semantic categorization tasks: Interactions between orthography, phonology, and semantics. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28(4): 686.

Holcomb, P. J., Grainger, J. & O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience* 14(6): 938–950.

Hopper, P. J. & Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language* 56(2): 251–299.

Horálek, K. (1979). *Русская грамматика. Том 1.* Praha: Academia.

Hothorn, T., Bühlmann, P., Dudoit, S., Molinaro, A. & Van Der Laan, M. J. (2006a). Survival ensembles. *Biostatistics* 7(3): 355–373.

Hothorn, T., Hornik, K. & Zeileis, A. (2006b). Unbiased recursive partitioning: A conditional inference framework. *Journal of Computational and Graphical Statistics* 15(3): 651–674.

Howes, D. H. & Solomon, R. L. (1951). Visual duration threshold as a function of word probability. *Journal of Experimental Psychology* 41(6): 401–410.

Huang, A. (2008). Similarity measures for text document clustering. *Proceedings of the Sixth New Zealand Computer Science Research Student Conference.* New Zealand. 49–56.

Hunston, S. & Francis, G. (2000). *Pattern grammar: a corpus-driven approach to lexical grammar of English.* Amsterdam: John Benjamins.

Huumo, T. (1996). Bound Spaces and the Semantic Interpretation of Existentials. *Linguistics* 34(2): 295-328.

Huumo, T. (2003). Incremental existence: The world according to the Finnish existential sentence. *Linguistics* 41(3): 461–493.

Huumo, T. (2005). How fictive dynamicity motivates aspect marking: The riddle of the Finnish quasi-resultative construction. *Cognitive Linguistics* 16(1): 133–144.

Hwang, J. D., Nielsen, R. D. & Palmer, M. (2010). Towards a domain independent semantics: Enhancing semantic representation with construction grammar. In Sahlgren, M. & Knutsson, O. (Eds.), *Proceedings of the NAACL HLT workshop on extracting and using constructions in computational linguistics*. NAACL HLT 2010. 1–8.

Hyndman, R., J. & Fan, Y. (1996). Sample quantiles in statistical packages. *The American Statistician* 50(4): 361–365.

Isachenko, A. V. (1974). On 'Have' and 'Be' languages. In Flier, M. S. (Ed.), *Slavic forum: Essays on linguistics and literature*. Mouton: The Hague. 43–77.

Israeli, A. (1997). *Semantics and pragmatics of the "Reflexive" verbs in Russian.* München: Slavistische Beiträge.

Iwata, S. (2008). *Locative alternation: A lexical-constructional approach.* Amsterdam and Philadelphia: John Benjamins.

Jackedoff, R. (1997). Twistin' the night away. *Language* 73(3): 534–559.

Jackendoff, R. (1990). *Semantic structures.* Cambridge: The MIT Press.

Jakobson, R. (1989 [1932]). Structure of the Russian verb. In Waugh, L. R. & Halle, M. (Eds.), *Russian and Slavic grammar. Studies 1931-1981.* Berlin and New York: Mouton Publishers. 1–14.

Jakobson, R. (1989 [1936]). Contribution to the general theory of case: General meanings of the Russian cases. In Waugh, L. R. & Halle, M. (Eds.), *Russian and Slavic grammar. Studies 1931-1981.* Berlin and New York: Mouton Publishers. 59–103.

Janda, L. A. (1986). *A semantic analysis of the Russian verbal prefixes: za-, pere-, do-, and ot-*. München: Verlag Otto Sagner.

Janda, L. A. (1993a). Cognitive linguistics as a continuation of the Jacobsonian tradition:The semantics of Russian and Czech reflexives. *American Contributions to The Eleventh International Congress of Slavists.* Columbus: Slavica. 310–319.

Janda, L. A. (1993b). *A geography of case Semantics. The Czech dative and the Russian instrumental.* Berlin and New York: Mouton de Gruyter.

Janda, L. A. (2007). Aspectual clusters of Russian verbs. *Studies in Language* 31(3): 607–648.

Janda, L. A. (2008a). Motion verbs and the development of aspect in Russian. *Scando-Slavica* 54(1): 179–197.

Janda, L. A. (2008b). Transitivity in Russian from a cognitive perspective. In Kustova, G. (Ed.), *Dinamičeskie modeli: Slovo. Predloženie. Tekst. Sbornik statej v čest' E. V. Padučevoj.* Moskva: Jazyki slavjanskoj kul'tury. 970–988.

Janda, L. A. & Divjak, D. (submitted). The role of non-canonical subjects in the overall grammar of a language: A case study of Russian. In Helasvuo, M.-L. & Huumo, T. (Eds.), *Canonical and non-canonical subjects in constructions.* Amsterdam: John Benjamins.

Janda, L. A. & Lyashevskaya, O. (2011). Grammatical profiles and the interaction of the lexicon with aspect, tense and mood in Russian. *Cognitive Linguistics* 24(4): 719–763.

Janda, L. A. & Nesset, T. (2010). Taking apart Russian RAZ-. *Slavic and East European Journal* 54(3): 476–501.

Jespersen, O. (1924). *The philosophy of grammar.* London: George Allen & Unwin.

Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination and* *reasoning*. Chicago: University of Chicago Press.

Kaufman, L. & Rousseeuw, P. J. (2005 [1990]). *Finding groups in data: An introduction to cluster analysis*. New Jersey: John Wiley & sons Inc.

Kay, P. (2002a). An informal sketch of a formal architecture for construction grammar. *Grammars* 5(1): 1–19.

Kay, P. (2002b). Patterns of coining. *http://www.icsi.berkeley.edu/~kay/coining.pdf. [20.03.2010]*.

Kay, P. & Fillmore, C. J. (1999). Grammatical constructions and linguistic generalizations: the What's X doing Y? construction. *Language* 75(1): 1–33.

Kazanina, N. (2011). Decomposition of prefixed words in Russian. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(6): 1371–1390.

Keenan, E. L. (1976). Towards a universal definition of "subject". In Li, C. N. (Ed.), *Subject and Topic*. New York: Academic Press. 303–333.

Keller, F. (2004). The entropy rate principle as a predictor of processing effort: An evaluation against eye-tracking data. In Dekang, L. & Dekai, W. (Eds.), *Proceedings of the conference on Empirical methods in natural language processing*. Barcelona: Association for Computational Linguistics. 317–324.

Kemmer, S. (1993). *The middle voice*. Amsterdam and Philadelphia: John Benjamins.

Kempe, V. & MacWhinney, B. (1999). Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes* 14(2): 129–171.

Kepser, S. & Reis, M. (2005). Evidence in linguistics. In Kepser, S. & Reis, M. (Eds.), *Linguistic evidence: Emperical theoretical and computational perspectives*. Berlin and New York: Mouton de Gruyter. 1–25.

