THE POLYSEMY OF AN "EMPTY" PREFIX: A CORPUS-BASED COGNITIVE SEMANTIC ANALYSIS OF THE RUSSIAN VERBAL PREFIX *PO-*

Nicholas Lance LeBlanc

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Slavic Languages and Literatures (Slavic Linguistics).

Chapel Hill
2010

Approved by
Laura A. Janda
Lawrence Feinberg
Radislav Lapushin
Christopher Putney
J. Michael Terry

ABSTRACT

NICHOLAS LANCE LEBLANC: The polysemy of an "empty prefix": A corpus-based cognitive semantic analysis of the Russian verbal prefix *po-*
(Under the direction of Laura Janda and Lawrence Feinberg)

This dissertation proposes a structured semantic account of the polysemous Russian verbal prefix *po-* within the theoretical framework of cognitive linguistics and using corpus linguistic methods. While scholarly consensus identifies five meanings for *po-* and an additional meaning in conjunction with the suffix *-yva-*, the relationships among these six meanings have not been fully explored. By means of a corpus-linguistic analysis I determine the semantic structure linking the various meanings of *po-*: I collect a randomly selected sample of *po-*prefixed verbs (with accompanying contexts) from the Russian National Corpus, the largest annotated Russian language corpus extant. The collected data is manually tagged for a number of collocational, syntactic, and semantic parameters to create a behavioral profile of *po-*. The behavioral profile is subjected to a hierarchical agglomerative cluster analysis, resulting in a dendrogram that illustrates varying degrees of connection among meanings. Meanings are grouped into clusters based on degree of similarity, and the intra- and inter-cluster differences are investigated by use of *z*-scores and *t*-values. I then apply cognitive linguistic concepts to motivate the semantic structure of *po-*, showing how this account both echoes and expands previous work on prefixal semantics.

I conclude that the meanings of *po-* can be grouped into two clusters: Cluster one is comprised of the attenuative, delimitative, ingressive, and resultative meanings. Cluster two contains the more peripheral distributive and intermittent-attenuative meanings. The resultative meaning is prototypical and indicates that the subject has traversed the metaphoric path implied by the base verb in its entirety. The remaining meanings are metaphoric and metonymic extensions of that central meaning. This view of the semantics of *po-* coincides with what is known about the historical development of the prefix.

The contributions of this dissertation are twofold: First, I have produced a cognitively-motivated description of the semantic structure of *po-* based on empirical data. Secondly, this analysis suggests that quantitative methods are useful not only in the study of lexemes and grammatical constructions, but also in prefixal semantics. In addition, I point out large groups of *po-*prefixed verbs largely untouched by the scholarly literature that deserve further study.

ACKNOWLEDGEMENTS

Without the help and support of my advisors, professors, colleagues, family, and friends, this dissertation would simply not have been possible. I am deeply indebted to the present and former faculty members of the Department of Slavic Languages and Literatures at UNC - Chapel Hill who, in one way or another, have assisted me on this journey: Lawrence Feinberg, Beth Holmgren, Laura Janda, Radislav Lapushin, Madeline Levine, Eleonora Magomedova, Christopher Putney, and Ivana Vuletic. I owe a particular thank you to Laura Janda, who from the beginning has offered me her unwavering support. I can only hope that this dissertation does justice to the time, energy, and hours of guidance she has so generously provided. My thanks also to Lawrence Feinberg for his help parsing many a novel po-prefixed verb, and to all the members of my committee for their time and insight: Lawrence Feinberg, Laura Janda, Radislav Lapushin, Christopher Putney, and Michael Terry. Furthermore, Radislav Lapushin and Eleonora Magomedova provided invaluable assistance translating difficult Russian-language data.

I am grateful for the help of a number of colleagues: Dagmar Divjak provided great assistance navigating the waters of statistically-analyzable corpora. Adrian Ilie's programming abilities were essential to the development of the software behind my data collection. Chuck Simmons helped translate my database dreams into reality, and Chris Wiesen of UNC's Odum Institute and Chris Cabanski both kindly assisted with the

computation and interpretation of the actual statistics. John Korba aided in the remote acquisition and processing of dictionary material. Other fellow graduate students, professionals, and friends also provided indispensable help and support: Jenny and Brian Barrier, William Meyer, Kevin Reese, Evan Waldheter, and John Wrobel.

My two roommates during this process deserve special mention: Maya Bringe kept me sane during the early days of the dissertation and helped me to stay the course during a period of doubts. Diane Caton likewise encouraged me to continue progressing steadily despite a number of setbacks. I owe a special thanks to my mother, Tina A. LeBlanc, who so patiently prodded me onward. And last, but far from least, I would like to thank my partner Chris Martinez for his care and devotion throughout this challenging endeavor, and for the countless hours of dissertation-related discussions he endured. He has facilitated the successful completion of this work in more ways than I can mention.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF FIGURES (Cont.)

Figure

1 Introduction

This dissertation is a cognitive linguistic investigation into the semantic structure of the Russian verbal prefix *po-* . In my analysis I make extensive use of corpus data and proven statistical techniques to uncover the relationships among the widely disparate meanings of *po-*. As such this dissertation takes advantage of recent developments in the fields of cognitive and corpus linguistics to make an empirically-substantiated contribution to our understanding of the semantics of Russian verbal prefixes. In this introduction I outline the flow of the dissertation, progressing from general historical and theoretical background, to methodological details of the study, to specific findings of the analysis, and finally to a summary highlighting the contributions of this dissertation and directions for future research.

The semantics of the Russian verbal prefix *po-* has been problematic for linguists ever since the prefix became an object of scholarly inquiry. In particular, the prefix *po-* presents two interesting problems: the delineation problem and the structuring problem. At the heart of the delineation problem lies the question "How many meanings does *po-* have?" As the reader will see in 2.1 – 2.2.1, this question was answered as lexicographers and linguists explored the semantics of *po-*: First lexicographers proposed lengthy catalogs of meanings, and after subsequent investigations linguists revised and condensed those catalogs, reaching a broad consensus (Isačenko 1960, Guiraud-Weber 1993, Zaliznjak & Šmelev 2000, Dickey 2007) that *po-* has five meanings, plus an

additional sixth meaning when paired with the suffix *-yva-*. These six meanings include the attenuative, delimitative, distributive, resultative, ingressive, and intermittent-attenuative senses (defined in 2.2.3). Each of these meanings reflects a sub-lexical and/or grammatical modification of the base verb to which *po-* is added. The term *sub-lexical* (Townsend 1975:118) refers to the fact that the prefix modifies the meaning of the base verb but does not add new "lexical" content, thus precluding the formation of derived Imperfectives in most cases.

The second problem posed by polysemous *po-* is the question this dissertation seeks to answer: What are the relationships (or structure) among the various meanings of *po-*? Russian verbal prefixes in general present difficult challenges for the semanticist, given that any single prefix may express several meanings that have seemingly nothing in common. Several structuralist solutions to prefixal polysemy have been proposed, most notably those by Flier (1975, 1984), van Schooneveld (1978), and Gallant (1979). Each of these structuralist descriptions seeks to explain prefixal meaning(s) in terms of a prefix's (un)markedness for a number of abstract binary features. In short, the structuralist approach seeks to handle the diversity of meanings via abstraction; a single abstract invariant meaning is posited, while all other meanings are considered contextualized derivatives of the invariant. But as Janda (1985, 1986) notes, it becomes difficult to reconnect the concrete meanings of the prefix with its invariant. In 2.2.2 I discuss the three major structuralist approaches to Russian verbal prefixal semantics mentioned above.

A brief discussion of the category of Aktionsart follows in 2.2.3. Variously known as *soveršaemosti* 'perfectivizations' (Isačenko 1960), *sposoby dejstvija* 'means of

the action', *actional classes* (Sasse 2002, Tatevosov 2002), or *procedurals* (Forsyth 1970), the category of Aktionsart (pl. *Aktionsarten*, from German 'types of action') consists of sub-lexical modifications of base verbs by prefixation. Historically the study of Aktionsart has been intimately connected to the study of *po-*: Having lost all its originally spatial meanings (see Dickey 2007 for discussion), five of the six present-day meanings of *po-* have been analyzed as various types of Aktionsarten (Isačenko 1960; Zaliznjak & Šmelev 2000). Although there has been some debate about the linguistic reality (or at least the utility) of Aktionsart as a category for Russian verbs (Krongauz 1998), this debate does not concern us here, since the meanings of *po-* exist regardless of their classification as Aktionsarten or simply as sub-lexical or grammatical modifications of the base verb, and also because I do not analyze *po-* with respect to any other prefixes. A detailed explanation of each meaning of *po-*, with examples, concludes 2.2.3.

Cognitive linguistics (2.3) tackles the problems of prefixal polysemy by appealing to mechanisms that play a part in other human cognitive processes, such as categorization based on family resemblances, metaphor, and metonymy. Unlike more traditional linguistic paradigms, cognitive linguistics holds that linguistic cognition is a subset of human cognition in general; the rules that govern the one apply to the other, and there are no functionally compartmentalized linguistic modules in the mind. As such cognitive linguistics attempts to explain language in a way consistent with current knowledge about human cognition – this goal, frequently referred to as the "cognitive commitment" (Lakoff 1990), encourages the formulation of psychologically plausible (while not necessarily psychologically real) theories about language phenomena. Section 2.3.1 details the intimate links between cognition and our shared experiences living in human

bodies. How that experience influences linguistic expression becomes the topic of 2.3.2. Several concepts basic to cognitive linguistics are outlined in 2.3.3, including radial categories, image schemas, metaphor, and metonymy – all of which play a vital role in my analysis of *po-* in Chapter 4. A preview of just how these concepts can be used to structure the semantics of *po-* is given in 2.3.4 (to be taken up again later in 4.4.2).

While cognitive linguistics provides the necessary theoretical background for this dissertation, corpus linguistics (2.4) provides the methodological framework. Corpus linguistics makes use of large bodies of collected, usually non-elicited linguistic data (called *corpora*) to investigate any number of linguistic phenomena, spanning the range from morphemes to lexemes, semantics to grammar. Corpora are powerful tools in the linguist's arsenal for several reasons: First, corpora provide an empirical basis for the study of language. Since corpora usually consist of collected data, as opposed to data gathered by elicitation or introspection, they provide significant buffers against experimental and researcher bias. While constructing a truly representative corpus in the statistical sense is fraught with difficulty (Kilgarriff & Grefenstette 2003), most corpora aim for the more realistic goal of *balance* by drawing on materials from a large number of texts, authors, genres, and in some cases, time periods. Secondly, electronic corpora can be extensively annotated – the grammatical, syntactic, and semantic properties of each word in the corpus can be stored in a format readily accessible to the researcher. Consequently, corpora are searchable entities that the linguist can use to gather large amounts of naturally-occurring language data, which are then submitted to any number of statistical analyses that can reveal patterns not readily apparent to the human eye. Underlying this quantitative, empirically-motivated approach to linguistic research is the

assumption that distributional similarity reflects functional similarity (Gries & Divjak 2008; Divjak 2010), or in more poetic terms, "You shall know a word by the company it keeps" (Firth 1957:11). That is, a word's (or in our case, a prefix's) context provides a window onto its meaning – a window that is gaining increasing popularity among cognitive semanticists (see Chapter 3 for references).

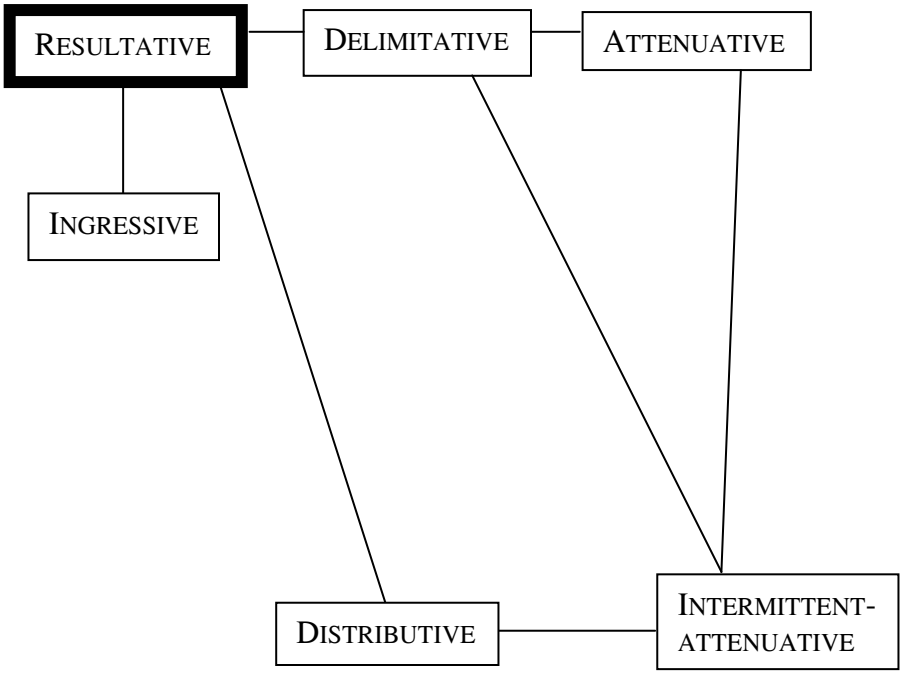
Having established the theoretical and methodological frameworks for this dissertation, in Chapter 3 I discuss all the details of how this study was conducted. I make use of the behavioral profile method (Hanks 1996; Gries & Divjak 2008; see Stefanowitsch & Gries 2003, Janda & Solovyev 2009 for similar methods), which involves selecting a corpus, collecting all records containing the item under study from that corpus (3.2), annotating the data (3.3), and subjecting the formatted data (3.4) to the appropriate statistical analyses (3.5). In this dissertation the Russian National Corpus (RNC), which at the time of data collection (July 2008) contained over 160,000,000 words, serves as my data source. To preserve the integrity of the data, however, the smaller manually-annotated portion of the RNC was used (about 2.5 million words in July 2008, now over six million words), from which 16,121 citations containing *po-* prefixed verbs were collected. Not all *po*-prefixed verbs provide insight into the semantics of *po-*, however; verbs like *polučit'*[p] / *polučat'*[i] 'receive' have *po-* in both the Perfective and Imperfective forms, and thus it is difficult to determine how *po-* modifies the meaning of the base verb – if *po-* modifies the verb's meaning at all. Thus a classification system based on the morphological behavior of *po*-prefixed verbs is presented in 3.3.1, and only those verbs where *po-* makes a discernable contribution to the meaning of the prefixed verb (i.e. expresses one of the six meanings detailed in 2.2.3)

are selected for further study. A random sample of such verbs (with their accompanying context from the corpus) are annotated for a number of grammatical and semantic parameters (3.3.2 – 3.3.4) and subjected to statistical analysis (3.4 – 3.5).