Kim, H. & Loh, W.-Y. (2001). Classification trees with unbiased multiway splits. *Journal of the American Statistical Association* 96(454): 589–604.

Kipper-Schuler, K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. Doctoral Dissertation, University of Pennsylvania.

Knjazev, J. P. (2007). Reciprocal constructions in Russian. In Nedjalkov (Ed.), *Reciprocal constructions*. Amsterdam: John benjamins. 673–708.

Koffka, K. (1935). *Principles of gestalt psychology*. New York: Harcourt, Brace & World.

Köhler, R. (1986). *Zur linguistischen Synergetik: Struktur und Dynamik der Lexik*. Bochum: Brockmeyer.

König, E. & Gast, V. (Eds.) (2008). *Reciprocals and reflexives: Theoretical and typological explorations*. Berlin and New York: Mouton de Gruyter.

Krasovitsky, A., Baerman, M., Brown, D. & Corbett, G. G. (2011). Changing semantic factors in case selection: Russian evidence from the last two centuries. *Morphology* 21(3): 573–592.

Krasovitsky, A., Long, A., Baerman, M., Brown, D. & Corbett, G. G. (2008). Predicate nouns in Russian. *Russian Linguistics* 32(2): 99–113.

Kruskal, J. B. (1983). An overview of sequence comparison: time warps, string edits, and macromolecules. *SIAM Review* 25(2): 201–237.

Kuperman, V., Stadthagen-Gonzalez, H. & Brysbaert, M. (in pres.). Age-of-acquisition ratings for 30 thousand English words. *Behavior Research Methods.*

Kutas, M. & Hillyard, S. A. (1980). Brain potentials during reading reflect word expectancy and semantic association. *Nature* 307(307): 161–163.

Kyröläinen, A.-J. (2008). Low-frequency constructions and salience: A case study on Russian verbs of motion of dative impersonal construction type. In Mustajoki, A., Kopotev, M. V., Birjulin, L. A. & Protasova, E. J. (Eds.), *Instrumentarij rusistiki: korpusnye podhody.* Helsinki: Department of Slavonic and Baltic Languages and Literatures. 176–197.

Kyröläinen, A.-J. (2012). Multiple valency frames: towards a density-based network model. *SLE 2012.* Stockholm.

Kyröläinen, A.-J. (submitted). From canon and monolith to clusters: A constructionist model of Subjecthood. In Helasvuo, M.-L. & Huumo, T. (Eds.), *Canonical and non-canonical subjects in constructions.* Amsterdam: John Benjamins. 47 pp.

Lagus, K., Kohonen, O. & Virpioja, S. (2009). Towards unsupervised learning of constructions from text. In Sahlgren, M. & Knutsson, O. (Eds.), *Proceedings of the workshop on extracting and using Constructions in NLP of 17th Nordic Conference on Computational Linguistics.*

Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind.* Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. (1980). *Metaphors we live by.* Chicago: University Press of Chicago.

Langacker, R. W. (1987). *Foundations of cognitive grammar. Theoretical prerequisites.* Stanford: Stanford University Press.

Langacker, R. W. (1988a). The nature of grammatical valence. In Rudzka-Ostin, B. (Ed.), *Topics in cognitive linguistics.* Amsterdam and Philadelphia: John Benjamins. 91–125.

Langacker, R. W. (1988b). A usage-based model. In Rudzka-Ostin, B. (Ed.), *Topics in cognitive linguistics.* Amsterdam and Philadelphia: John Benjamins. 127–161.

Langacker, R. W. (1988c). A view of linguistic semantics. In Rudzka-Ostyn, B. (Ed.), *Topics in cognitive linguistics.* Amsterdam and Philadelphia: John Benjamins. 49–90.

Langacker, R. W. (1990). Subjectification. *Cognitive Linguistics* 1(1): 5–38.

Langacker, R. W. (1991). *Foundations of cognitive grammar. Descriptive application.* Stanford: Stanford University Press.

Langacker, R. W. (1993). Reference-point constructions. *Cognitive Linguistics* 4(1): 1–38.

Langacker, R. W. (1999). *Grammar and conceptualization.* Berlin and New York: Mouton de Gruyter.

Langacker, R. W. (2002). *Concept, image and symbol. The cognitive basis of grammar.* Berlin and New York: Mouton de Gruyter.

Langacker, R. W. (2009). *Investigations in cognitive grammar.* Berlin and New York: Mouton de Gruyter.

Langacker, R. W. & Munro, P. (1975). Passives and their meaning. *Language* 75(1): 63–111.

Laszlo, S. & Federmeier, K. D. (2009). A beautiful day in the neighborhood: An event-related potential study of lexical relationships and prediction in context. *Journal of Memory and Language* 61(3): 326–338.

Lavidas, N. & Papangeli, D. (2007). Deponency in the diachrony of Greek. In Baerman, M., Corbett, G. G., Brown, D. & Hippisley, A. (Eds.), *Deponency and morphological mismatches*. Oxford: Oxford University Press. 97–126.

Lehmann, C., Koenig, T., Jelic, V., John, R. E., Lars-Olof, W., Dogde, Y. & Dierks, T. (2007). Application and comparison of classification algorithms for recognition of Alzheimer's disease in electrical brain activity (EEG). *Journal of Neuroscience Methods* 161(2): 343–350.

Leino, J. & Östman, J.-O. (2005). Constructions and variability. In Fried, M. & Boas, H. C. (Eds.), *Grammatical constructions: Back to the roots*. Amsterdam: John Benjamins. 191–213.

Leinonen, M. (1982). Roles and responsibilities: passivity in Russian and Finnish. *Scando-Slavica* 28(1): 201–208.

Leinonen, M. (1984). Quantifier sentences in Finnish and in Russian. *Studia Slavica Finlandensia* 1: 52–72.

Leinonen, M. (1985). *Impersonal sentences in Finnish and Russian: Syntactic and semantic properties*. Helsinki: Slavica Helsingiensia.

Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago and London: The University of Chicago Press.

Levin, B. & Rappaport Hovav, M. (2005). *Argument Realization*. Cambridge: Cambridge University Press.

Levinson, S. C. & Evans, N. (2010). Time for a sea-change in linguistics: Response to comments on 'The myth of language universals'. *Lingua* 120(12): 2733–2758.

Levinson, S. C. & Meira, S. (2003). 'Natural Concepts' in the spatial topological domain-adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language* 79(3): 485–516.

Levinson, S. C. & Wilkins, D. P. (2006). The background to the study of the language of space. In Levinson, S. C. & Wilkins, D. P. (Eds.), *Grammars of space*. Cambridge: Cambridge University Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition* 106(3): 1126–1177.

Liaw, A. & Wiener, M. (2002). Classification and regression by randomForest. *R News* 2(3): 18–22.

Lichtenberk, F. (1985). Multiple uses of reciprocal constructions. *Australian Journal of Linguistics* 5(1): 19–41.

Lidz, J. & Gleitman, L. R. (2004). Yes, we still need Universal Grammar. *Cognition* 94(1): 85–93.

Lidz, J. & Waxman, S. (2004). Reaffirming the poverty of the stimulus argument: A reply to the replies. *Cognition* 93(2): 157–165.

Lin, Y. & Jeon, Y. (2006). Random forests and adaptive nearest neighbors. *Journal of American Statistical Association* 101(747): 578–590.

Lunetta, K. L., Hayward, B. L., Segal, J. & Van Eederwegh, P. (2004). Screening large-scale association study data: Exploiting interactions using random forests. *BMC Genetics* 5(32): 1–13.

Lyons, J. (1977). *Semantics.* Cambrigde: Cambridge University Press.

Macwhinney, B. (2001). Emergentist approaches to language. In Bybee, J. L. & Hopper, P. J. (Eds.), *Frequency and the emergence of linguistic structure.* Amsterdam and Philadelphia: John Benjamins. 449–470.

Manney, L. J. (2000). *Middle voice in modern Greek: Meaning and function of an inflectional category.* Amsterdam and Philadelphia: John Benjamins.

Manning, C. & Schütze, H. (1999). *Foundations of statistical natural language processing.* Cambridge: MIT Press.

Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science (learning, development, and conceptual change).* Cambrigde: MA: MIT Press.

Markman, V. G. (2008). The case of predicates (revisited): Predicate instrumental in Russian and its restrictions. *Journal of Slavic Linguistics* 16(2): 187–246.

McDonald, S. & Shillcock, R. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech* 44(3): 295–323.

Mel'čuk, I. (1993). The inflectional category of voice: towards a more rigorous definition. In Comrie, B. & Polinsky, M. (Eds.), *Causatives and transitivity.* Amsterdam and Philadelphia: Benjamins. 1–46.

Mel'čuk, I. (1997). Grammatical case, basic verbal constructions, and voice in Maasai: Towards a better analysis of the concepts. In Dressler, W. U., Prinzhorn, M. & Rennison, J. R. (Eds.), *Advances in Morphology.* Berlin and New York: Mouton de Gruyter. 131–170.

Mel'čuk, I. (2004). Actants in semantics and syntax: I: Actants in semantics. *Linguistics* 42(1): 1–66.

Meyer, R. (2010). Reflexive passives and impersonals in North Slavonic languages: A diachronic view. *Russian Linguistics* 34(3): 285–306.

Michaelis, L. A. (2004). Type shifting in construction grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics* 15(1): 1–67.

Milin, P., Durdevic, F. & Moscoso del Prado Martín, F. (2009). The simultaneous effects of inflectional paradigms and classes on lexical recognition: Evidence from Serbian. *Journal of Memory and Language* 60(1): 50–64.

Miller, G. & Johnson-Laird, P. (1976). *Language and perception.* Cambridge: Cambridge University Press.

Miller, G. A., Beckwith, R., Fellbaum, C., Gross, D. & Miller, K. J. (1990). Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography* 3(4): 235–244.

Moisl, H. & Jones, V. (2005). Cluster analysis of the Newcastle Electronic Corpus of Tyneside English: A comparison of methods. *Literary and Linguistic Computing* 20(125-146): 125.

Moore, J. & Perlmutter, D. M. (2000). What does it take to be a dative subject? *Natural Language & Linguistic Theory* 18(2): 373–416.

Moscoso del Prado Martín, F., Kostić, A. & Baayen, R. H. (2004). Putting the bits together: An information theoretical perspective on morphological processing. *Cognition* 94(1): 1–18.

Mosteller, F. & Tukey, J. W. (1977). *Data analysis and regression: A second course in statistics*. Reading: Addison-Wesley.

Moyse-Faurie, C. (2008). Constructions expressing middle, reflexive and reciprocal situations in some Oceanic languages. In König, E. & Gast, V. (Eds.), *Reciprocals and reflexives: Theoretical and typological explorations*. Berlin and New York: Mouton de Gruyter. 105–168.

Murray, W. & Forster, K. I. (2004). Serial mechanisms in lexical access: The rank hypothesis. *Psychological Review* 111(3): 721–756.

Mustajoki, A. (2006). The Integrum database as a powerful tool in research on contemporary Russian. In Никипорец-Такигава, Г. (Ed.), *Integrum: точные методы и гуманитарные науки*. Москва: Летний сад.

Newmeyer, F. J. (2003). Grammar is grammar and usage is usage. *Language* 79(4): 682–707.

Nichols, J. (1981). *Predicate nominals. A partial surface syntax of Russian*. Berkeley: University of California Press.

Nicodemus, K. K., Malley, J. D., Strobl, C. & Ziegler, A. (2010). The behaviour of random forest permutation-based variable importance measures under predictor correlation. *BMC Bioinformatics* 11(110): 1–13.

Nielsen, R. D. & Pradhan, S. (2004). Mixing weak learners in semantic parsing. In Dekang, L. & Dekai, W. (Eds.), *Proceedings of EMNLP-2004*. 80–87.

Noë, A. (2004). *Action in perception*. Cambridge/Massachusetts: The MIT Press.

Nørgård-Sørensen, J. (2010). What languages must convey: The construction-based syntax of Old Russian. *Acta Linguistica Hafniensia* 42(1): 46–59.

North, B. V., Curtis, D. & Sham, P. C. (2002). A note on the calculation of empirical P values from Monte Carlo procedures. *American Journal of Human Genetics* 71(2): 439–441.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General* 115(1): 39–57.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14(1): 54–65.

Osherson, D. N., Wilkie, O., Smith, E. E., Lopez, A. & Shafir, E. (1990). Category-based induction. *Psychological Review* 97(2): 185–200.

Östman, J.-O. & Fried, M. (Eds.) (2005). *Construction Grammars: Cognitive grounding and theoretical extensions*. Amsterdam: John Benjamins.

Paducheva, E. (2003). Is there an 'ANTICAUSATIVE' component in the semantics of decausatives? *Journal of Slavic Linguistics* 11(1): 173–198.

Paducheva, E. (2008). Locative and existential meaning of Russian *быть*. *Russian Linguistics* 32(3): 147–158.

Paluničić, F., Ferreira, H. C., Swart, T. G. & Clarke, W. A. (2009). Modelling distances between genetically related languages using an extended weighted Levenshtein distance. *Southern African Linguistics and Applied Language Studies* 27(4): 381–389.

Parashar, A. & Vinay, C. (2008). Framework for clustering linguistic data, its algorithm for clustering and implementation using Perl. *COIT-2008*.

Pastizzo, M. J. & Feldman, L. B. (2009). Multiple dimensions of relatedness among words: Conjoint effects of form and meaning in word recognition. *Mental Lexicon* 4(1): 1–25.