In Chapter 4 I present the results of the statistical analysis (4.1 – 4.3) and I use the conceptual tools of cognitive linguistics to interpret those results meaningfully (4.4). More specifically, I use a hierarchical agglomerative cluster analysis (HAC; see 3.5 for introduction and 4.2 for discussion) to analyze the behavioral profile previously constructed in Chapter 3. The results of the HAC analysis show that the six meanings of *po-* fall into two groups, called clusters: The first cluster consists of the attenuative, delimitative, ingressive, and resultative meanings, while the second cluster consists of the distributive and the intermittent-attenuative. This grouping remains remarkably consistent, even when sets of variables from different linguistic levels (semantic, grammatical, and discourse-level variables) are used to run supplemental HAC analyses. Using additional statistical techniques (*t*-values and *z*-scores; see 4.3 for discussion) I uncover which grammatical and semantic parameters distinguish the two clusters from one another, and which parameters differentiate the members within each cluster. In this manner a very detailed picture of each meaning of *po-* begins to take shape (4.3.1 – 4.3.2). Having determined the structure among the meanings of *po-*, I then propose solutions to two cognitive questions (4.4): First I suggest which of those meanings is most likely the prototypical member of the category (4.4.1) – namely, the resultative. There are several pieces of evidence that support the prototypicality of the resultative, each discussed in turn in 4.4.1: The resultative is among the first meanings to be grouped in the HAC analysis (4.2 explains why this is significant); the resultative is by far the

most frequently occurring meaning in the corpus, both in terms of the number of verbs expressing resultative meaning and in terms of how frequently those verbs occur (56.4% of all verbs and 67.6% of all citations in the sample); the resultative is the diachronically primary sense (Dickey 2007); and the resultative possesses the strongest family resemblances to the other members of the category. Next I turn my attention to the cognitive links among the different senses of *po-* (4.4.2). All six meanings can best be understood as metonymic (and in one case metaphoric) extensions from the resultative: The resultative indicates the full traversal of the metaphoric "path" expressed in the base verb, and via several PART-WHOLE metonymies the remaining meanings are easily motivated. The semantic structure of *po-* can thus be represented graphically by the following diagram (identical to Fig. 4.1-2):

Fig. 1-1



In Fig. 1-1 the darker box around the resultative symbolizes its prototypicality. The various metonymies and metaphors connecting the meanings are indicated by lines between the boxes. While this diagram is not to scale, the relative distance between the boxes represents the relative semantic "distance" between each meaning, as uncovered by the HAC analysis in 4.2.

Chapter 5 places the results of Chapter 4 in the broader context of cognitive and corpus linguistics and highlights the contributions of this dissertation. In sum, I have produced an empirically substantiated, psychologically plausible solution to the problem I intended to address (5.1) – namely, to discover the structure among the strongly divergent meanings of the Russian verbal prefix *po-*. More generally, I show that corpus-based studies can be just as fruitful in the investigation of the semantics of morphemes as they are in the study of independent lexemes (Gries 2006; Gries & Divjak 2008; Janda & Solovyev 2009) and abstract grammatical constructions (Stefanowitsch & Gries 2003) – empirical investigations of prefixal semantics in Russian have much to offer the cognitive semanticist. And given the importance of metonymy in the structural solution I propose, this dissertation is part of a renewed interest in the role metonymy plays not only in the lexicon, but also in grammar (Janda 2008a, 2010b, and forthcoming; Nesset 2009). Finally I point out interesting directions for future research in 5.2: On the basis of the semantic nuances uncovered for each meaning (4.3), I suggest that similar studies of other prefixes could elucidate the semantic structure of those prefixes as well, in addition to supplementing our understanding of the nature of those Aktionsarten formed by multiple prefixes (5.2.1). The other types of *po-* prefixed verbs excluded from this study (see 3.3.1) could offer further insight into the semantics of *po-*, especially from a

diachronic perspective. I conclude with a brief discussion of those remaining types and proposals for the additional study of each.

Throughout this dissertation, the following conventions will be observed: Language-specific categories will be capitalized, while cross-linguistic categories will not be. Contrast for example the category of Perfective verbs in Russian with the more generalized category of perfective that can be fitted to a number of languages. Perfective verbs will be denoted by a superscript *p*, while Imperfective verbs will be marked with a superscript *i*: *pisat'*^{i} 'write' vs. *popisat'*^{p} 'write (a while)'. Words denoting new or important concepts are **bolded** when first introduced. Conceptual categories, metaphors, and metonymies are written in SMALL CAPS. All examples introduced in the text are taken from the Russian National Corpus unless otherwise noted.

2 Prefixal semantics : A history of ideas and approaches

2.1 Introduction

In this chapter I have two primary goals: I will locate my research within the larger historical context of work on the polysemy of Russian verbal prefixes, and I will develop the theoretical background for my actual study (discussed in Chapters 3 and 4). I begin Chapter 2 by examining two historical approaches (Section 2.2) to the semantics of Russian prefixes in general and to the semantics of *po-* in particular: atomism (2.2.1) and several structuralist approaches (2.2.2). This section concludes with a look at the verbal category of Aktionsart in Russian (2.2.3) and how it relates to the six generally-accepted meanings of *po-*. I then show how the framework of cognitive linguistics can aid our understanding of prefixal polysemy (2.3). After explaining important concepts (2.3.1 – 2.3.3), I present a semantic analysis of *po-* from a cognitive perspective (2.3.4). A survey of useful concepts from the field of corpus linguistics (2.4) follows. In the final section of this chapter (2.5), I discuss the cognitive-corpus approach adopted in this dissertation and how that approach is well-suited to accomplishing the central aim of this dissertation: to explain the structured relationships among the meanings of *po-* in a methodologically sound, psychologically plausible, and empirically verifiable way. Having concluded the theoretical discussion in Chapter 2, I move to a detailed explanation of my methodology in Chapter 3, followed by a discussion of results in Chapter 4, and general conclusions in Chapter 5.

2.2 Historical approaches to prefixal semantics: Atomist & structuralist perspectives

In 2.2 I examine how two approaches to prefixal semantics (atomism and structuralism) have handled the polysemy of *po-*. First I discuss the atomist approach (2.2.1), which seeks to provide detailed catalogs of meanings but fails to elucidate the structure among those meanings, in addition to lacking a "cut off" mechanism for determining when a usage in context constitutes a distinct sense. Next I discuss three well-known structuralist approaches to prefixal meaning (Gallant 1979; Flier 1975, 1984; Van Schooneveld 1978; see 2.2.2). The structuralists attempt to resolve the lack of structure in the atomist approach by positing a single invariant meaning for each prefix, but because these invariant meanings are usually very abstract, it becomes difficult to relate them to concrete usages of those same prefixes. I also provide a brief excursus on the concept of markedness and how it relates to the structuralist approaches to prefixal semantics in 2.2.2.1. The notion of Aktionsarten, or sublexical senses of prefixes, is the topic of 2.2.3, where I describe the six widely-accepted meanings of *po-* that are used later in statistical analyses (see Chapters 3 & 4).

2.2.1 Atomism

The approach to prefixal meaning found in reference works can aptly be termed the atomist approach. **Atomism** (the traditional approach) entails the creation of detailed lists of all meanings of a prefix – no matter how obscure or similar to the other meanings of that same prefix. While this approach does provide a thorough description of prefixal meaning(s), it does not illustrate the prefix's semantic structure: No inter-meaning relationships are pointed out, and thus the meanings of a prefix seem like independent

atoms, completely isolated and self-sufficient. Atomism may leave one with the (false) impression that meanings of a prefix are indeed unrelated, having come into existence by mere happenstance; for the language student (and sometimes the linguist) the task is to learn these lists by heart and hope for the best, without asking too much about why and how these meanings came to be. A comparison of the entries for *po-* from several reference works will illustrate the questions left unsolved by the atomistic approach:

1. *Grammatika russkogo jazyka* [Grammar of the Russian Language] (1960)
   8 meanings

2. *Prefiksacja czasownika we współczesnym języku* [Prefixation of the verb in the modern language] (Bogusławski 1963)
   9 meanings, 15 sub-contexts for meaning 6

3. *Slovar' russkogo jazyka v četyrex tomax* [Russian Language Dictionary in four volumes] (1959)
   9 meanings, plus simple perfectivization

4. *Slovar' sovremennogo russkogo literaturnogo jazyka* [Dictionary of the contemporary Russian literary language] (1950-1965)
   6 meanings

5. *Russian Grammar* (Unbegaun 1967:258-259)
   5 meanings

6. *Tolkovyj slovar' russkogo jazyka* [Explanatory dictionary of the Russian language] (Ožegov & Švedova 1996)
   5 meanings

7. *Russian: A practical grammar with exercises* (Pulkina & Zakhava-Nekrasova 1974)
   5 meanings

8. *The Oxford Russian Dictionary* (1992)
   3 meanings

9. *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim* [Grammatical structure of Russian compared to Slovak] (Isačenko 1965)
   5 Aktionsarten (meanings) listed that use *po-*

10. *Vvedenie v russkuju aspektologiju* [Introduction to Russian aspectology] (Zaliznjak & Šmelev 2000)
    5 Aktionsarten listed that use *po-*

11. *Russkij glagol: formy i ix funkcii* [The Russian verb: Form and Function] (Andrews et. al 2004:105-106)
    5 meanings

We see that lexicographers and grammarians assign *po-* between three and nine meanings – or even more if you count the sub-contexts. So how many meanings does *po-* have? Atomism unfortunately does not provide a stable answer.

To what extent do these various inventories of meanings overlap? No one reference work lists all the potential meanings of *po-*; while none of the works I surveyed listed more than nine meanings, among them all I found 12 distinct meanings. Even Bogusławski's (1963) seemingly exhaustive text lacks five meanings found in other works. Four meanings are given in only one of the works surveyed. Here is a list of the meanings I encountered, followed by the number of reference works that mentioned those meanings:

1. Delimitativity[1]: The action is limited in time, without natural endpoint; often the action has short duration. (10)

2. Distributivity: Either the action is performed by a number of subjects, or it is directed towards a number of objects. (10)

3. Completion of an action: The action has reached its (natural/expected) result; "simple perfectivity"; the resultative meaning. (9)

4. Ingressivity: The *po-*prefixed verb indicates the inception of the action named by the verb. (9)

[1] Three reference works recognize an overlap between the delimitative and the attenuative meanings of *po-*.

5. Intermittent-attenuative: The action occurred with interruptions and with weakened intensity. Several works do not include the stipulation "with weakened intensity". (9)

6. Attenuative: The action occurs with less-than-usual intensity. (8)

7. Completion of an action in one attempt/motion; short duration, momentaneousness. (3)

8. Change in spatial conditions or characteristics. (2)

9. Incrementality: The action is completed gradually, incrementally, not all at once. (1)

10. Directed, goal-oriented motion; specific to verbs denoting some sort of movement. (1)

11. Acquisition of a quality, property. (1)

12. Specification of the action/emphasis: The prefix *po-* indicates that the action is occurring at that very moment, as opposed to a usual or habitual occurrence. (1)

This list is reproduced, with explanatory examples, as Appendix 1: Meanings of *po-* culled from the reference works surveyed.

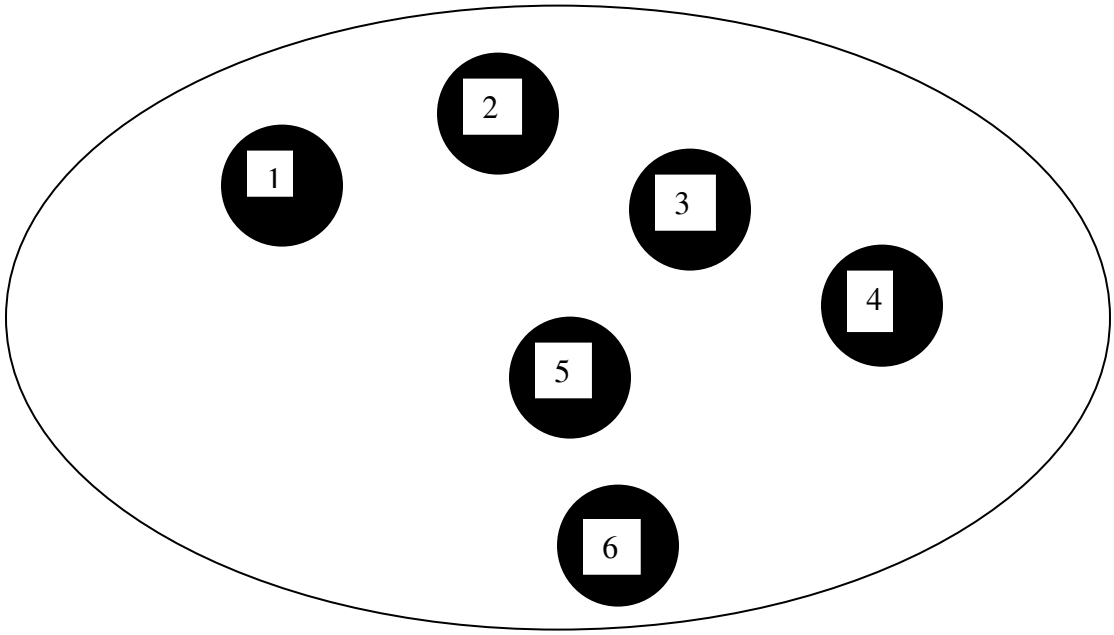
Meaning number three ("simple perfectivity") is of particular interest to our present discussion. The problem of "empty" prefixes has long been a topic of debate in Slavic linguistic studies (Forsyth 1970, Tixonov 1958, Isačenko 1960; note also Janda & Nesset forthcoming and the "Exploring Emptiness" project underway at the University of Tromsø). Scholars can be divided into two camps with respect to this issue: those who believe in empty prefixes, and those who believe that no prefix is ever semantically empty. The first camp is epitomized by Tixonov (1958, 1961, 1962), who maintains that some prefixes (*po-* and *s-* being prime examples) do not add any semantic content to the prefixed verb. Instead, these prefixes function only to form the Perfective counterpart of a verb: *po-* + *blagodarit'*^{i} 'thank' → *poblagodarit'*^{p} 'thank', *s-* + *delat'*^{i} 'do'→ *sdelat'*^{p}

'do'. Forsyth (1970) accepts the existence of empty prefixes, but with the stipulation that if a prefixed verb has a derived Imperfective partner (*pridumat'*^{p} → *pridumyvat'*^{i} 'think up'; cf. *dumat'*^{i} 'think'), then the prefix cannot be empty – if the prefix were empty, then why come up with a derived Imperfective instead of using the simpler, unprefixed Imperfective? Isačenko (1960) highlights this very issue in his argument against the idea of "empty" prefixes: If the prefix were truly empty, then the unprefixed (Imperfective) base verb and the derived Imperfective would be synonymous. And if these forms were synonymous, one would have to agree that *čitat'*^{i} 'read' is synonymous with *pročityvat'*^{i} (derived from *pročitat'*^{p} 'read (through)'). Unfortunately, subjective judgments of synonymy or near-synonymy can never lay this debate to rest; at present the Exploring Emptiness project at the University of Tromsø holds best promise for resolving the problem of "empty" prefixation.

Returning to our list of meanings of *po-*, one can easily see that the atomist approach to prefixal meaning can be described as a simplified set-theory approach to prefixal semantics[2]. Using the framework of set theory, we can say that the separate senses of *po-* are members of the set "Meanings of *po-*" and we can represent a hypothetical set consisting of six senses graphically:

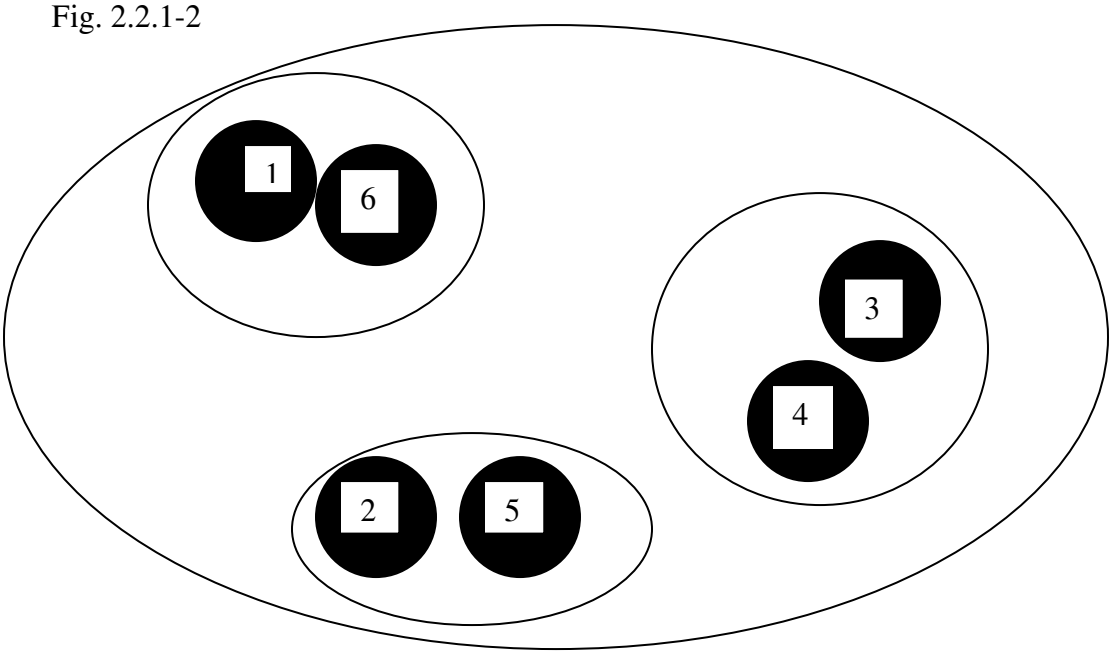
[2] It is true that set theory evolved as a mathematical tool and was not directly employed by the "atomists" in their approach to prefixal semantics. Nevertheless, set theory and atomism are so similar in logic that I think describing the latter in terms of the former is justified; see Janda (1985) and Ludwig (1995) for similar descriptions. It should also be noted that the version of set theory presented here is quite simplified and should not be taken as representative of the status of set theory in linguistics today.