Paul, H. (1989). *Prinzipien der Sprachgeschichte.* Halle: Max Niemeyer.

Penke, M. & Rosenbach, A. (2004). What counts as evidence in linguistics? An introduction. *Studies in Language* 28(3): 480–526.

Perlmutter, D. & Moore, J. (2002). Language-internal explanation: The distribution of Russian impersonals. *Language* 78(4): 619–649.

Perlmutter, D. M. (1983). Personal vs. impersonal constructions. *Natural Language & Linguistic Theory* 1(1): 141–200.

Pierrehumbert, J. B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In Bybee, J. L. & Hopper, P. J. (Eds.), *Frequency and the emergence of linguistic structure.* Amsterdam and Philadelphia: John Benjamins. 135–157.

Pierrehumbert, J. B. (2003). Probabilistic phonology: discrimination and robustness. In Bod, R., Hay, J. & Jannedy, S. (Eds.), *Probabilistic linguistics.* Cambridge/Mass.: MIT Press. 177–228.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure.* Cambridge: The MIT Press.

Pinker, S. (1999). *Words and rules: The ingredients of language.* New York: Basic Books.

Pinker, S. & Ullman, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Sciences* 6(11): 456–463.

Plag, I. & Baayen, R. H. (2010). Suffix ordering and morphological processing. *Language* 85(1): 109–152.

Podlesskaya, V. I. & Rakhilina, E. V. (1999). External possession, reflexivization and body parts in Russian. In Payne, D. L. & Barshi, I. (Eds.), *External Possession.* Amsterdam and Philadelphia: John Benjamins. 505–521.

Pollard, C. & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar.* Chicago: University of Chicago Press

Primus, B. (1999). *Cases and thematic roles.* Tübingen: Niemeyer.

Pustet, R. (2003). *Copulas: Universals in the categorization of the lexicon.* New York: Oxford University Press.

R Development Core Team (2011). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Ramscar, M. (2002). The role of meaning in inflection: Why the past tense doesn't require a rule. *Cognitive Psychology* 45(1): 45–49.

Rappaport Hovav, M. & Levin, B. (1998). Building verb meanings. In Butt, M. & Geuder, W. (Eds.), *The projection of arguments.* Stanford: CSLI Publications. 97–134.

Rhymes (2011). Большой словарь рифм. Версия 3.0.6.

Rohrer, T. (2007). Embodiment and experientialism. In Geeraerts, D. & Cuyckens, H. (Eds.), *The Oxford handbook of cognitive linguistics.* Oxford and New York: The Oxford University Press. 25–47.

Rosch, E. (1978). Principles of categorization. In Rosch, E. & Lloyd, B. B. (Eds.), *Cognition and categorization.* Hillsdale, New Jersey: Lawrence Erlbaum associates. 27–47.

Rosch, E. & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7(4): 573–605.

Sag, I. A. (2010). Sign-based Construction Grammar: An informal synopsis. In Boas, H. C. & Sag, I. A. (Eds.), *Sign-Based Construction Grammar.* 39–170.

Sahlgren, M. & Knutsson, O. (2010). Workshop on extracting and using constructions in computational linguistics. *Proceedings of the NAACL HLT workshop on extracting and using constructions in computational linguistics.*

Sapir, E. (1955 [1921]). *Language: An introduction to the study of speech.* San Diego, New York and London: Harcourt Brace & Company.

Schapire, R. E. (1990). The strength of weak learnability. *Machine Learning* 5: 197–227.

Scheibman, J. (2002). *Point of view and grammar. Structural patterns of subjectivity in American English conversation.* Amsterdam and Philadelphia: John Benjamins.

Schenker, A. M. (1986). On the reflexive verbs in Russian. *International Journal of Slavic Linguistics and Poetics* XXXIII: 27–41.

Schulte im Walde, S. (2006). Experiments on the automatic induction of German semantic verb classes. *Computational Linguistics* 32(2): 159–194.

Segal, M. R. (2004). Machine learning benchmarks and random forest regression. *Center for Bioinformatics and Molecular Biostatistics.* San Francisco: University of California. 1–14.

Seo, S. (2001). *The frequency of null subject in Russian, Polish, Czech, Bulgarian and Serbo-Croatian: An analysis according to morphosyntactic environments*. Doctoral Dissertation, Indiana University at Bloomington.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* 27(4): 379–423.

Shi, T. & Horvart, S. (2006). Unsupervised learning with random forest predictors. *Journal of Computational and Graphical Statistics* 15(1): 118–138.

Shibatani, M. (1996). Applicatives and benefactives: A cognitive account. In Masayoshi, S. & Thompson, S. A. (Eds.), *Grammatical constructions: Their form and meaning.* Oxford: Clarendon Press. 157–194.

Shibatani, M. (1998). Voice in Philippine languages. In Shibatani, M. (Ed.), *Passive and voice.* Amsterdam: Jonh benjamins. 85–142.

Siewierska, A. (1984). *The passive: A comparative linguistic analysis.* London: Croom Helm.

Siewierska, A. (2004). *Person.* Cambridge: Cambridge University Press.

Sigurðsson, H. Á. (2002). To be an oblique subject: Russian vs. Icelandic. *Natural Language & Linguistic Theory* 20(4): 691–724.

Silverstein, M. (1976). Hierarchy of features and ergativity. In Dixon, R. M. W. (Ed.), *Grammatical categories in Australian languages.* Canberra: Australian Institute of Aboriginal Studies. 112–171.

Sinclair, J. M. & Mauranen, A. (2006). *Linear unit grammar: Integrating speech and writing.* Amsterdam: John Benjamins.

Slobin, D. I. (2006). What makes manner of motion salient? Explorations in linguistic typology, discourse and cognition. In S., H. M. a. R. (Ed.), *Space in languages: Linguistic systems and cognitive categories.* Amsterdam and Philadelphia: John Benjamins. 59–81.

Smolka, E., Komlósi, S. & Rösler, F. (2009). When semantics means less than morphology: The processing of German prefixed verbs. *Language and Cognitive Processes* 24(3): 337–375.

Sokolova, M. & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management* 45: 427–437.

Solstad, T. & Lyngfelt, B. (2006). Perspectives on demotion. In Solstad, T. & Lyngfelt, B. (Eds.), *Demoting the agent: passive, middle and other voice phenomena* Amsterdam and Philadelphia: John Benjamins. 1–20.

Sonnenhauser, B. (2010). The event structure of verbs of emotion in Russian. *Russian Linguistics* 34(3): 331–353.

Spivey, M. (2007). *The continuity of the mind.* Oxford and New York: Oxford University Press.

Stassen, L. (1997). *Intransitive Predication.* Oxford: Clarendon.

Stefanowitsch, A. (2008). Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics* 19(3): 513–531.

Stefanowitsch, A. (2010). Empirical cognitive semantics: some thoughts. In Glynn, D. & Fischer, K. (Eds.), *Quantitative methods in cognitive semantics: Corpus-driven approaches.* Berlin: Mouton de Gruyter. 357–380.

Stefanowitsch, A. & Gries, S. T. (2003). Collostructions: investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2): 209–243.

Steyvers, M. & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science* 29(1): 41–78.

Stokhof, M. & van Lambalgen, M. (2011). Comments-to-comments. *Theoretical Linguistics* 37(1): 79–94.

Strehl, A. & Ghosh, J. (2000). Impact of similarity measures on web-page clustering. *AAAI Technical Report WS-00-01.* 58–64.

Strobl, C. (2008). *Statistical issues in machine learning: Towards reliable split selection and variable importance measures.* Göttingen: Cuvillier Verlag.

Strobl, C., Boulesteix, A.-L., Kneib, T., Augustin, T. & Zeileis, A. (2008). Conditional variable importance for random forests. *BMC Bioinformatics* 9(307).

Strobl, C., Boulesteix, A.-L., Zeileis, A. & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics* 8(25).

Strobl, C., Hothorn, T. & Zeileis, A. (2009a). Party on! A new, conditional variable importance measure for random forests available in the party package. *The R Journal* 1(2): 14–17.

Strobl, C., Malley, J. & Tutz, G. (2009b). An introduction to recursive partitioning: Rationale, application and characteristics of classification and regression trees, bagging and random forests. *Psychological Methods* 14(4): 323–348.

Strobl, C. & Zeiles, A. (2008). Danger: High Power! - Exploring the statistical properties of a test for random forest variable importance. *COMPSTAT 2008 - Proceedings in Computational Statistics.* Heidelberg: Physica Verlag. 59–66.

Strokel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *Journal of Child Language* 29: 251–274.

Sturtevant, E. H. (1931). The origin of the medio-passive. *Language* 7(4): 242–251.

Suárez, M. (2004). An inferential conception of scientific representation. *Philosophy of Science* 71(5): 767–779.

Suárez, M. & Cartwright, N. (2008). Theories: Tools versus models. *Studies in History and Philosophy of Modern Physics* 39(62-81): 62.

Surdeanu, M., Harabagiu, S., Williams, J. & Aarseth, P. (2003). Using predicate-argument structures for information extraction. In Proceedings of ACL 2003. 8–15.

Suttle, L. & Goldberg, A. E. (2011). The partial productivity of constructions as induction. *Linguistics* 49(6): 1237–1269.

Talmy, L. (1975). Semantics and syntax of motion. In Kimball, J. P. (Ed.), *Syntax and semantics.* New York: Academic Press. 181–238.

Talmy, L. (2000a). *Toward a cognitive semantics: Concept structuring systems.* Cambridge: MA: MIT Press.

Talmy, L. (2000b). *Toward a cognitive semantics: Typology and process in concept structuring.* Cambridge: MA: MIT Press.

Talmy, L. (2006). Foreword. In Gonzalez-Marquez, M., Mittelberg, I., Coulson, S. & Spivey, M. J. (Eds.), *Methods in cognitive linguistics.* Amsterdam and Philadelphia: John Benjamins. xi–xxi.

Tesnière, L. (1959). *Eléments de syntaxe structurale.* Paris: klincksieck.

Timberlake, A. (1986). The semantics of the case in Russian predicate complements. *Russian Linguistics* 10(2): 137–165.

Timberlake, A. (2004). *A reference grammar of Russian.* Cambridge.

Tomasello, M. (2003). *Constructing a Language: A usage-based theory of language acquisition.* Cambridge: Harvard University Press.

Tomasello, M. (2004). Syntax or semantics. Response to Lidz et al. *Cognition* 93(2): 75–165.

Tuggy, D. (1993). Ambiguity, polysemy and vagueness. *Cognitive Linguistics* 4(3): 273–290.

Tversky, A. & Gati, I. (1978). Studies of similarity. In Rosch, E. & Lloyd, B. B. (Eds.), *Cognition and categorization.* Hillsdale, New Jersey: Lawrence Erlbaum associates. 79–98.

Van Valin, R. D., Jr. & LaPolla, R. J. (1997). *Syntax: structure, meaning and function.* Cambridge: Cambridge University Press.

Vendler, Z. (1957). Verbs and times. *The Philosophical Review* 66: 143–160.

Verhagen, A. (2005). *Constructions of intersubjectivity: Discourse, syntax, and cognition.* New York: Oxford University Press.

Verhagen, A. (2008). Intersubjectivity and explanation in linguistics: A reply to Hinzen and van Lambalgen. *Cognitive Linguistics* 19(1): 125–143.

Vitevitch, M. (2008). What can graph theory tell us about word learning and lexical retrieval. *Journal of Speech, Language, and Hearing Research* 51: 408–422.

Westerhout, E. (2009). Definition extraction using linguistic and structural features. *RANLP 2009.*

Whitten, W. B., Newton Suter, W. & Frank, M. L. (1979). Bidirectional synonym ratings of 464 noun pairs. *Journal of Verbal Learning and Verbal Behavior* 18(1): 109–127.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis.* London/New York: Springer.

Wiemer, B. (2004). The evolution of passives as grammatical constructions in Northern Slavic and Baltic languages. In Bisang, W., Himmelmann, N. P. & Wiemer, B. (Eds.), *What makes Grammaticalization? A look from its fringes and its components.* Berlin and New York: Mouton de Gruyter. 271–331.

Wierzbicka, A. (1986). What's in a noun? (Or: how do nouns differ in meaning from adjectives?). *Studies in Language* 10: 353–389.

Williams, A. (1993). The argument structure of *sja*-predicates. *Journal of Slavic Linguistics* 1(1): 167–190.

Williams, A. (1999). Prototype marker or reflexive marker: Russian *-sja* and categorical change. In De Stadler, L. & Eyrich, C. (Eds.), *1993 Proceedings of the International Cognitive Linguistics Conference. Issues in Cognitive Linguistics.* Berlin: Mouton de Gruyter. 277–295.

Wu, B., Abbot, T., Fishman, D., McMurray, W., Mor, G., Stone, K., Ward, D., Williams, K. & Hongyu, Z. (2003). Comparison of statistical methods for classification of ovarian cancer using mass spectrometry data. *Bioinformatics* 19(12): 1636–1643.

Xu, P. & Jelinek, F. (2007). Random forests and the data sparseness problem in language modeling. *Computer Speech & Language* 21(1): 105–152.

Xu, Z., Aranoff, M. & Anshen, F. (2007). Deponency in Latin. In Baerman, M., Corbett, G. G., Brown, D. & Hippisley, A. (Eds.), *Deponency and morphological mismatches.* Oxford: Oxford University Press. 127–143.