Fig. 2.2.1-1



In this figure we see the meanings of *po-* (represented by the numbers in each black circle) as members within the set "Meanings of *po-*", and anything outside the boundary of the set (the thin line of the larger oval) is not a meaning of *po-*. While this approach possesses some appeal, it is not capable of describing the relationship(s) among the various meanings. For instance, the attenuative and delimitative meanings are widely recognized as related – both limit the action in some way, whether in time or intensity. Perhaps meanings such as "multiple performances of the action on multiple objects or by multiple subjects" and "action performed with interruption/repeatedly" (numbers two and five in our list above) are also related. Given that some meanings are more similar than others, we could revise the hypothetical set in Fig. 2.2.1-1, grouping similar meanings into subsets of the larger set "Meanings of *po-*":

Fig. 2.2.1-2



But the problem remains: We may have established a relationship between meanings such as attenuation and delimitativity (both of which limit the action in some way), but we still have no established relationship between disparate meanings of *po-*, like delimitativity and distributivity (that is, performance of an action on multiple objects, or by multiple subjects). Fig. 2.2.1-2 groups together meanings like delimitativity and distributivity but leaves us to believe they are related only because both meanings are expressed by the same morpheme. But why would this single morpheme *po-* express such widely divergent meanings? This is not a case of two historically different forms falling together; as far as we know, there was only one morphological *po-* in early Slavic (see Němec 1954 for a discussion of the early meanings of *po-*). Could there be a hierarchy among meanings? If so, how can one determine which, if any, meanings are primary, and which are best relegated to the status of context-conditioned sub-meanings? Atomism unfortunately offers no answers to these questions.

Leaving aside the issue of relationships among meanings, we are still confronted with the fact that atomism possesses no theoretical mechanism to explain another issue: How do speakers know which meaning is intended by a *po*-prefixed verb, given that speakers have so many to chose from? For example, when the prefix *po-* is added to the verb *guljat'*[i] 'walk about, stroll', the resulting *poguljat'*[p] means 'to walk about, stroll for a while'. Another verb, *stroit'*[i] 'build', when prefixed by *po-*, attains the meaning 'build to completion'. Other verbs depend on context to disambiguate. Take the verb *dumat'*[i] 'think' for example. It can have at least two different meanings when prefixed by *po-*:

> ***Podumal***[p] *i predložil Mukovskomu vstupit' v rok-gruppu.*
> 'He **thought** about it **a while** and invited Mukovskij to join the rock-group.' (delimitative meaning)
>
> *Ja ešče* ***podumala***[p]*: nu ego k čertu, pust' zvonit, ne do nego.*
> 'Furthermore, I **thought**: Well, to hell with him, let him call, I don't care about him.' (resultative meaning)

Perhaps even more interesting are verbs that behave like *stavit'*[i] 'put, place (upright)' when prefixed by *po-*: Whereas *postavit'*[p] usually means 'put, place (upright)' it can also mean 'supply, provide', and in that meaning it has the derived Imperfective *postavljat'*[i] 'supply, provide'. Here the meaning of *po-* in *postavljat'*[i] is no longer clear, as it seems that in this figurative sense the meanings of the prefix and stem fuse to such a degree that speakers must derive a new Imperfective, instead of simply assigning the figurative meaning 'supply, provide' to the simplex *stavit'*[i]. Overall, the method by which speakers select a specific meaning from all potential meanings of *po-* remains to be explained.

Finally there is one additional problem not resolved within the atomist framework: If you look in virtually any Russian dictionary, many prefixes share the same meanings. How then do speakers select a single prefix to use, if the meaning they intend

could just as easily be expressed by some other prefix? For example, *Russian Grammar* (1982) lists the meaning *dovesti do rezul'tata/konca dejstvija* 'bring to the result/end of the action' for 22 of the 25 Russian prefixes (both productive and unproductive prefixes; only *v-*, *nad-*, and *pere-* are not assigned this meaning). The Ožegov-Švedova (2005) dictionary lists "*načalo dejstvija* (beginning of a process)" as a meaning for the prefixes *voz-*, *za-*, and *po-*, and yet *za-* is used in well over 100 verbs to express this meaning while only about 35 *po-*prefixed verbs and a handful of *voz-*prefixed verbs express the meaning "beginning of a process". If there is no difference in meaning among these prefixes, then how does one explain this difference in usage frequency? Mere chance seems unlikely. Janda (1985) offers an interesting case study that shows only a low degree of interchangeability among *za-*, *pere-*, *do-*, and *ot-* in verbs where more than one prefix could potentially be used to express the meaning of 'excess' (a meaning assigned to each of these prefixes in various references) – it appears that the meaning of the base verb and perhaps even the context of the verb (presence/absence of adverbs, properties of the complements, other modifiers) favor the use of one prefix to the exclusion of the others. Again, the atomist approach lacks a mechanism to handle the disambiguating effects of base verb semantics and larger context, and cannot explain the complex interplay of prefixal meaning and other factors without resorting to even longer catalogs of meanings and sub-meanings.

To sum up, the atomist approach provides the necessary foundation for further investigation of Russian verbal prefixal meaning, while leaving some questions unanswered: Which meanings of a prefix are primary? Can all meanings listed by lexicographers/grammarians be considered separate, independent meanings of *po-*? How

can this one morpheme indicate meanings as divergent as "start of a process" and "multiple instances of an action"? Is there any relationship between these meanings at all? If so, what is it, and how is it motivated? Lexicographers, grammarians, and scholars such as Bogusławski (1963) have done a superb job of delineating the endpoints in this web of meaning, but the threads connecting each meaning are left hidden from view. Such is the task that the structuralists undertook in their treatment of prefixal polysemy. We will now turn our attention to how the work of structuralist scholars shed some light on these issues – and to the new questions raised by the structuralist approach.

2.2.2 The structuralist approaches

Recognizing the need for a clearer understanding of prefixal semantics, a number of scholars (Flier 1975, 1984; Gallant 1979; van Schooneveld 1958, 1978) have sought unity of prefixal meaning within a structuralist framework. Whereas atomism created catalogs of unrelated meanings for a given prefix, structuralism posits a single abstract meaning for each prefix, usually defined in terms of the presence or absence of binary semantic features. Structuralism arose as a new approach to linguistic inquiry in the Prague Linguistics circle in the 1930s; Roman Jakobson, Nikolai Trubetzkoy, Sergei Karcevskiy, Jan Mukařovský, and Vilém Mathesius were among the founding members. Although originally intended for the study of linguistic objects, structuralism quickly expanded to deal with problems in a variety of fields – anthropology, literature, psychology – however, I will not attempt to summarize such a broad movement here. Rather, I will focus specifically on how some scholars used the structuralist framework to investigate Russian prefixal semantics – a historical departure from the atomist approach

of earlier works. The structuralist approach presents prefixal meanings as highly abstract, often in terms of the presence/absence of (or un/markedness for) binary features. This abstraction is a mechanism intended to handle the diversity of meanings presented by any given prefix, often with an eye toward geometrically symmetrical representations (cf. van Schooneveld 1978). The Danish linguist Louis Hjelmslev (1935), in discussing the semantics of case in Russian, makes a remark that is equally representative of the structuralist study of Russian prefixes: "A case, like linguistic units in general, does not mean several different things; it means one single thing – it carries a single abstract concept, from which concrete applications can be derived". Thus we can see how far the linguistic pendulum has swung: Scholars like Bogusławski (1963) meticulously dissected the meanings of *po-*, seeking a semantic description that would explain as many contexts as possible. Structural analyses of Russian verbal prefixes, however, seek to uncover the underlying *unity* in prefixal meaning; the contextual meanings given by Bogusławski (1963) simply fall out as a natural consequence of that underlying meaning interacting with context. Here I will discuss the contributions of the three scholars mentioned earlier – Gallant (1979), Flier (1975, 1984), and van Schooneveld (1958, 1978). First I will briefly divert our attention to the concept of "markedness" (2.2.2.1) and what it means within a structuralist framework, and then I will show how structuralist approaches advanced our understanding of Russian prefixal semantics (2.2.2.2 – 2.2.2.4). In so doing I will point out the new questions raised (and left unanswered) by structuralism, followed by a summary of the structuralist contributions to the topic (2.2.2.5). This summary will serve as a segue into our discussion of how the cognitive and corpus linguistic frameworks integrate the contributions of the atomist and structuralist approaches into a

new understanding of prefixal polysemy, an understanding that actively engages findings in human psychology and neuroscience.

2.2.2.1 Overview of markedness

In order to understand a discussion of structuralist approaches to prefixal semantics, one must first have an understanding of the concept of **markedness**, which can be broadly defined as "an asymmetric relationship between two or more elements" (Janda 1995), that is, a relationship between the marked item(s) and the unmarked item(s). The term markedness was first used by two pioneering members of the Prague Linguistic Circle, Roman Jakobson and Nikolai Trubetzkoy, and has since found its way into many branches of linguistic inquiry, from structuralism to generative grammar, from phonology all the way to cultural studies (Moravcsik & Wirth 1986). As we will see later, the concept of markedness fits nicely within the cognitive linguistic framework as well, with the unmarked term corresponding to category members closer to the prototype and marked terms corresponding to more peripheral category members (Janda 1993, 1995).

From the vast literature on markedness, Haspelmath (2006:25) distills at least twelve different senses of "markedness". In our discussion of structuralism, we will restrict our notion of markedness to the realm of semantics while ignoring the meaning(s) of markedness in other spheres. Jakobson's (1971:3-4) well known example *osel* 'donkey' vs. *oslica* 'donkey' illustrates the concept well: *oslica* is marked for feminine sex, meaning that the use of the term always signals to the listener that the donkey is female. *Osel*, on the other hand, is not marked for sex at all: Though grammatically masculine, *osel* can also refer to a generalized donkey whose sex is not known (or just

irrelevant). As Haspelmath (2006:29) sums it up, "the difference between marked and unmarked is not between A and non-A, but between A and *indifference to non-A*" (emphasis added). Under the convention for representing markedness, *oslica* would be described as [+female] while *osel* would be [-female]. Again, [-female] does not mean "not female" but rather "indifference to whether female or not-female". The quality [female] is referred to as a **feature** (a notion borrowed from phonological distinctive feature theory), and the opposition of the type [+female] vs. [-female] is called a **privative** or **binary opposition**, since the terms *osel* and *oslica* must be marked [+female] or [-female] – there is no other option. As we will see in the following survey of Gallant (1979), Flier (1975, 1984), and van Schooneveld (1958, 1978), the concept of markedness plays an essential role in structuralist approaches to prefixal semantics.

2.2.2.2 Gallant (1979)

Gallant attempts to define Russian verbal prefixes in terms of a single relational feature, either [+horizontal] or [+vertical], and a number of additional "frame features". As a case study Gallant takes the prefix *vz-*, which he defines as [+horizontal] and [+transgression], the latter being one of the possible frame features. Using these features, Gallant describes the direction – whether literal or figurative – of the verbal action. For instance, adding *vz-* to the verb *kopat'*^{i} 'dig' gives *vskopat'*^{p} 'dig up', as in

> *Podrostok spal do poludnja, toroplivo zavtrakal i, poka ego ne zastavili* ***vskopat'***^{p} *grjadku u sebja na ogorode, toropilsja na pomošč' k Ane.*
> 'The teenager slept until noon, hurriedly ate breakfast and, until they forced him **to dig up** the rows in his vegetable garden, hurried to Anja's aid.'

Here the action of the verb indicates a transgression of the horizontal surface of the earth – that is, in order for it to be dug up, some implement must cut through the horizontal

surface of the soil. Aside from direct physical references, the feature-specifications of [+horizontal] and [+transgression] function together in a figurative sense as well:

> *Ja vas prošu ne šutja: kogda vam **vzdumaetsja**[p] obo mne govorit' durno, voz'mite lučše nož i zarež'te menja – ja dumaju, èto vam ne budet očen' trudno.'*
> 'Joking aside, I ask you: when it **crosses your mind** to speak badly of me, it would be better to take a knife and slit my throat – I don't think it will be hard for you.'

Here the prefix *vz-* in *vzdumaetsja*[p] 'will cross (your) mind' indicates a figurative transgression of the threshold (figuratively understood as a horizontal surface, perhaps) between what enters the mind as a thought and what simply never occurs to a person.

Gallant (1979) claims that prefixes do not add semantic content to the base verb but rather select and highlight some pre-existing meaning within the verb itself. This notion that prefixes simply *select* meaning, as opposed to *adding* meaning to the base verb, offers a partial explanation of the "choosiness" of Russian verbs – that is, verbs combine with one prefix instead of another even though both prefixes have the same apparent denotation. In Gallant's system, some pre-existing meanings within the verb itself are compatible with a prefix's feature-specification[3]. All that remains is to figure out the feature-specifications that will make a given prefix the logical choice for a given verb. For instance, if we could ascertain the exact feature-specifications for the prefixes *pere-*, *ot-*, *do-*, and *za-*, we could explain the non-interchangeability of these prefixes mentioned previously (cf. Janda 1985), despite the fact that they all seem to mean 'do to excess' when combined with some verbs. But while Gallant's theory of meaning-

[3] The notion of prefix-base compatibility is known as the "overlap hypothesis" (Janda & Nesset forthcoming; see also van Schooneveld 1958 and Vey 1952) and is often cited as the reason why the prefixes in Perfective "partner" verbs in Russian appear to be semantically empty. It is important to note that Gallant (1979:37) extends the overlap hypothesis from the realm of prefixed aspectual partner verbs (e.g. *napisat'*[p] / *pisat'*[i] 'write') to cases of lexical derivation, such as *pod-* 'under' + *pisat'*[i] 'write' → *podpisat'*[p] / *podpisyvat'*[i] 'sign'. In this respect Gallant (1979) differs substantially from the traditional understanding of prefix-base relations.

selection over meaning-addition may seem appealing, it does not explain all the data included in his study. Of the 200 or so base verbs that Gallant analyzes, 12 (6%; see Gallant 1979:445-454) of these base verbs have no meaning in common with a *vz-* prefixed form of the same verb. Compare for instance *vz-* + *brat'sja*^{i} 'undertake, begin doing' → *vzobrat'sja*^{p} / *vzbirat'sja*^{i} 'climb' – here there is no meaning already present in *brat'sja*^{i} 'undertake, begin doing' that *vz-* highlights. Rather, it would appear that *vz-* is adding the meaning 'up, upwards' to a concrete, physical understanding of *brat'sja*^{i} 'take (oneself)', thus producing (literally) 'take oneself up, climb'. Even if we were to argue that *brat'sja*^{i} has lost its original, physical meaning, making it a poor test case, the fact remains that even a purely literal, concrete meaning of *brat'sja*^{i} 'take (oneself)' does not possess a notion of upward movement for *vz-* to highlight. In short, while Gallant's (1979) feature-based analysis of *vz-* points the way towards the unity behind prefixal semantics, features alone do not suffice to explain his data.

2.2.2.3 Flier (1975, 1984)

Flier has produced two works in which he attempts to describe the semantics of Russian verbal prefixation within a structuralist framework. Here I will outline how both of these works relate to the semantics of *po-*, followed by my own review of the strengths and weaknesses of Flier's arguments.

Like Gallant (1979), Flier (1975:219) asserts that "Both prefixes and prepositions can be likened to prisms or cameras which alter the perspective of a given object. *Neither is additive*; rather, they *flesh out perspectives inherent in the object* itself" [emphasis added]. But unlike Gallant (1979), Flier (1975, 1984) seeks to define the actual semantic

content of the prefix itself. Flier (1975) employs a feature-based approach that focuses on the invariant meaning of *po-* – that is, the single underlying meaning of the prefix, of which all other meanings are context-induced connotations. While Gallant posits the two features [+horizontal] and [+transgression] to describe *vz-*, Flier specifies *po-* using the three features [+spanned, +lateral, +domainial]. The feature [+spanned] indicates that *po-* (like *pere-*, *pro-*, and *ob-*, the other three prefixes Flier examines) focuses on the inceptive, terminal, and/or lateral limits of the verbal action (which Flier calls the "domain"). *Po-* is also marked as [+lateral], which distinguishes it from *pere-* and *pro-*, both marked [-lateral]: *Pere-* and *pro-* focus on metaphoric travel from beginning to end of the verbal action, without reference to any "lateral" limits. The following examples illustrate this difference:

> *On* ***perešel***^{p} ***ulicu****, kupil gazetu i sel na skamejku*.
> 'He **crossed the street**, bought a newspaper, and sat down on the bench.'
>
> *On* ***prošel***^{p} ***čerez zal.***
> 'He **passed through the hall**.'

Both *perešel*^{p} 'crossed' and *prošel*^{p} 'passed through' ignore anything associated with the periphery of the action – instead they focus all attention on getting directly from point A to point B. According to Flier (1975), verbs in *po-* are marked [+lateral], and thus do not ignore the "lateral" aspects of the verbal action:

> *On vse veščі* ***pobrosal***^{p} *v jaščik*. (Flier 1975:225)
> 'He **threw** all of his things into the box.'

Unlike *pere-* in *perebrosat'*^{p} 'throw (one after the other), the *po-* in *pobrosat'*^{p} 'throw (all or many)' does not focus attention on the consecutive nature of the action as much as it focuses on the cumulative result that all the things ended up in the box, whether one after the other or perhaps in groups, helter-skelter.

Finally *po-* is marked [+domainial], which indicates that the prefix focuses on the part of the action within the domain; *ob-*, which is marked [-domainial], focuses on action outside of or extending beyond the domain. The difference can easily be seen in the following two examples:

> *Rebjata **pošli**[p] za nim v magazin.*
> 'The children **walked** behind him into the store.'

> *Ja pribavil gazu, **obošel**[p] gruzovik, perestroilsja ešče dal'še vlevo.*
> 'I added some gas, **walked around** the truck, moved over even further to the left.'

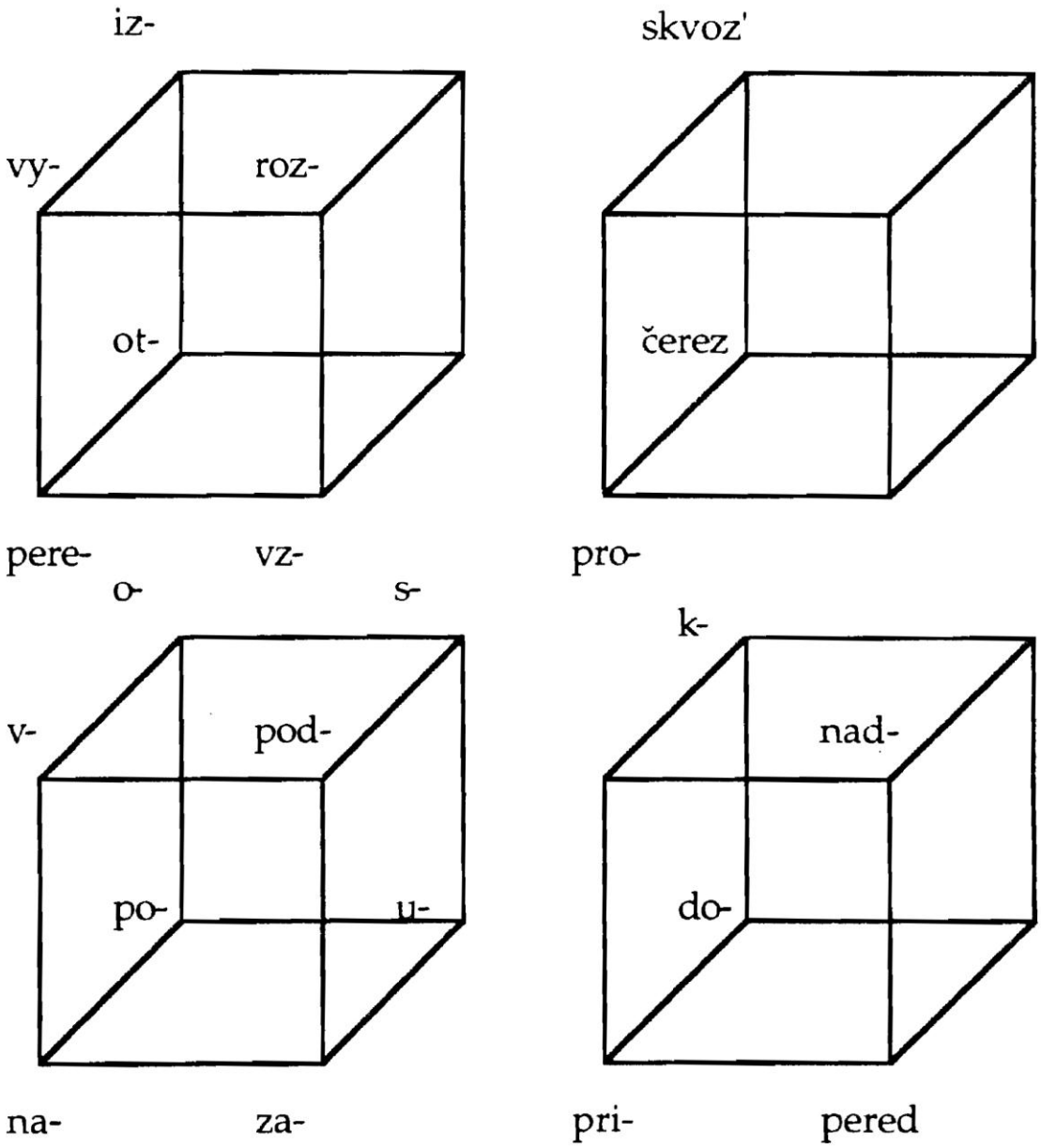
Flier (1985) extends his theoretical work on the semantics of *po-* in a subsequent article, where he states that *po-*, in its delimitative function, "combine[s] only with verbs denoting atelic activities, which are both controllable (and hence delimitable) and nonprogressive" (1985:56). With reference to the feature hierarchy, Flier (1985) notes that *po-* pays attention to the metaphoric "contours" of the action without describing its internal structure. In this way *po-* is to verbs what containers are to mass nouns (Flier 1985; Mehlig 2004, 1996): *po-* delimits the verbal action of *čitat'*[i] 'read' in *počitat'*[p] 'read for a while' much like *čaška* 'cup' delimits *čaj* 'tea' in *čaška čaja* 'a cup of tea'.

Unfortunately, Flier devotes no subsequent work to the other meanings of *po-*, and thus the problem encountered earlier remains: How do we get from a single, abstract invariant meaning to the multiplicity of meanings for *po-*? Even with his account of the *po-* delimitative, Flier (1985) makes a number of references to metaphorical interpretations of the meaning of *po-*, and he goes so far as to suggest that all verbal prefixes refer to "metaphoric spatial configurations" (1985:46). But Flier nowhere explains the mechanism by which these metaphors work, what forces constrain them, and how something as malleable as metaphor can interact with a system based on strict

hierarchies of feature markings. While Flier's (1975, 1985) work marks a significant advance over earlier descriptions of the semantics of *po-*, there remains considerable work to be done.

2.2.2.4 Van Schooneveld (1978)

Van Schooneveld's *Semantic Transmutations* (1978) is easily the most ambitious and thorough structuralist approach to Russian prefixal and prepositional semantics (he does not distinguish between the two) ever completed. Inspired by Jakobson's (1958) cubic representation of case meaning in Russian, van Schooneveld crafts a set of four cubes, the corners of which represent the combinations of features that make up a given prefix's invariant meaning. Markedness is indicated by location on the cube: The unmarked members of an opposition are at the bottom, and markedness moves first up, then right, and to the back. A visual aid will assist in seeing how the prefixes/prepositions are distributed in this system (reproduced from Ludwig 1995):



Van Schooneveld's system encompasses six features: plurality, pre-identity, verification, demarcatedness, cancellation, and objectiveness. Due to the complex nature of van Schooneveld's work, I will not attempt to explain the inner workings of this system here; suffice it to say that van Schooneveld (1978:106-107) states that *po-* is only marked [+extension], which basically means that "the object of *po-* is minimally affected by its

relation to its modified, that is, it is minimally affected by the narrated situation". From this vague definition, van Schooneveld is able to deduce various sub-meanings by relying on context and pragmatic inferences. Putting aside a few creative adaptations of data, one still sees that van Schooneveld's cubic semantic representation encounters the same difficulties that Gallant (1979) and Flier (1975, 1985) grapple with: How does one derive the specific instantiations of meaning from the abstract invariant? The relationship between invariant and concrete usage remains tenuous.

2.2.2.5 Summary of the structuralist approaches

The structuralist approaches to prefixal semantics proposed by Gallant (1979), Flier (1975, 1985), and van Schooneveld (1978) constitute a significant advance over the atomist approach. Each of these three scholars seeks to establish the semantic unity underlying each prefix, and each is convinced that prefixal meaning is not a random amalgam of meanings accumulated during the development of Russian but rather a coherent, logically motivated semantic system. However, in seeking to overcome the previous unstructured approach of the atomists, the structuralists create strict semantic hierarchies and abstract geometries that are hard to connect with the variety of meanings encountered in everyday Russian usage. In addition, purely structuralist approaches to prefixal meaning are difficult to integrate with later research on human linguistic cognition. Several studies (Bierwisch & Schreuder 1992; Fodor et al. 1975; see also Dąbrowska 2004: 106-107) have suggested that comprehension is *not* dependent on decomposition of words into lexical features (for example, *bachelor* would not be understood in terms of the proposed lexical features [+/-adult], [+/-male], [+/-never

married]). Likewise, the psychological plausibility of features such as [+/-lateral] and [+/-domanial] (used by Flier (1975) to define *po-*, *pere-*, *pro-*, and *ob-*) – and whether such features are used to compute meaning – is yet to be substantiated. In light of these issues, another approach to prefixal meaning is needed, some way of bridging the gap between abstraction and instantiation, without denying the importance of either. Such an approach can be found in the framework of cognitive linguistics, to which we will turn our attention in Section 2.3. But first a few words about the category of Aktionsart and the most widely recognized meanings of *po-* are in order.

2.2.3 Aktionsarten and the meanings of *po-*

As the reader saw in Section 2.2.1 (Atomism), lexicographers and scholars have assigned *po-* a litany of meanings. However, most modern treatments of *po-* (Isačenko 1960; Guiraud-Weber 1993; Mehlig 1996; Zaliznjak & Šmelev 2000) recognize five basic meanings of this prefix: attenuative, delimitative, distributive, ingressive, and resultative. When the prefix *po-* is combined with the suffix *-yva-* in the same verb, a sixth meaning – the intermittent-attenuative – is recognized. Variously referred to as *soveršaemosti* 'perfectivizations' (Isačenko 1960), *sposoby dejstvija* 'means of the action', *Aktionsarten* 'types of action' (German), *actional classes* (Sasse 2002, Tatevosov 2002), or *procedurals* (Forsyth 1970), five of these six meanings (all but the resultative) are sublexical (Townsend 1975:118) – that is, the prefix modifies the meaning expressed by the verb, but does not add new "lexical" content. In keeping with tradition, I will refer to this class of meaning (and any of its members) as **Aktionsart** (pl. **Aktionsarten**), wherever appropriate. In this section I describe the notion of Aktionsart and its relevance

in Russian, briefly summarizing the development of the concept from the 1600s to present. I discuss a few problems with the notion of Aktionsarten in general and with the five Aktionsarten expressed by *po-*, and I present the revised stance on Aktionsart used in the remainder of this dissertation, which can be summed up thus: Whether Aktionsarten exist as a "real" Russian verbal category is a debated issue that I will not presume to answer here. Instead, I can safely say that the five Aktionsarten encoded by *po-* reflect five of the basic meanings of this prefix; *po-* has lost any historically prior spatial meanings and is left with these five sublexical meanings plus the resultative. Finally, I will illustrate the six generally-accepted meanings expressed by *po-*, pointing out the distinguishing characteristics of each.

The study of Aktionsart in Russian is closely tied to the history of the study of aspect in Russian; I will highlight a few turning points in Russian aspectology and the study of Aktionsart here (for more thorough discussion see Młynarczyk 2004). It should be noted that close ties between the German and Russian scholarly communities of the time period under consideration allowed a cross-fertilization of ideas, whereby the theories of Russian scholars influenced German aspectology, and vice versa.

The term aspect (Russian *vid*) first appeared in the work of Meletij Smotrickij in the early 1600s, where aspects were considered part of the *tense* system of Russian (see Lomonosov 1764 for a similar use of the term). The first to use the term *vid* 'aspect' to describe non-tense distinctions was by Nikolaj Greč in his 1827 grammar of Russian. However, Greč's concept of *vid* was broad and not limited to the binary Perfective-Imperfective division we know today – rather, *vidy* expressed "accessory circumstances by which are more closely defined the signification and the extent of the action" (qtd.

Binnick 1991:140). The vagueness of Greč's concept *vid* enabled German grammarians to utilize this concept in their own linguistic investigations, despite the inherent differences between Russian and German; in German Greč's *vidy* 'aspects' became known as *Aktionsarten* 'manners of action'. So by the time the German grammarian Brugman described Aktionsart as "the manner in which the action proceeds" (1904; translation mine), there were only two conceptual categories for describing verbal action in Russian and German: tense and Aktionsart (aka *vidy*). Sigurd Agrell, however, altered the scholarly discourse on aspect and tense with his 1908 work on Polish verbs, in which he formally distinguished three verbal categories: tense, aspect, and Aktionsart. Under Agrell's system, aspect included only the binary distinction of Imperfective and Perfective; other modifications of verbal action were now subsumed under the category Aktionsart. By the 1930s (Młynarczyk 2004:36-37) this three-way classification found general acceptance among Slavists, and in 1960 A.V. Isačenko produced the most detailed description of Aktionsarten in contemporary Russian. A somewhat revised description appears in Zaliznjak & Šmelev (2000), but the substance of Isačenko's description remains unchanged. Isačenko's (1960) work forms the basis of the discussion that follows.

Isačenko (1960: 210) defines Aktionsart as "a certain semantic modification of a verb, indicating exactly how the action expressed by the verb is completed". For example:

> *Na Ukraine bešenyj kot* ***pokusal***^{p} *četyrex čelovek.*[4]
> 'In Ukraine a rabid cat **bit** four people.'