Yarkoni, T. & Balota, D. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review* 15(5): 971–979.

Zdorenko, T. (2010). Subject omission in Russian: A study of the Russian National Corpus. In Gries, S. T., Wulff, S. & Davies, M. (Eds.), *Corpus-linguistic applications: Current studies, new directions.* Amsterdam and New york: Rodopi. 119–133.

Ziegler, J. C. & Conrad, P. (1998). No more problems in Coltheart's neighborhood: resolving neighborhood conflicts in the lexical decision task. *Cognition* 68(2): B53–B62.

Zimmerling, A. (2009). Dative subjects and semi-expletive pronouns in Russian. In Zybatow, G., Junghanns, U., Lenertová, D. & Biskup, P. (Eds.), *Studies in formal Slavic phonology, morphology, syntax, semantics and information structure. Proceedings of FDS 7, Leipzig 2007.* Frankfurt am Main: Peter Lang. 253–265.

Zipf, G. K. (1965 [1935]). *The psycho-biology of language.* Cambridge: M. I. T.

Zlatev, J. (2007). Spatial semantics. In Geeraerts, D. & Cuyckens, H. (Eds.), *The Oxford handbook of cognitive linguistics.* Oxford and New York: Oxford University Press. 318–350.

Азарова, И. В., Синопальникова, А. А. & Яворская, М. В. (2004). Принципы построения wordnet-тезауруса RussNet. *Диалог-2004.* Наука.

Апресян, Ю. Д. (1974). *Лексическая семантика: Синонимические средства языка.* Москва: Наука.

Апресян, Ю. Д. (1980). *Типы информации для поверхностно-семантического компонента модели «Смысл ↔ Текст».* Wien: Wiener Slawistischer Almanach. Sonderband 1.

Апресян, Ю. Д. (1986). Дейксис в лексике грамматике и наивная модель мира. *Семиотика и информатика* 28: 5–53.

Апресян, Ю. Д. (1995a). *Избранные труды. Том 1 : Интегральное описание языка и системная лексикография.* Москва: Школа Языки русской культуры.

Апресян, Ю. Д. (1995b). *Избранные труды. Том 2 : Интегральное описание языка и системная лексикография.* Москва: Школа Языки русской культуры.

Апресян, Ю. Д. (2002). Взаимодействие лексики и грамматики: лексикографический аспект. *Русский язык в научном освещении* 3(1): 10–29.

Апресян, Ю. Д. (2005). О московской семантической школе. *Вопросы языкознания* 1: 3–30.

Арутюнова, Н. Д. (1976). *Предложение и его смысл. Логико-семантические проблемы.* Москва: Наука.

Арутюнова, Н. Д. (1999). Язык и мир человека. Москва: Языки русской культуры.

Богуславский, И. М. (1998). Сфера действия и начинательности и актуальное членение: втягивание ремы. *Семиотика и информатика* 36: 8–18.

Бондарко, А. В. (1990). Темпоральность. In Бондарко, А. В. (Ed.), *Теория функциональной грамматики. Темпоральность. Модальность.* Ленинград: Наука. 5–58.

Бондарко, А. В. (1991). Семантика лица. In Бондарко, А. В. (Ed.), *Теория функциональной грамматики : Персональность, Залоговость.* Санкт-Петербург: Наука. 5–40.

Бондарко, А. В. (2002). *Теория значения в системе функциональной грамматики на материале русского языка.* Москва: Языки славянской культуры.

Буланин, Л. Л. (1986). *Категория залога в современном русском языке.* Ленинград: Ленинградский университет.

Булаховский, Л. А. (1954). *Русский литературный язык первой половины XIX века.* Москва.

Булыгина, Т. В. & Шмелев, А. Д. (1997). *Языковая концептуализация мира (на материале русского языка).* Москва: Школа Языки русской культуры.

Введенская, Л. А., Павлова, Л. Г. & Кашаева, Е. Ю. (2005). *Русский язык и культура речи.* Ростов н/Д: Феникс.

Вимер, Б. (2001). Аспектуальные парадигмы и лексическое значение русских и литовских глаголов (Опыт сопоставления с точки зрения лексикализации и грамматикализации). *Вопросы языкознания* 2: 26–58.

Виноградов, В. В. (1972). *Русский язык (грамматическое учение о слове).* Москва: Высшая школа.

Виноградов, В. В. (1975). *Избранные труды. Исследования по русской грамматике.* Москва: Наука.

Виноградов, В. В. (Ed.) (1994). *История слов. Часть 1.* Москва: Толк.

Гаврилова, В. И. (1999). Сознательные действия, стихийные процессы и ситуация создания и снятия прегады. In Арутюнова, Н. Д. & Рябцева, Н. К. (Eds.), *Логический анализ языка: языки динамического мира*. Москва: Дубна. 159–174.

Галкина-Федорук, Е. М. (1958). *Безличные предложения в современном русском языке*. Москва: Изд. Московского университета.

Гладкий, А. В. (2007). О точных и математических методах в лингвистике и других гуманитарных науках Текст. *Вопросы языкознания* 5: 22–38.

Даль, В. И. (1903-1909). Толковый словарь живого великорусского языка. In Бодуэн де Куртенэ, И. А. (Ed.).

Данков, В. Н. (1981). *Историческая грамматика русского языка. Выражение залоговых отношений у глагола*. Москва: Высшая школа.

Денисов, П. Н. & Морковкин, В. В. (2002). Словарь сочетаемости слов русского языка. Москва: Астрель.

Долинина, И. Б. (1991). Ревлексивность и каузативность (категориальная семантика ревлексивных конструкций, соотносительных с каузативными конструкциями). In Бондарко, А. В. (Ed.), *Персональность. Залоговость*. Ленинград: Наука. 327–345.

Евгеньева, А. П. (1999). Словарь русского языка: В 4-х т. Москва: Русский язык.

Ефремова, Т. Ф. (2000). Новый словарь русского языка. Толково-словообразовательный. Москва: Русский язык.

Жолковский, А. К. & Мельчук, И. А. (1965). О возможном методе и инструментах семантического синтеза. *Науч.-техн.информ.* 3: 23–28.

Зализняк, А. А. (1980). *Грамматический словарь русского языка*. Москва: Русский язык.

Зализняк, А. А., Микаэлян, И. Л. & Шмелев, А. Д. (2010). Видовая коррелятивность в русском языке: в защиту видовой пары. *Вопросы языкознания* 1: 3–23.

Зарицкий, Н. С. (1961). *Формы и функции возвратных глаголов (на материале древнерусского языка)*. Киев: Издателъство Киевского университета.

Засорина, Л. Н. (1977). Частотный словарь русского языка. Москва: Русский язык.