[4] <http://www.dni.ru/incidents/2009/10/20/177633.html>

In this sentence the verb *pokusal*^{p} 'bit' belongs to the Distributive Aktionsart, which indicates that the action is distributed over a number of objects (or people, as in this case). Here the prefix *po-* not only perfectivizes the verb, but modifies the way the action is completed. Like many Aktionsarten, the Distributive Aktionsart can be formed by more than one prefix – in this case the prefix *pere-* also expresses distributive meaning, albeit of a slightly different shade.

The meanings expressed by Aktionsarten fall somewhere between "empty" perfectivization (that is, the prefixed verb is semantically identical to the unprefixed verb, save for the change in aspect) and the creation of a verb that no longer means the same as its unprefixed counterpart. In Russian three types of prefixal modification are generally recognized: "empty" prefixation, sublexical (Aktionsart) prefixation, and lexical prefixation. In **"empty" prefixation**, the prefix simply perfectivizes the verb and adds no additional semantic or grammatical information. Example: *po-* + *darit'*^{i} 'give (as a gift) → *podarit'*^{p} 'give (as a gift) [perfectivization only]. While the "emptiness" of the prefix in such cases is the traditional view (Tixonov 1998, Vinogradov 2001), many scholars challenge the notion that the prefix is semantically "empty" in such cases (Komárek 1984; Dickey 2006:12; Janda 2007; Zaliznjak & Šmelev 2008:81-82; Janda & Nesset forthcoming). Instead, they posit that the meaning of the prefix overlaps with the meaning of the verb so much that the prefix only *appears* to be "empty". This view is known as *subsumption* or the *overlap hypothesis*, and is the view that I adopt here. The prefixed and unprefixed verbs in cases of "empty" prefixation are traditionally thought to form an aspectual pair (Vinogradov 1938; Šaxmatov 1941; Maslov 1948; Bondarko 1983; Isačenko 1960; Zaliznjak & Šmelev 2000; Dickey 2006).

In **lexical prefixation** not only does the prefix perfectivize the verb, but it adds new semantic content as well. Example:

pere- 'through, again' + *čitat'*^{i} 'read' → *perečitat'*^{p} 're-read'

The addition of new semantic content often allows the derivation of an Imperfective verb by suffixation (in bold):

perečitat'^{p} / *perečit**yva**t'*^{i} 're-read'

In cases of lexical prefixation the spatial meanings of prefixes are often apparent:

pro- 'through' + *idti*^{i} 'go, walk' and *xodit'*^{i} 'go, walk' → *projti*^{p} / *proxodit'*^{i} 'go, walk through'

vy- 'out' + *kupit'*^{p} 'buy' → *vykupit'*^{p} / *vykupat'*^{i} 'buy out, ransom'

As noted before, Aktionsarten are the result of **sublexical prefixation**, whereby the prefix adds new semantic content to the verb, but not enough new content to distinguish the newly-prefixed verb as a separate lexical entity – consequently, formation of Imperfectives by suffixation is usually not allowed. Aktionsarten are formed with a variety of prefixes, and the number of Aktionsarten recognized in Russian ranges from 11 (Isačenko 1960) to 14 (Zaliznjak & Šmelev 2000). More than one prefix may form a given Aktionsart – for instance, the Distributive Aktionsart (which indicates that the action is performed upon multiple objects or by multiple subjects) can be formed either by *po-* or *pere-*, with different verbs preferring one prefix over the other. In Aktionsarten the spatial meanings of prefixes are seldom if ever expressed; instead, the prefix modifies the verb's semantics with respect to time or intensity. Compare for instance:

- *projti*^{p} / *proxodit'*^{i} 'go through' (*pro-* 'through') Lexical prefixation of the original *idti*^{i} / *xodit*^{i} 'go'. The spatial meaning is evident.

- *prosidet'*^{p} 'sit (for a specific amount of time)' Sublexical prefixation – an example of the Perdurative Aktionsart.

We could thus generalize as follows: Aktionsart is a verbal category that is identifiable by a morphological/derivational criterion and a semantic criterion: Aktionsarten include prefixed (or, less frequently, suffixed) verbs that cannot form a derived Imperfective[5]. The prefix (or suffix) introduces a sublexical change into the prefixed verb; this sublexical change can be a modification of the verb's meaning with regards to time or intensity, but not with regards to space or other domains.

Several problems exist with individual Aktionsarten and with the notion of Aktionsart itself in Russian. First, some prefixed verbs can have either an Aktionsart meaning or express simple perfectivity, depending on the context. For example:

> *Ja* ***podumala***^{p} *– možet štany nadet'.*
> 'I **thought**, "Maybe I'll put on pants."' [simple perfectivity]

> *Ja ešče* ***podumala***^{p} *o tom, čto èto očen' real'no i v Samare, gde očen' mnogo rynkov.*
> 'I **thought** some more about how that's also very feasible in Samara, where there are a lot of markets.' [Delimitative Aktionsart]

One could argue that "simple perfectivity" is in many ways semantically equivalent to the Resultative Aktionsart (Guiraud-Weber 1993:58, Townsend 1975:121). In that case, my point here becomes that some verbs are semantically ambiguous and can represent one or more Aktionsarten or meanings. Again, only context (not morphology or general semantics of the isolated verb form) can distinguish between these meanings:

[5] The one obvious exception is the Intermittent-attenuative Aktionsart, which is Imperfective and thus secondarily derived Imperfectives are of no concern in this case.

Na "Skoroj pomošči" Artema, k pimeru, zadolžennost' po zarplate kolossal'naja. Dva-tri mesjaca medikam den'gi ***povyplačivajut***^{p}*, i opjat' zatiš'e.*[6]
'The Artema Emergency Service, for instance, has gone hugely in debt paying wages. For two or three months they **will pay out wages** to the medics, and again all is quiet.'
(Delimitative Aktionsart: indicates the action occurred for some time.)

My ***povyplačivali***^{p} *vse položennye den'gi – xotja èto bylo nelegko.*[7]
'We paid out all the money owed – although it wasn't easy.'
(Distributive Aktionsart: indicates that the action involves a multiplicity/all of the objects.)

Not only are verbs frequently ambiguous between purely Perfective and Aktionsart meaning or between two different Aktionsart meanings, but the line between lexical and sublexical modification can itself be blurry. As Townsend aptly notes,

> [I]t is frequently difficult to decide whether a given prefixed perfective should be characterized as lexical or sublexical. The possibility of formation of a derived imperfective, which is sometimes adduced to prove the presence of a new "independent" meaning, is an unreliable criterion, for many obviously sublexical types are capable of forming derived imperfectives, whether or not dictionaries list all of them[8]. In the case of many prefixes, sublexical and lexical meanings will seem to overlap, and one may argue whether a lexical change has taken place or whether the action has merely been modified in some way with respect to time or intensity. (1975:121)

In a similar vein Krongauz (1998) criticizes the current understanding of Aktionsart as a generalization that obscures pertinent linguistic facts: "The simplicity of the system of temporal Aktionsarten…turns out to be no more than a pleasant illusion" (1998:128, translation mine). According to Krongauz, the difficulty in abstracting away from the specific meanings of prefixes is immediately visible in the sub-division of Aktionsarten into sub-Aktionsarten – for instance, Zaliznjak & Šmelev (2000) divide the

---

[6] Internet example, cited LeBlanc (2006). The phrase *dva-tri mesjaca* 'for two or three months' indicates that the verb in this example expresses delimitative meaning – phrases indicating that the action occurred *nekotoroe vremja* 'for some time' are considered common indicators of delimitative meaning.
[7] Internet example. The quantifier *vse* 'all' in the direct object frequently accompanies distributive verbs, as in this example.
[8] Townsend (1975) notes verbs prefixed in *pro-* in particular as capable of forming derived Imperfectives that are not always recognized by major dictionaries.

Ingressive Aktionsart into the Inchoative (prefix *za-*) and Ingressive proper (prefix *po-*, *vz-*). Unlike the category of Aspect, which applies to all verbs in Russian regardless of semantics, Aktionsarten seem to be highly selective of which verbs are admissible. The prefix *za-* is used to form the Ingressive (Inchoative) Aktionsart, which marks the beginning of an action. However, *za-* cannot be affixed to just any verb to indicate this meaning, and it appears that the temporal structure of the verbal action is not sufficient to account for this selectivity. Compare, for instance, the acceptable *Motor* ***zarabotal***^{p} 'The motor started up [started working]' versus the strange/unacceptable **Ja* ***zarabotal***^{p} *v sem' časov* 'I started working at seven o'clock' (Krongauz 1998:126). To account for this difference in acceptability, we must appeal to connotations and additional semantic content associated with *za-*; temporal structure of the action is insufficient. For these reasons Krongauz finds the idea that Aktionsarten represent an independent verbal category erroneous.

A final critique on the viability of Aktionsart as a verbal category in Russian centers on the multi-prefixal nature of many Aktionsarten. If we look at Zaliznjak & Šmelev's (2000) catalog of Aktionsarten, we find that of the 14 Aktionsarten listed, six are formed with more than one prefix. The Resultative-intensive Aktionsart, for instance, can be formed by the prefix/suffix combinations *do-…-sja*, *za-…-sja*, *raz-…-sja*, *iz-…-sja*, *u-…-sja*, or *vy-…-sja*. Much is already known about the original spatial meanings of prefixes, and how these meanings could have given rise to more abstract meanings (Gallant 1977; Janda 1985 & 1986; Dickey 2007). But to date no one has convincingly motivated or explained how disparate prefixes came to express one and the same Aktionsart. Nor has anyone demonstrated that the semantic similarities among

differently-prefixed verbs of the same Aktionsart are greater than their dissimilarities, justifying their inclusion in the same Aktionsart. That is, the prefix/suffix combinations *do-…-sja*, *za-…-sja*, *raz-…-sja*, *iz-…-sja*, *u-…-sja*, or *vy-…-sja* of the Resultative-intensive Aktionsart are probably not semantically identical. While all these combinations can potentially express the meaning of "excessiveness", does this similarity warrant grouping verbs with these affixes under a single Aktionsart? What of the differences between these prefix/suffix combinations – how are they weighted against their similarity? An empirical justification of the category is currently lacking.

Answers to questions regarding the linguistic reality or validity of Aktionsart in Russian are far beyond the scope of this dissertation. In fact, an answer is not necessary to investigate the semantics of *po-*. *Po-* has lost its original spatial meaning in modern Russian, and is left with six widely recognized meanings (if one includes the prefix/suffix combination *po-…-yva-,* as I do in this dissertation). Although five of these meanings[9] are themselves considered prime examples of Aktionsarten, their status as Aktionsarten is irrelevant since I will not be analyzing them with respect to other prefixes; that is, all the meanings of *po-* exist, regardless of whether we call them Aktionsarten or simply "meanings", and indeed virtually all the literature to date equates the separate senses of *po-* with its Aktionsarten. For the purposes of this dissertation I accept the six recognized meanings of *po-* as the starting point of my analysis; I do not seek to determine anew what the meanings of *po-* are, since the existing scholarship has already succeeded in doing so. Instead I aim to determine what the relationships among those meanings are – a problem that has thus far eluded a satisfactory solution. Along the way I uncover

[9] That is, all but the resultative meaning, which is often equated with "simple perfectivity".

additional trends associated with each meaning (4.3), but this is incidental to my larger goal. In the pages that follow I sketch the semantics of each of these six meanings: resultative, delimitative, attenuative, distributive, ingressive, and intermittent-attenuative.

Meanings of *po-*:

1. **Resultative**. For the purposes of this dissertation, the resultative meaning of *po-* is the same as the "empty" meaning of the prefix (Guiraud-Weber 1993:58, Townsend 1975:121): In this meaning *po-* indicates that the action expressed by the verb was carried out to completion. Verbs in the resultative meaning are formed from telic predicates. Resultative verbs are always Perfective and are the "natural Perfectives" (Janda 2007:624) of many verbs – that is, the prefixed Perfective verb is denotationally equivalent to the unprefixed Imperfective form. Besides perfectivizing the verb, *po-* provides no readily apparent semantic content. Example:

> *Gorbačeva nado vašego* ***povesit'***$^{\text{P}}$*!*
> 'They need to **hang** your Gorbačev!'

2. **Delimitative**. In this meaning the prefix *po-* puts a boundary on (delimits) an action that does not have an inherent endpoint. Although often translated as "do X for a (little) while", the time period delimited does not necessarily have to be short. Delimitatives are formed from verbal predicates that are activities (Vendler 1957), meaning that they cannot be completed and have duration (Janda 2007). Having no inherent goal or endpoint, these actions are considered homogenous, akin to mass nouns in Russian (Mehlig 1996, 2004) – any sub-phase of the action would be identical to the action as a whole, in much the same way that a cup of milk is identical to the milk in a gallon jug,

differences in volume aside. Verbs of the delimitative meaning are most frequently formed by adding *po-* to an unprefixed Imperfective verb, but in rare cases this meaning also obtains when *po-* is added to a prefixed derived Imperfective. Here is an example of each type, respectively:

> *Davajte* ***pogovorim***^{p} *predmetno vse-taki o vzryve v Moskve.*
> 'All the same let's **talk (a while)** seriously about the explosion in Moscow.'
>
> *Dlja vsex želajuščix pogonjat' mjač, pokidat' ego v korzinu ili, na xudoj konec,* ***poperebrasyvat'***^{p} *ego čerez setku otkryty dveri sekcij po futbolu, basketbolu i volejbolu.*[10]
> 'For everyone desiring to chase after a ball, to dunk it in a basket or, if worse comes to worst, **to spend some time tossing** it through a net, the doors of the football, basketball, and volleyball sections are open.'

Po-prefixed verbs in the delimitative meaning are always Perfective and do not form derived Imperfectives. Dickey (2006) cites functional similarities between verbs of the delimitative and resultative meanings and even points out cases where the delimitative meaning is all but absent in atelic predicates. He considers delimitatives to be the Perfective "partners" of atelic Imperfective verbs, despite the addition of sublexical content. We will return to the overlap between resultative and delimitative meaning in some verbs in Section 3.3.3.

3. **Attenuative**. The attenuative meaning of *po-* is very similar to the delimitative meaning, except here the action is not limited in time, but rather in intensity (Isačenko 1960:238-239; Zaliznjak & Šmelev 2000:120). In this meaning the prefix *po-* most frequently occurs with already prefixed Perfective verbs, resulting in a double-prefixed verb. The resulting verb means 'do X a little, slightly, at less-than-usual intensity' – note that unlike the delimitative meaning, the focus is not on the duration of the verbal action:

[10] Internet example, cited in LeBlanc (2006).

*čem-nibud' **porazvleč'**[p] bol'nogo, otvleč' ego vnimanie ot slučajnoj bolezni*[11]
'to **entertain** the patient **a little** with something, distract his attention from the incidental illness'

Po-prefixed verbs in the attenuative meaning are always Perfective. Attenuative verbs do not form derived Imperfectives.

4. **Distributive**. In this meaning, the prefix *po-* indicates that the verbal action is performed upon multiple (or all) objects, or that the action is performed by multiple (or all) subjects (Isačenko 1960:286-294; Zaliznjak & Šmelev 2000:124). In either case the object or subject is frequently modified by *vse* 'all'. Verbs in this meaning can be formed by adding *po-* to an unprefixed Imperfective stem (*pobrosat'*[p] 'throw (all or many)'), a prefixed derived Imperfective (*povyplačivat'*[p] 'pay out (all or many)'), or a prefixed Perfective (*pozaperet'*[p] 'lock (all or many)').

Pobrosal[p] *vse svoi magaziny, osobnjaki i jaxty, zajavilsja v kolledž.*
'He **threw away all** his stores, mansions and yachts, and applied to college.'