Зельдович, Г. М. (2010). Синтетический пассив на *-ся*: Почему его (почти) нет. *Вопросы языкознания* 2: 3–36.

Золотова, Г. А. (1981). О субъекте предложения в современном русском языке. *Филологические науки* 1: 33–42.

Золотова, Г. А. (2000a). Понятие личности/безличности и его интепретации. *Russian Linguistics* 24(2): 103–115.

Золотова, Г. А. (2000b). Понятие личности/безличности и его интепретации. *Russian Linguistics* 24(2): 103-115.

Золотова, Г. А. (2005 [1973]). *Очерк функционального синтаксиса русского языка*. Москва: КомКнига.

Золотова, Г. А., Онипенко, Н. К. & Сидорова, М. Ю. (1998). *Коммуникативная грамматика русского языка*. Москва: РАН.

Исаченко, А. В. (1960). *Русский язык в сопоставлении со словацким*. Братислава.

Калашникова, К. В. & Сай, С. С. (2006). Системные отношения между классами русских рефлексивных глаголов в связи с их частотными характеристиками (по данным корпусного исследования). In Храковский, В. С., Дмитренко, С. Ю. & Заика, Н. М. (Eds.), *Проблемы типологии и общей лингвистики. Международная конференция, посвященная 100 летию со дня рождения проф. А. А. Холодовича.* Санкт-Петербург: Нестор. 56–54.

Князев, Ю. П. (1999). Обозначение направленного движения в русском языке. In Арутюнова, Н. Д. & Рябцева, Н. К. (Eds.), *Логический анализ языка: языки динамического мира.* Москва: Дубна. 182-192.

Князев, Ю. П. (2007). *Грамматическая семантика. Русский язык в типологической перспективе.* Москва: Языки сланянских культур.

Коломацкий, Д. И. (2009). *Дистрибуция русских пассивных форм: корпусное исследование.* Московский государственный университет им. М. В. Ломоносова.

Копотев, М. В. (2008). К построению частотной грамматики русского языка: падежная система по корпусным данным. In Мустайоки, А., Копотев, М. В., Бирюлин, Л. А. & Протасова, Е. Ю. (Eds.), *Инструментарий русистики: корпусные подходы.* Helsinki: Department of Slavonic and Baltic Languages and Literatures. 136–151.

Копотев, М. В. & Мустайоки, А. (2008). Современная корпусная русистика. In Мустайоки, А., Копотев, М. В., Бирюлин, Л. А. & Протасова, Е. Ю. (Eds.), *Инструментарий русистики: корпусные подходы.* Helsinki: Department of Slavonic and Baltic Languages and Literatures. 7–24.

Королев, Э. И. (1968). Количественные характеристики смысловых классов возвратных глаголов (неопубл).

Кретов, А. А. (2009). Анализ семантических помет в НКРЯ. *Национальный корпус русского языка: 2006—2008. Новые результаты и перспективы.* СПб.: Нестор-История. 240–257.

Кронгауз, М. А. (1993). Семантика русского глагола и его словообразовательные возможности. *Russian Linguistics* 17(1): 15–36.

Кронгауз, М. А. (1997). Исследования в области глагольной префиксации: современное положение дел и перспективы. In Кронгауз, М. & Пайар, Д. (Eds.), *Глагольная префиксация в русском языке: сборник статей.* Москва: Русские словари. 4–28.

Крысин, Л. П. (2007a). Русская литературная норма и современная речевая практика. *Русский язык в научном освещении* 14(2): 5–17.

Крысин, Л. П. (2007b). Современный русский язык : лексическая семантика, лексикология, фразеология, лексикография. Москва: Академия.

Крысько, В. Б. (1984). Транзитивность возвратных глаголов в русском языке XI–XVIII вв. *Вестник Ленингр. ун-та.* 2: 79–84.

Крысько, В. Б. (2006). *Исторический синтаксис русского языка: Объект и переходность.* Москва: Азбуковник.

Кузнецов, С. А. (2009 [1998]). Большой толковый словарь русского языка. Санкт-Петербург: Норинт.

Кузнецова, М. В. (1984). *К истории формирования возвратных глаголов (на материале языка памятников древнеславянской письменности).*

Кузнецова, Э. В. (1989). *Лексико-семантические группы русских глаголов.* Иркутск: Иркутск. ун-та.

Кустова, Г. И. (2002). Семантические аспекты лексических функций (глаголы со значением 'начаться'?/ 'кончиться'?). In Арутюнова, Н. Д. (Ed.), *Логический анализ языка. Семантика начала и конца.* Москва: Индрик. 69–82.

Кустова, Г. И., Ляшевская, О. Н., Падучева, Е. В. & Рахилина, Е. В. (2005). Семантическая разметка лексики в Национальном корпусе русского языка: принципы, проблемы, перспективы. *Национальный корпус русского языка: 2003-2005. Результаты и перспективы.* М. 155–174.

Лагута, О. Н. & Тимофеева, М. К. (2007). Национальный корпус русского языка и Интегрум: итоги и перспективы. *Русский язык в научном освещении* 14(2): 113–132.

Левенштейн, В. И. (1966). Двоичные коды с исправлением выпадений, вставок и замещений символов. *Доклады Академий Наук СССР* 163(10): 845–848.

Ляшевская, О. Н. & Шаров, С. А. (2009). Частотный словарь современного русского языка (на материалах Национального корпуса русского языка). Электронная версия издания. Москва: Азбуковник.

Майсак, Т. А. & Рахилина, Е. В. (1999). Семантика и статистика: глагол *идти* на фоне других глаголов движения. In Арутюнова, Н. Д. & Рябцева, Н. К. (Eds.), *Логический анализ языка: языки динамического мира.* Москва: Дубна. 53–66.

Мейе, А. (2001 [1951]). *Общеславянский язык.* Москва: Прогресс.

Мельчук, И. А. (1995). *Русский язык в модели Смысл-Текст.* Москва/Вена: Школа Языки русской культуры.

Мельчук, И. А. & Холодович, А. А. (1970). К теории грамматического залога. *Народы Азии и Африка* 4: 111–124.

Милославский, И. Г. (1978). Какому залогу принадлежит глагол «нравиться»? In Холодович, А. А. (Ed.), *Проблемы теории грамматического залога.* Ленинград: Наука. 208–213.

Мучник, И. П. (1971). *Грамматические категории глагола и имени в современном русском литературном языке.* Москва: Наука.

Недялков, В. П. (1971). *Каузативные конструкции в немецком языке. Аналитический каузатив.* Наука: Ленинград.

Недялков, В. П. (1978). Заметки по типологии рефлексивных деагентативных конструкций (опыт исчисления). *Проблемы теории грамматического залога.* Ленинград: Наука. 28–42.

Недялков, В. П. & Сильницкий, Г. Г. (1969). Типология каузативных конструкций. In Холодович, А. А. (Ed.), *Типология каузативных конструкций. Морфологический каузатив.* Ленинград: Наука. 5–19.