*Vse **porazezžalis'**[p], **porazexalis'**[p] po dačam.*
'Everyone **left** (by vehicle), **went their own way** to their dachas.'

Po-prefixed verbs in the distributive meaning are always Perfective. Distributive verbs do not form derived Imperfectives.

5. **Ingressive**. In this meaning *po-* focuses attention on the initial phase of the action, while implying that the action was likely completed. The ingressive meaning of *po-* deserves special consideration here. Isačenko (1960:224-230) defines the Ingressive Aktionsart as a focus on the commencement of an action, and of the three prefixes that

[11] <http://psichology.vuzlib.net/book_o738_page_16.html>

can form the Ingressive Aktionsart (*za-*, *vz-/vs-/vzo-*, and *po-*)[12], *po-* is the most constrained in the range of stems it can prefix. In Isačenko's view, *po-* can impart ingressive meaning to two unrelated groups of verbs: directed motion verbs such as *pojti*^{p} 'go, set off' and *poletet'*^{p} 'fly (off)' and a diverse array of other verbs, including *podut'*^{p} 'begin to blow', *poljubit'*^{p} 'come to love', *počuvstvovat'*^{p} 'begin to feel', and *poslyšat'sja*^{p} 'begin to be heard'. With directed motion verbs, the ingressive meaning is unstable and in some syntactic contexts can disappear: *Ona* ***pošla***^{p} *v kino* can mean 'She set out for the movies' or 'She went to the movies (and arrived there; she's no longer here)', depending on the speaker's intention.

The ingressives that do not involve directed verbs of motion are interesting for at least two reasons. First, we can divide these into two groups. The first group includes verbs that indicate motion of some sort, but are not considered "verbs of motion" in the narrow sense (that is, they do not have a bifurcated Imperfective consisting of unidirectional and multidirectional forms). This group includes verbs like *podut'*^{p} 'begin to blow' and *pomčat'sja*^{p} 'rush, speed off'. Since these verbs bear semantic similarity to the verbs of motion proper (that is, they all refer to motion of some sort), it is not difficult to see why the ingressive meaning would obtain with the prefix *po-* in these cases. The second group includes verbs of perception such as *poljubit'*^{p} 'come to love', *počuvstvovat'*^{p} 'begin to feel', and *poslyšat'sja*^{p} 'begin to be heard'. The only thing linking this second group to the other ingressive verbs is the "[*i*]*zvestnyj ottenok načinatel'nosti*" [certain nuance of commencement] (Isačenko 1960:231). While these verbs are traditionally considered ingressive, Isačenko seems non-committal about their

[12] The prefix *raz-* can also express Ingressive meaning (Zaliznjak & Šmelev 2000; Janda & Nesset forthcoming), but is omitted from Isačenko's (1960) catalog.

status – he says only that Zemskaja (1955) considers them ingressive. Since neither Isačenko nor Zemskaja give any other constraints for excluding verbs from this group, one may wonder why other verbs like *ponravit'sja*[p] 'like' and *podružit'sja*[p] 'become friends' are not considered ingressive as well. Like *poljubit'*[p] 'come to love', *ponravit'sja*[p] 'like' and *podružit'sja*[p] 'become friends' both refer to mental (or social) states that come into being at some point and remain so for some time afterwards. Indeed this is the stance that Zaliznjak & Šmelev (2000:110-111) take when they connect the ingressive meaning to verbs like *pokazat'sja*[p] 'seem, appear', *ponravit'sja*[p] 'like', *počudit'sja*[p] 'seem, appear', and *pomereščit'sja*[p] 'seem, appear'. However, they note that this last series of verbs often simply denote the action itself, without focus on the inception of the action. It is not traditional, however, to consider these last verbs as ingressive, and the reasons for categorizing them as ingressive are nebulously defined at best – one must be careful not to confuse the ingressive meaning with the perfect meaning of past tense Perfective verbs in Russian. For these reasons, I will only classify the following verbs as ingressive in this analysis, where appropriate:

- *po*-prefixed unidirectional verbs of motion: *pojti*[p] 'set off (on foot)'
- *po*-prefixed verbs expressing motion: *podut'*[p] '(begin to) blow', *pomčat'sja*[p] 'rush, speed off'
- verbs of perception/mental state, such as *poljubit'*[p] 'come to love', *počuvstvovat'*[p] 'begin to feel', *poslyšat'sja*[p] 'begin to be heard'

The others will be classified according to their more obvious meanings. *Po*-prefixed verbs in the ingressive meaning are always Perfective. Ingressive verbs do not form derived Imperfectives.

6. **Intermittent-attenuative**. In this meaning *po-* does not act alone, but is always accompanied by (a variant of) the suffix *-yva-*. Despite its productivity, some recent

treatments omit this meaning from semantic investigations of *po-* (see Dickey 2007). Unlike verbs of the other meanings, intermittent-attenuative verbs are always Imperfective and mean 'do X a little, with less-than-full intensity, from time to time' (Isačenko 1960:279-283, Zaliznjak & Šmelev 2000:122-124).

> *Tol'ko vot muž vse čašče* ***pogulival***^{i} *na storone, ne udeljaja žene vnimanie…*
> 'It's just that the husband **was cheating (a little, from time to time)** more and more often, not paying any attention to his wife….'

The verbal action in these cases usually refers to an atelic activity that could occur in multiple episodes – this includes multidirectional verbs of motion. However, this notion of iterativity (that the action is performed repeatedly in distinct episodes) is weak or absent in some verbs with the *po-…-yva-* combination (Isačenko 1960:280), and in some cases speakers seem to be re-construing the meaning of the prefixed verb altogether. Take this example from my data:

> *Vot u vas net takogo vpečatlenija (èto vopros ko vsem prisutstvujuščim) / čto sredstva massovoj informacii* ***očen' pobaivajutsja***^{i} *sejčas vlastej?*
> 'So you don't get the impression (this is a question to everyone present) / that the media **is very afraid** of the government now?'

The use of *očen'* 'very' along with the verb *pobaivat'sja*^{i} 'be afraid (a little)' seems to contradict the usual, intermittent-attenuative meaning of the verb. At this time I do not have sufficient data to say whether this constitutes a shift in usage, whether this usage is acceptable with a subset of verbs, or whether this is simply an idiosyncratic example.

While these may be the six generally-accepted meanings of *po-*, not all instances of *po-*prefixed verbs fall easily into these six categories – in some contexts two different readings seem equally plausible. In the data used for this dissertation, a number of these ambiguous cases turned up. Usually the ambiguity was between a resultative reading of

the *po*-prefixed verb and some other meaning: ingressive, distributive, or delimitative. A native speaker informant was consulted for help with ambiguous cases. In my data these cases were the exception, not the rule (71 out of the 1,000 randomly-selected instances of *po*-prefixed verbs could potentially be ambiguous). I will discuss how I dealt with these cases of ambiguity in Section 3.3.3. For now we will turn our attention to the cognitive linguistic paradigm and discuss what insights it gives into the semantics of *po*-.

2.3 The cognitive approach

Cognitive linguistics arose in the 1980s in part as a response to the questions left unanswered by the atomist and structuralist approaches, and in part as a response to new research on human cognition. During the decades to follow, the movement produced an international organization (the International Cognitive Linguistics Association) with national chapters in Belgium, China, Finland, France, Germany, Japan, Korea, North America, Poland, Russia, and Spain. The cognitive linguistic approach is at once a continuance of earlier trends in linguistic research and a departure from some traditional ideas. Though space will not permit me to summarize the entire movement here (see Janda 2006 & 2010a for a full discussion), I will discuss several overarching themes in cognitive linguistics that are pertinent to this dissertation: the nature of cognition and the role of embodiment (2.3.1), the nature of linguistic categories (2.3.2), and how linguistic knowledge can be organized via various cognitive structures, such as radial categories, idealized cognitive models, and image schemas (2.3.3). All of this information will help the reader to understand the proposed semantic analysis of the prefix *po*- given in 2.3.4.

2.3.1 Cognition and embodiment

Unlike most formal approaches to linguistics, cognitive linguistics attempts to explain linguistic phenomena in terms of general cognitive mechanisms. Many scholars (Chomsky, 1965; Fodor 1983; Pinker 1994; Carston 1996; Coltheart 1999; Flombaum et al. 2002; Hauser et al. 2002; Pulvermuller 2003; Pinker & Jackendoff 2005) have focused on the unique properties of language, often employing formal models and seeking out linguistic universals. Unfortunately space does not permit me to do justice to the generative linguistic framework here, and so I restrict myself to the very simplified observation that most generative approaches posit a number of specialized modules in the mind/brain, each of which handles linguistic information from various levels (phonology, semantics, syntax), and all of which are interactively linked by complex, rule-governed processing operations. Cognitive linguistics, however, does not focus on language as a separate cognitive entity – instead the cognitive framework appeals to general cognitive mechanisms in its quest to understand linguistic cognition. This stance is well-grounded in light of a wide range of evidence: Experiments done by Tanenhaus et al. (1995) indicate that in both reading and parsing spoken language, semantic, visual, and syntactic information are integrated to understand an utterance. In reviewing a host of neurocomputational and linguistic studies, Feldman (2006:7) notes that "[w]hat is technically called "aspect" in linguistics— the way we conceptualize the structure of events, reason about events, and express events in language— appears to stem from the neural structure of our system of motor control". Also, the interdependence of linguistic and non-linguistic cognition is underscored in persons with compromised cognitive ability – as Dąbrowska (2004) demonstrates, even in cases of aphasia and other language

deficits, linguistic and cognitive deficits go hand-in-hand (see also Lehečková's (2001, 2003) work on aphasia). With these and a host of similar findings in mind, the cognitive perspective begins with the assumption that the processes governing general human cognition also govern linguistic cognition. This assumption allows cognitive linguists to integrate advances in neurobiology and psycholinguistics as they investigate linguistic phenomena – with positive results.

Not only is linguistic cognition inextricably linked to the rest of cognition, but it is likewise informed by our experiences in the world – namely, our experience of interacting with the outside world through our bodies. Human bodies impose certain constraints on us (like having only two eyes, standing upright on two legs) that filter our experience with the world, and these constraints influence linguistic cognition in important ways. As Janda (2004) points out, human interactions with solid objects vs. fluid substances form the experiential base on which the Russian grammatical categories Perfective and Imperfective are founded. Being limited in space and time likewise influences the human perception of reality: Changing the time or the place from which one observes an event can alter a person's perception of that event. Again the Russian verbal categories Perfective and Imperfective illustrate the effects of perception and construal on the linguistic representation of events:

> ***Čitaja***^{i} ***knigu****, ja natknulsja na stroki o Zubre.*
> '**Reading the book**, I stumbled upon lines about Zubr.'
>
> ***Pročitav***^{p} ***knigu*** *Stendalja ob ital'janskoj živopisi, zainteresovalsja eju.*
> '**Having read** Stendhal's **book** on Italian painting, I got interested in it.'

The act of reading a book is largely the same, regardless of the size or type of book being read. The two examples above, however, use different aspects – Imperfective in the first,

Perfective in the second – to encode this activity, despite the essential similarity of reading the books in question (and the fact that both events occur in the past). The difference lies in the speaker's temporal perspective (in the midst of the action in the first example, looking back after completing the act in the second), and this embodied perspectival difference is reflected in the choice of Imperfective *čitat'* 'read' and Perfective *pročitat'* 'read (through)', respectively. Construal (which is influenced by perception) also has ramifications for the process of categorization, the next topic in our introduction to the cognitive linguistic framework.

2.3.2 Cognition and linguistic categorization

If we postulate that linguistic categorization is an instance of general categorization, and that our embodied experience influences how we categorize objects/events in the world, we would expect linguistic categorization to be affected by both. Indeed that is exactly what we encounter: In interacting with their environments, humans categorize objects, events, and even other humans in order to access knowledge about these objects/events/ other humans, make predictions, and choose behaviors appropriate to the interaction. And while many things we encounter on a daily basis fit into one category or another, other things defy rigid pigeon-holing into this or that group. Labov's (1973) experiments with the categories CUP and BOWL illustrate this nicely. Subjects were presented line drawings of cups and bowls of various shapes and sizes. Labov found that as the apparent width of the items depicted increased, there was no point at which all subjects agreed the picture represented a bowl instead of a cup — the boundaries of these categories are not clearly delineated. Set theory and Venn diagrams – both used in traditional approaches to

semantics and which rely on clear boundaries between members and non-members of a category – do not capture the "fuzzy" nature of real-world category boundaries. The same can be said regarding linguistic categories: The classes noun, adjective, verb, adverb, conjunction may indeed describe many lexemes/morphemes in a language, but some words (or word-parts) do not quite fit this scheme. Taylor (1995:176-180) presents an analysis of the categories WORD and AFFIX, showing that there are linguistic units within language, such as English *the*, that fall between these two categories. But by acknowledging the "fuzziness" of some category boundaries, I am not asserting that no categories have clear boundaries. Rather, I maintain that boundaries – lists of necessary and sufficient conditions –are not sufficient to explain the structure of many real-world categories. We need additional theoretical tools if we are to describe the nature of categorization more accurately.

How, then, are cognitive categories organized? For an answer we will turn to the notion of cognitive **prototypes**, which pervade all areas of human thought. In sum, prototype theory (Lakoff 1982, 1987) states that members of a category do not gain their membership in that category by possessing a certain set of necessary characteristics. Wittgenstein illustrates this notion in his *Philosophical Investigations* (1973) when he examines members of the category *SPIEL* 'game'. As Wittgenstein shows, a careful look at all games reveals that there is no one set of underlying characteristics that all members of the category share. For instance, while most games involve multiple players, solitaire is not excluded from being a game. Though most games have winners, ring-around-the-rosie is an example of a children's game without winners. Likewise chess lacks the usual criterion of amusement, and the format differences between baseball, poker, and chess

should suffice to show that finding a set of characteristics necessary for membership in the category GAME is futile.

Unlike the Aristotelian notion of certain, defining attributes being necessary for category membership, prototype theory places the burden of category membership on a given member's resemblance to the category's **prototypical member(s)**. Croft & Cruse (2004:74-105), Lakoff (1987), and Rosch (1973a, 1973b, 1978) discuss the structure of linguistic and perceptual categories (with their prototypical and non-prototypical members) in detail. For instance, ROBIN is a prototypical member of the category BIRD: adult robins have feathers and beaks, can fly, lay eggs, build nests, and possess the body shape that would be expected of a bird. Penguins, on the other hand, lack some of these characteristics: They cannot fly, do not build nests, and their feathers are different from those of a robin. Nevertheless, they are still members of the category BIRD — not because they possess a certain set of characteristics (some small dinosaurs had beaks, feathers, built nests, etc., but were certainly not birds in the usual sense), but because they possess some resemblance to the prototypical bird. While it is true that some categories have clear boundaries separating members from non-members (the categories DOG and CAT are not ambiguous, despite numerous similarities between the two), others do not. Bybee & Moder (1983) extend this method of categorization to grammatical categories in their study of ablaut and overall verb shapes in the strong verbs of English (*sing*, *sang*, *sung,* for example). On the basis of historical and experimental data, Bybee & Moder (1983) find that strong verbs in English are organized around a prototype of the shape /s/ + consonant cluster + /ɪ/ + /ŋ/ (the verb *string/strung* is a prototypical example). When participants in an experiment were asked to form the past tense of a series of nonce verbs,

the likelihood that the nonce verbs would be analyzed as strong (i.e. given a past tense like *sung* or *strung*) was directly proportional to the resemblance of the nonce verbs to the prototype. Sometimes it was the presence of a consonant cluster beginning in /s/ that determined the classification of the nonce verbs as strong, whereas other times the presence of a nasal or velar ending (both similar to the prototypical /ŋ/ ending) led participants to form past tense forms like *strung*. However, much like in Wittgenstein's explanation of GAME, no set of necessary-and-sufficient conditions could be isolated as the basis for membership in the category of strong verbs. Thus Bybee & Moder conclude that "speakers of natural language form categorizations of linguistic objects in the same way that they form categorizations of natural and cultural objects" (Bybee & Moder 1983:267). In other words, the method of classifying robins and penguins also applies to linguistic units.

As mentioned earlier, the similarity or dissimilarity of an item to the prototype will determine its inclusion in or exclusion from the category. Since category members will resemble the prototype to varying degrees, some members will be more prototypical, while others are peripheral members – in our example of the category BIRD, the robin is a prototypical member, while penguins are peripheral. In this manner the cognitive linguistic approach to category membership is able to provide internal structure to the category – recall that earlier approaches to prefixal meaning (2.1) were unable to describe any structure amongst catalogs of disjointed meanings (the atomist approach) or among the contextualized instances of the invariant (the structuralist approaches).

2.3.3 Imagining meaning: radial categories, ICMs, image schemas

Now that I have set the general stage for a cognitive analysis of *po-*, I will turn my attention to several additional concepts vital for that analysis: radial categories/networks, idealized cognitive models (ICMs), and image schemas. I will discuss each of these concepts in turn, followed by a few words on the notion of motivation (as opposed to prediction). It will then be possible to look at a detailed cognitive model of *po-* in Section 2.3.4.

Linguistic categories can be represented as **radial categories**, with a central, prototypical member connected to less prototypical/more peripheral members via various cognitive mechanisms, usually metonymy and metaphor (Lakoff 1987:91-114)[13]. A radial category is often depicted as a network of interconnected nodes, with each node representing a certain subcategory (or, in the case of polysemy, a certain meaning), and the lines connecting nodes represent different types of cognitive links between those subcategories/meanings. Take for instance the word *mother* in English. Typically a mother is the woman who gives birth to a child and subsequently nurtures it to adulthood and beyond. This meaning is given as primary in most dictionaries ("female parent" is the usual definition) and is the prototypical representative of the category MOTHER. However, there are other types of mothers who, although still encoded by the word *mother*, are more peripheral members of the category. For instance, in cases of adoption a birth mother is indeed a female parent (sharing half her DNA with the child) that does not nurture the child after birth. An adoptive mother shares no genetic relationship with her

[13] The details of category organization into radial networks are debated and developed differently by various authors, especially with regard to online construal (cf. Croft & Cruse 2004). Here I will be using Lakoff's (1987) model for the sake of simplicity, because its logic is basic to many similar treatments, and because it meets the needs of this analysis.

child, but assumes all the responsibilities associated with motherhood. A surrogate mother neither shares genetic material with the child nor does she provide any post-natal care, but is still considered a type of mother by virtue of carrying the child in her womb and subsequently giving birth.

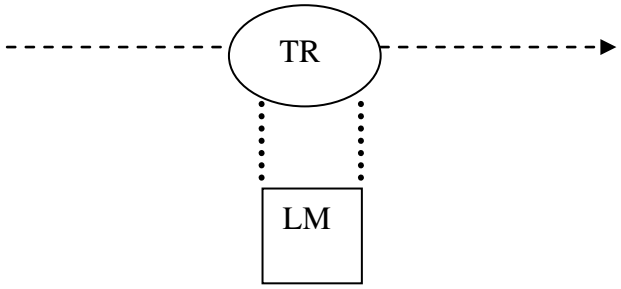
These three types of mothers – birth, adoptive, surrogate – are non-prototypical instances of the category mother. Each corresponds (with various degrees of fit) to the **idealized cognitive model (ICM)** (Lakoff 1987:68-90; Croft & Cruse 2004:28-32) of motherhood. In simple terms an ICM is a mental representation of a concept and all the background assumptions associated with that concept. An ICM thus represents not only a word's denotation, but also the word's connotations. Since ICMs are abstractions generated by multiple interactions with our environments, they do not always match reality perfectly. When a situation matches the ICM well – that is, most of the background assumptions are fulfilled – we can say that that instance is a prototypical case of the category under discussion. For instance, in Western culture the ICM of motherhood rests on several assumptions which include the following: The woman gives birth to the child. The woman and child share half their genes. The woman cares for the child at least until adulthood. Other less obvious assumptions are also part of this ICM: The mother lives with the child. The mother is a woman (cf. a recent case of a transgender person living as a man who gave birth to a child). The mother feels strong emotional attachment to the child. In the case of the prototypical mother, most if not all of these assumptions are true. In the non-prototypical cases, however, the reality of the situation does not correspond to the ICM in one or more ways: Birth mothers do not provide the post-natal care assumed by the ICM of motherhood. Adoptive mothers do not

share the assumed genetic link to their children. Surrogate mothers are perhaps the most peripheral case – they neither share genes with their children, nor do they rear them. ICMs explain (at least in part) the existence of prototype effects: Instances of a concept that match the ICM closely will be prototypical (the stereotypical mother), whereas instances that are only a partial match (birth, adoptive, surrogate mothers) are less prototypical. These "partial matches" are really metonymic relationships where each specific type of mother fulfills a part of the fully-specified ICM of MOTHER (a part-whole metonymy): birth mothers are genetically related to the child and give birth to it, adoptive mothers provide emotional support and care for the child, and surrogate mothers at least carry the child during gestation. In addition to these metonymic extensions of the category MOTHER, metaphoric extensions are also possible, as in the old adage *Necessity is the mother of invention*. Here there is a metaphoric mapping from the domain of human relations (*mother*) to the domain of mental creativity (*invention*) – just as a woman is the origin of the child, spurs the growth of the fetus, and ultimately "produces" the child, so is *necessity* the point of origin and facilitator of the thought and experimentation that ultimately produces *invention.* Thus we have a group of mothers related by family resemblances – attributes are shared by connected members in the category, but there is no one attribute that all members of the category hold in common.

The senses of a polysemous word can be similarly represented as a radial network, with one or more senses that are central/prototypical and other, peripheral senses connected to the prototypical cases by cognitive links (usually metaphor and metonymy). Whereas we have already seen how the various types of mothers all relate to the category MOTHER, other types of cognitive operations link the meanings of

polysemous terms like *po-*. But before we discuss those links, first we must understand how the meanings of verbal prefixes are represented in a cognitive analysis, that is, by image schemas. In short, an **image schema** (Johnson 1987, Lakoff 1987) is an abstract pattern in cognition that represents recurring relationships in our embodied experience of the world. Image schemas come in various kinds, but the type that is most pertinent to the study of *po-* usually involves a **trajector (TR)** that stands in relation to a **landmark (LM)**. The LM provides a point of reference for locating the TR. The usefulness of these image schemas can perhaps best be seen in the well-known example of the English preposition *over*[14]. We will start by examining only one of the many senses of *over,* namely the 'above-across' sense. We can represent the meaning of *over* image-schematically:

Fig. 2.3.3-1



In this image schema, the TR traces a path (indicated by the dashed arrow) above some LM. The dotted lines extending between TR and LM represent the extreme boundaries indicated by the LM. Although the drawing implies no contact between TR and LM, contact is possible; the drawing should be understood schematically – i.e. without commitment to contact or non-contact between TR and LM. This schema captures the

[14] Here I am drawing on Brugman's (1981) and Lindner's (1981) accounts, as adapted by Lakoff (1987). I am using an English-language example to establish the basic concepts in the mind of the reader before proceeding to an actual image-schematic analysis of *po-* in Section 2.3.4.

meaning of *over* in sentences like *The plane flew over the building*, in which the plane is the TR, the building is the LM, and *over* describes the relationship between the two. In some cases the LM is only implied, as in *The plane flew over*. Changing the manner of locomotion does not change the schema: The sentences *Jane drove over the fallen branch* and *John ran over the hill* still represent the same relationship between TR (Jane and John) and LM (fallen branch and hill).

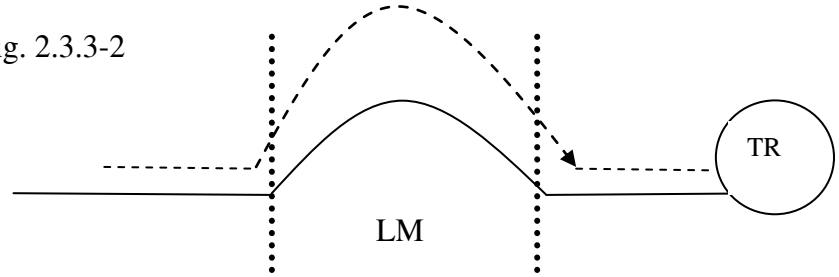
The different senses of *over* are the result of metaphoric and metonymic extension from the prototypical image schema in Fig. 2.3.3-1. **Metaphor** allows the extension of a meaning from one domain into another. A semantic **domain** includes any knowledge and conceptual structure relating to some aspect of our embodied experience in the world (Lakoff 1987) and serves as background for our understanding of specific concepts (also known as the **profile/base** relation; see Croft & Cruse 2004:15-16). A metaphor juxtaposes two semantically unrelated concepts and allows us to understand one in terms of the other by "mapping" the semantic content of the **source domain** into the **target domain**. For instance, the meaning of *over* represented in Fig. 2.3.3-1 relates to physical motion or physical location. However, we can metaphorically map our understanding of the physical world onto our emotional experience (and thus gain new insights into that experience) via the metaphors OBSTACLES ARE VERTICAL OBJECTS and LIFE IS A JOURNEY. This mapping motivates the use of *over* in sentences like *Josh needs to get over his divorce*. The divorce is understood as a vertical, physical object, above and across which Josh must travel on his journey of life. This meaning of *over* can be represented much the same as the meaning in *The plane flew over the building*, except that the domain is no longer the physical, but the psycho-social realm.

Like metaphor, **metonymy** is another process by which meanings can be derived. Whereas metaphor relies on perceived or imaginatively created similarity between two entities, metonymy relies on association or (cognitive) contiguity between two entities that belong to the same domain or ICM (Kövecses 2002:145; Peirsman & Geeraerts 2006). And whereas metaphor is most often concerned with understanding, metonymy is often about reference. Lakoff & Johnson (1980) list several common metonymic relations in which one entity stands for or is the conceptual source for another: PART for WHOLE (***Rednecks*** *aren't welcome here*), PRODUCER for PRODUCT (*I just bought a new* ***Louis Vuitton***), CONTAINER for CONTENTS (*That* ***carton*** [of eggs, for example] *is spoiled, throw it away*). Metonymy is sometimes involved in diachronic semantic shift – Geeraerts (1997:68-79, cited Dickey 2007:19) notes that Dutch *winkel* originally meant 'corner' and was used to refer to the street corner where shops were often located. Using the metonymy whereby LOCATION stands for the THING LOCATED, speakers referred to the shop at the street corner as *winkel,* and with time 'shop' became the new meaning of *winkel,* as it is in Dutch today. Though a number of relations can be posited among the many metonymies at work in language (see Peirsman & Geeraerts 2006 for discussion), in this dissertation we will make use of the generalization that "[m]etonymy is present when one item (a VEHICLE) is used to access another item (a TARGET)" (Janda forthcoming; see also Kövecses 2002).

Returning to the example of *over,* metonymy is responsible for several extensions of the prototypical meaning. In some senses the focus on motion (as in *The plane flew over the building*) is metonymically shifted to the endpoint location (MOTION STANDS FOR

DESTINATION), yielding the meaning found in *Sam lives over the hill*. We can depict this meaning graphically:

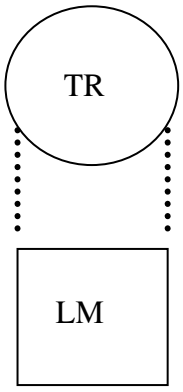
Fig. 2.3.3-2



The TR, Sam, lives at some location beyond the hill which can only be reached by traveling a path over (above and across) the hill (the LM). Although there is no actual motion denoted in the sentence, a path is still strongly implied.

Other senses lack the notion of a path altogether. Note the variant of the 'above' meaning of *over* in this sentence: *The painting is hanging over the fireplace*.

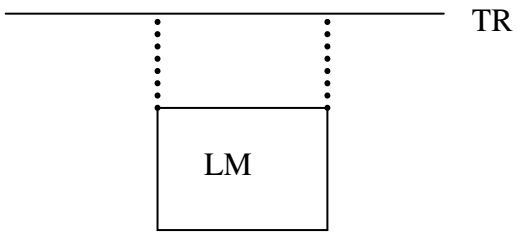
Fig. 2.3.3-3



The painting (TR) is simply located above the fireplace (LM), and no movement or path is implied. This meaning shares the 'above' relation expressed in Fig. 2.3.3-1 (*The plane flew over the building*), but lacks a path for the TR – again, this is a metonymic relationship whereby only parts of the prototypical TR-LM configuration in Fig. 2.3.3-1

are selected as relevant. By an additional modification – changing the shape of the TR – we arrive at the meaning of *over* in sentences like *The power line stretches over the yard*:

Fig. 2.3.3-4



The power line is represented by a cognitively one-dimensional TR that, although it does not move, extends beyond the outer boundaries of the LM (the yard). Just as the schema in Fig. 2.3.3-4 is based on the schema in Fig. 2.3.3-3, additional meanings of *over* can be derived in chain-like fashion.

As the reader will see in 2.3.4 and again in 4.4.2, metaphor and metonymy play a role in the semantics of *po-* as well. For instance, the attenuative meaning is related to the delimitative by a metaphoric transfer from the domain of TIME to INTENSITY. The ingressive meaning is a metonymic extension of the resultative, where only the initial phase of the path expressed by the verb is highlighted as relevant, and the completion of the action is backgrounded. Furthermore, Zaliznjak & Šmelev (2000:109) claim that the frequent ambiguity between ingressive or resultative readings of directed motion verbs (such as *pojti*^{p} 'set out, go (on foot)') is due to this metonymic relationship, whereby the inception of the action (which is usually expressed in ingressive verbs) stands for the entire course of the action, bringing us closer to a resultative reading of the same verb (cf. Dickey 2007:37). Note that these metonymic relationships are not arbitrary – rather, they have logical motivations in our experience of the world.

This brings us to the final concept that we will review in our brief tour of cognitive linguistic theory: **motivation** (Lakoff 1987:96). Much of scientific research concerns itself not only with an explanation of facts, but also with predictions about facts yet to be uncovered. Indeed, the strength of a theory is usually measured by its predictive power. Language, however, is not entirely predictable, as its development and employment depend on somewhat unpredictable actors – humans. If language were completely predictable, we would be hard pressed to explain dialectal variation and diachronic changes that affect some speech communities differently from others. This does not mean, however, that since linguistic systems are not entirely predictable, they are completely arbitrary. Motivation provides a middle option between these two extremes. A motivated account of linguistic phenomena (whether semantic, syntactic, phonetic, etc.) makes sense of the system, pointing out which domains of experience or construals of experience are pertinent to that system. For instance, the image-schematic account of *over* explains what the meanings of *over* are, how the meanings of *over* are related, and what aspects of experience are relevant to the semantics of the preposition. We would expect there to be logical, psychologically plausible ways of relating the various meanings of *over* if the analysis is to be at all tenable. The model of *over* does not and cannot predict exactly how the meaning of the preposition will change with time, how those meanings may differ in yet un-studied dialectal contexts, and it cannot make precise, retrograde "predictions" about the historical semantic development of *over* – we might make informed guesses based on the model, but these guesses are not predictions in the strict sense that, if they are scientifically (in)validated, they offer incontrovertible "proof" of the model's correctness . In this fashion this dissertation seeks to motivate the

semantic structure of *po-* in a psychologically plausible way that takes into account all the available factual information, without making scientifically verifiable predictions about a phenomenon that is only partially predictable. That is not to say that because I am not making scientific predictions in the technical sense, my analysis becomes a theoretical free-for-all. As the reader will see in Sections 2.5, 2.4, and throughout Chapter 3, statistical analysis (a verifiable, quantitative approach) is the basis of my analysis of *po-*. Given the partially unpredictable (and yet non-arbitrary) nature of language, a motivated account of *po-* grounded in corpus data is a reasonable goal.