Никипорец-Такигава, Г. Ю. (Ed.) (2006). *Integrum: точные методы и гуманитарные науки.* Москва: Летний сад.

Норман, Б. Ю. (2004). Возвратные глаголы-неологизмы в русском языке и синтактические-предпосылки их образования. In Храковский, В. С., Мальчуков, А. Л. & Дмитренко, С. Ю. (Eds.), *40 лет Санкт-Петербургской типологической школе.* Москва: Знак. 394–406.

Падучева, Е. В. (1999). Глаголы движения и их стативные дериваты (в связи с так называемым движением времени). In Арутюнова, Н. Д. & Рябцева, Н. К. (Eds.), *Логический анализ языка: языки динамического мир.* Москва: Дубна. 87–107.

Падучева, Е. В. (2001). Каузативний глагол и декаузатив в русском языке. *Русский язык в научном освещении* 1(1): 52–79.

Падучева, Е. В. (2002). Диатеза и диатетический сдвиг. *Russian Linguistics* 26(2): 179–215.

Падучева, Е. В. (2004). *Динамические модели в семантике лексики.* Москва: Языки славянской культуры.

Падучева, Е. В. (2006). Генитив дополнения в отрицательном предложении. *Вопросы языкознания* 6: 21–43.

Паневова, Я. (1978). Критерии для установления облигаторных партиципантов глагола. In Храковский, В. С. (Ed.), *Проблемы теории грамматического залога.* Ленинград: Наука. 50–79.

Перцов, Н. В. (2003). Возвратные страдательные формы русского глагола в связи с проблемой существования в морфологии. *Вопросы языкознания* 4: 43–71.

Пешковский, А. М. (1938). *Русский синтаксис в научном освещении* Москва.

Плунгян, В. А. (2000). *Общая Морфология.* Москва: Эдиториал УРСС.

Ровнова, О. Г. (1998). Имперфективация глагола в русских диалектах (с точки зрения синхронии и диахронии). In Черткова, М. Ю. (Ed.), *Типология вида: проблемы, поиски, решения.* Москва. 396–404.

Розина, Р. И. (1999). Движение в физическом и ментальном пространстве. In Арутюнова, Н. Д. & Рябцева, Н. К. (Eds.), *Логический анализ языка: языки динамического мира.* Москва: Дубна. 108–118.

Сай, С. С. (2007). Прагматически обусловленные возвратные конструкции ”опущенного объекта” в русском языке. *Вопросы языкознания* 2: 75–91.

Сиротинина, О. Б. (2006 [1965]). *Порядок слов в русском языке* Москва: КомКнига.

Соболев, А. Н. (2005). Заметка о так называемых глаголав imperfectiva tantum в русском языке. *Russian Linguistics* 29(2): 189–199.

Соболевский, А. И. (2004 [1907]). *Труды по истории русского языка: Очерки из истории русского языка; Лекции по истории русского языка (репринтное издание).* Москва: Языки славянской культуры.

Тихонов, А. Н. (1985a). Словообразовательный словарь русского языка. Москва: Русский язык.

Тихонов, А. Н. (1985b). Словообразовательный словарь русского языка. Москва: Русский язык.

Тихонов, А. Н. (1998). *Русский глагол: проблемы теории и лексикографии.* Москва: Academia.

Фасмер, М. (1986). Этимологический словарь русского языка. В четырех томах. Москва: Прогресс.

Фортунатов, Ф. Ф. (1899). О залогах русского глагола. *Известия Отд. русского языка и словесности АН.* 1153–1158.

Храковский, В. С. (1974). Пассивные конструкции. In Холодович, А. А.

(Ed.), *Типология пассивных конструкций: диатезы и залоги.* Ленинград: Наука. 5–45.

Храковский, В. С. (1978a). Залог и рефлексив. In Храковский, В. С. (Ed.), *Проблемы теории грамматического залога.* Ленинград: Наука. 50–61.

Храковский, В. С. (Ed.) (1978b). *Проблемы теории грамматического залога.* Ленинград: Наука.

Храковский, В. С. (1981). Диатеза и референтность. In Храковский, В. С. (Ed.), *Залоговые конструкции в разноструктурных языках.* Ленинград: Наука. 5–38.

Храковский, В. С. (1987). Фазовость. Семантика фазовости и средства ее выражения. In Бондарко, А. В. (Ed.), *Теория функциональной грамматики. Введение. Аспектуальность. Временная локализованность. Таксис.* Ленинград: Наука. 153–209.

Храковский, В. С. (1991). Пассивные конструкции. In Бондарко, А. В. (Ed.), *Теория функциональной грамматики – Персональность. Залоговость.* Санкт-Петербург: Наука. 141–181.

Храковский, В. С. (2004). Концепция диатез залогов. In Храковский, В. С., Мальчуков, А. Л. & Дмитренко, С. Ю. (Eds.), *40 лет Санкт-Петербургской типологической школе.* Москва: Знак. 505–519.

Храковский, В. С. (2005). Аспектуальные тройки и видовые пары. *Русский язык в научном освещении* 9(1): 46–59.

Цейтлин, С. Н. (1978). Возвратные глаголы и детская речь. In Храковский, В. С. (Ed.), *Проблемы теории грамматического залога.* Ленинград: Наука. 193–197.

Шапир, М. И. (2005). "Тебе числа и меры нет": О возможностях и границах "точных методов" в гуманитарных науках. *Вопросы языкознания* 1: 43–62.

Шахматов, А. А. (1925). *Синтаксис русского языка. Выпуск I: учение о предложении.* Ленинград: АН СССР.

Шведова, Н. Ю. (2001). Еще раз о глаголе *бить. Вопросы языкознания* 2: 3–12.

Шведова, Н. Ю. & другие (Eds.) (1982). *Русская грамматика. Том 1.* Москва: Наука.

Шелякин, М. А. (1983). *Категория вида и способы действия русского глагола: теоретические основы.* Таллин: Валгус.

Ширшов, И. А. (2004). Толковый словообразовательный словарь русского языка. Москва: Русские словари.

Якобсон, Р. О. (1985a). *Избранные работы.* Москва: ПРОГРЕСС.

Якобсон, Р. О. (1985b). К общему учению о падеже. In Звегинцева, В. А. (Ed.), *Избранные работы. Переводы с английского, немецкого, французского.* Москва: ПРОГРЕСС. 133–175.

Янко-Триницкая, Н. А. (1962). *Возвратные глаголы в современном русском языке.* Москва: Академия наук.

Яхонтов, С. Е. (1981). Выражение рефлексивности в китайском языке. In Храковский, В. С. (Ed.), *Залоговые конструкции в разноструктурных языках.* Ленинград: Наука. 146–159.