2.3.4 Cognitive analysis of prefixal semantics

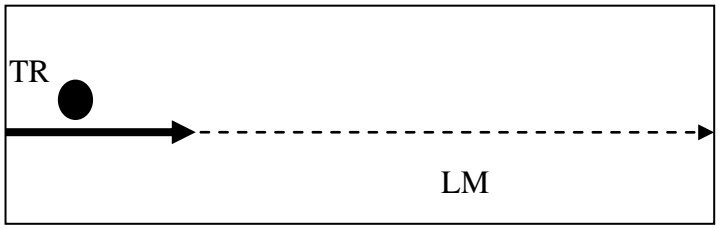
It should be noted that cognitive linguistic accounts of prefixal meaning are in a sense picking up where the atomist and structuralist approaches left off: The cognitive accounts seek both to uncover the basic meanings of a prefix and to find unity in the diversity of those meanings, albeit by relying on a different set of theoretical tools. Detailed analyses of several Slavic prefixes have been proposed: *vy-* in Polish (Rudzka-Ostyn 1983); *za-, pere-, ot-*, and *do-* in Russian (Janda 1985, 1986); *pere-* and *nad-* in Bulgarian (Tchizmarova 2005, 2006); select meanings of several Czech and Russian prefixes, including *po-* (Shull 2003); and from a diachronic perspective, *po-* in Russian (Dickey 2007). I now turn my attention to this last analysis of *po-* (Dickey 2007).

Dickey's (2007) work on *po-* is a diachronic analysis that aims to explain a suspected shift in prototype that occurred over the course of several centuries. Although my analysis in Chapter 4 is entirely synchronic, Dickey's work is useful inasmuch as it is the only cognitive treatment of *po-* that attempts to model the semantics of the prefix in

its entirety. Dickey (2006) does not explicitly present an image-schematic interpretation of each meaning of *po-* nor does he present a graphical representation of *po-*'s semantic network. The analysis that I present not only recounts but expands many of Dickey's (2007) ideas, and I also address one meaning omitted by Dickey (2007). My interpretation of Dickey's proposed network serves two purposes: 1) Since this is the only network account of *po-* available, it forms a useful point of comparison for my statistical analysis in Chapter 4, and 2) it shows the reader how the principles of Section 2.3.3 apply to *po-*. I treat my agreements and disagreements with Dickey's model in Section 4.4. For a complete description of how the meanings relate to one another, see 4.4.2.

Drawing on work by Shull (2003), Dickey (2007:18) asserts that most meanings of *po-* stem from an original PATH/SURFACE CONTACT meaning in Old Russian. Unlike other Russian verbal prefixes, however, *po-* has since lost its spatial meaning (Camus 1998:101, Guiraud-Weber 1993:57, Tixonov 1998:101). The remaining six meanings have to do with the domain of TIME (and INTENSITY in some cases) – the original spatial sense of *po-* was metaphorically re-construed via the TIME IS SPACE metaphor, allowing the expansion of *po-* into new semantic territory. The ingressive meaning stems from a meaning of PATH/PARTIAL-TRAJECTORY, in which both the path and the trajectory are understood metaphorically as aspects of an action's progression through time. We could represent the ingressive meaning image-schematically like this:

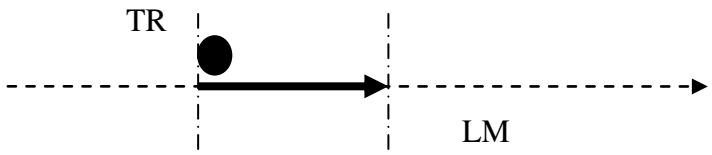
Fig. 2.3.4-1



Here the TR is the subject of the *po*-prefixed verb, and the LM is the canonical course of action implied by the base verb (Shull 2003:152-153). The dashed horizontal line represents the full trajectory of the action encoded by the base verb, and the darkened arrow represents some subset of that trajectory – here *po-* only stipulates that some portion of the trajectory has been covered, and is non-committal regarding whether the TR reaches the end of the trajectory. This schematic "lack of commitment" to covering the full possible trajectory accounts for the ambiguity of sentences like *Ivan pošel*^{p} *v bar*: The sentence can be interpreted as *Ivan set out for the bar*, indicating that the speaker knows he has left and, barring unforeseen circumstances, will probably arrive there as planned. Or the sentence can be interpreted *Ivan went to the bar*, implying that he reached his destination – this is a case of metonymy, where the INITIAL SUBEVENT (setting out for the bar) STANDS FOR THE COMPLEX EVENT (setting out for, traveling to, and arriving at the bar). Context usually disambiguates between the two interpretations. As a reminder, image schemas are schematic: They do not represent an aspect of our experience point for point, line for line, but can represent generalizations and incomplete "pictures" where some facets of reality are judged significant (and are thus clearly represented), whereas other aspects of reality are selectively omitted. The two-dimensional drawings I give here are only representational; I do not claim that the mind stores exact copies of the pictures I give here as real cognitive structures.

Dickey (2007) suggests that the delimitative meaning can be defined as RELATIVE DELIMITATION. It closely resembles the ingressive meaning (PATH/PARTIAL-TRAJECTORY), except that the path/trajectory has no inherent endpoints. As such, the prefix can only delimit a certain portion of the action's potentially infinite timeline.

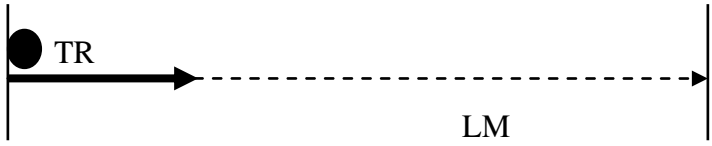
Fig. 2.3.4-2



Here again the TR is the subject of the verb, and the LM is the canonical course of action implied by the base verb. The path here is potentially limitless, since this meaning of *po-* only occurs with activity verbs (Vendler 1957) – verbs that encode atelic actions (having no natural endpoint). *Po-* delimits a portion of the action relative to that trajectory.

The attenuative meaning can be schematicized similarly. With the attenuative, however, there has been a metaphoric transfer from the source domain TIME (which is relevant in the delimitative meaning) to the target domain of INTENSITY. Just as the delimitative signals an action that occurs over a small portion of the potential duration of the event encoded by the verb, so the attenuative signals that the action occurs at only a fraction of the potential intensity associated with that verb. The prefix no longer modifies a course through time, but instead it modulates an imagined scale of intensity.

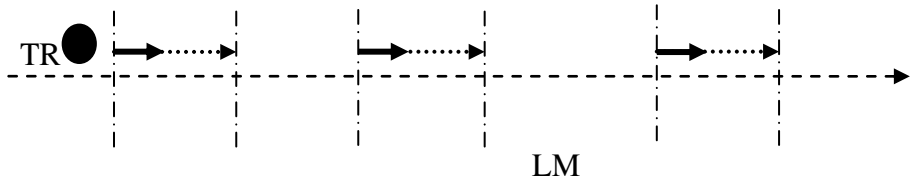
Fig. 2.3.4-3



What is relevant is that only a portion of the intensity scale is realized during the performance of the action, yielding the frequent translation 'do X slightly/a little/incrementally'.

The intermittent-attenuative meaning of *po-…-yva-* ('do X slightly, from time to time) resembles both the attenuative and to some extent the delimitative. Though Dickey (2007) does not examine this meaning, we can represent it image-schematically thus:

Fig. 2.3.4-4

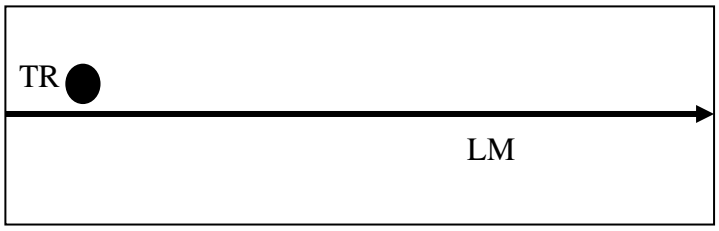


In short, the delimitative and/or attenuative meanings are reduplicated along a temporal trajectory – an action is performed at less-than-full intensity, at irregularly spaced intervals along a timeline. The relationship between the delimitative/attenuative and the intermittent-attenuative is an example of yet another metonymy where a SINGLE EVENT IS THE SOURCE FOR A COLLECTION OF SIMILAR EVENTS (see Kövecses 2002; Peirsman & Geeraerts 2006). Once again the TR is the subject of the verb and the LM is the canonical course of action represented by the base verb. *Po*-prefixed verbs of the intermittent-attenuative meaning are typically atelic, so the metaphoric trajectory here is unbounded. Because the action encoded by the verb has no natural terminus, any period of performance is thus relatively delimited.

At least in Dickey's view, the resultative meaning is perhaps most connected to the ingressive, which ultimately paved the way for the delimitative to replace the

resultative as the prototype in the semantic network of *po-*[15]. The relationship between the resultative and ingressive is a case of PART-WHOLE metonymy: the resultative meaning implies that the action has been performed from beginning to end, while the ingressive focuses on the initial portion of an action. Here is an image-schematic representation of the resultative:

Fig. 2.3.4-5



Once again, the TR is the subject of the verb, the LM represents the canonical course of action encoded by the base verb, and the path marks a complete trajectory from inception of the action to termination, beyond which the action cannot naturally proceed.

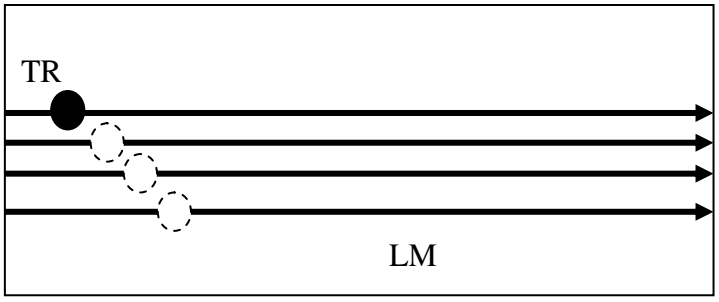
For the final meaning of *po-*, the distributive, I will depart from Dickey's (2007) interpretation, as it appears somewhat tenuous[16] (Dickey 2007:25), and instead I will present an image-schematic interpretation that preserves the logic of more traditional approaches (cf. Isačenko 1960:288). The distributive meaning is much like the resultative, only multiplied:

---

[15] I will offer my own take on the centrality and relatedness of the meanings of *po-* in Chapter 4. In this chapter, however, Dickey's (2007) analysis meets our illustrative purposes.

[16] After explaining how the distributive fits into his semantic model of *po-*, Dickey (2007:25) comments "Inasmuch as this solution seems far-fetched […]", hinting that he anticipates some doubt on the part of (at least some of) his readers. My aim here is not to disprove or discredit Dickey's proposed meaning of the distributive. Instead I favor the solution I offer here because it is simpler, is more easily motivated by the model, and bears a more direct relationship to the other members of the network.

Fig. 2.3.4-6



Like the intermittent-attenuative, the distributive is another instance of the metonymy A SINGLE EVENT IS THE SOURCE FOR A COLLECTION OF SIMILAR EVENTS – whereas *po-* signals the completion of a single event in the resultative meaning, its meaning is extended to indicate the completion of a group of similar events in the distributive meaning. The TR is the subject of the verb, and the LM is the canonical course of action encoded by the base verb. The TR can be plural or singular; the schema is non-committal regarding the subject's number. What is significant is that the action is performed multiple times (whether by multiple subjects or upon multiple objects), and that the TR completes the possible trajectory from beginning to end.

As the reader may have noted, the image schemas for *po-* are relatively simple by comparison to image schemas for other Slavic verbal prefixes (cf. Rudzka-Ostyn 1983a, Janda 1986, Tchizmarova 2006). This is due largely to the fact that *po-* has lost its spatial meanings, and as a result the LM becomes the action encoded by the verb itself, not an external, physical landmark (Shull 2003:147-180, Dickey 2006:14). It is for this reason that the resultative meaning of *po-* has been often equated with "simple perfectivity" (Guiraud-Weber 1993).

Having thus represented the six recognized meanings of *po-* as image-schemas, I have yet to address at least two important issues. The first involves the internal structure

of *po-* as a semantic category. While Dickey (2007) proposes a radial network with the delimitative and ingressive meanings at the center, this is not the only plausible arrangement. Secondly, the links between meanings that I present are based primarily on introspective judgments; while suggesting general links between senses, Dickey (2007) leaves the specifics of his inter-meaning connections undefined. I believe that both of these issues can be addressed, however, with recourse to a representative linguistic sample, i.e. corpus data. I introduce the concepts of corpus linguistics and the analyses corpora make possible in Section 2.4. I then describe an objective, statistically sound method of using corpus data to investigate the semantic network of *po-* in Chapter 3. And finally in Chapter 4 I use the empirical results of that method to determine the structure of *po-*'s semantic network, determining the prototypical meaning and the links among meanings.

2.4 Corpus linguistics – basic concepts and assumptions

In this section I introduce the basic concepts of corpus linguistics. First I define what a corpus is, and I discuss the notions of representativeness, searchability, and annotation. I then present two assumptions underlying the use of corpora that are particularly relevant to this dissertation, and outline a few issues that must be kept in mind when using corpus data to study linguistic phenomena. This basic introduction to corpus linguistic concepts sets the stage for the analysis presented in Chapters 3 and 4.

Corpus linguistics is not a theoretical framework for linguistic analysis. While it does make some theoretical assumptions from the outset, in reality corpus linguistics is a methodology – a means of arriving at linguistic knowledge that relies on **corpora**. A

corpus is a body of linguistic data that is collected in such a way that it is as representative as possible of its target language or area of linguistic inquiry. For instance, a corpus of spoken American English can consist of either audio recordings of American English speakers, or transcriptions of utterances produced by American English speakers. Corpora that aim to represent a language of a given time period (say, modern British English) draw from a wide variety of genres and formats, ranging from high literature to newspaper articles, from poetry to internet blogs. Depending on their purpose, corpora vary widely in size, though most that are of any scholarly use range from several million to well over a hundred million words.

There are three aspects of corpora that make them powerful tools in linguistic investigation (McEnery & Wilson 2001): representativeness, searchability, and annotation of linguistic information. As mentioned earlier, the documents included in a given corpus are carefully chosen so that the corpus is as representative of the target language as possible. In so doing, it is important to accurately define the boundaries of the population under study: Instead of having a corpus of "Brazilian Portuguese", a corpus may rather seek to represent "written Brazilian Portuguese from 1925 to 1975". The compilers of such a corpus would then employ the same random sampling techniques used in the sciences to gather a random sample of Brazilian Portuguese works from that time period, perhaps using a national bibliography or annual press guides. When collecting speech for inclusion in a corpus, demographic sampling techniques akin to those used in public opinion polls are used. Conclusions drawn from precisely-defined corpora may or may not be broadly applicable to the language, depending on both the nature of the corpus and the linguistic object of study – Biber (1993) found that common

items are so evenly distributed in linguistic production that even small corpora suffice for their study, whereas rarer items require larger, more carefully structured (and thus representative) corpora if the researcher is to get any realistic data on their distribution. However, for larger general-purpose corpora, representativeness – if it is achievable at all – is a far more difficult task. The creation of a corpus representative of a *language* is fraught with theoretical and pragmatic questions for which there are currently no answers (Kilgarriff & Grefenstette 2003). Therefore most general-purpose corpora seek the more modest (and attainable) goal of *balance*, meaning that they include texts from the wide array of formats, styles, genres, and registers available in a language. Yet despite the incomplete nature of corpora, they offer an empirical basis for semantic studies, allowing researcher bias to be minimized much more than with introspective analyses.

The remaining two features of corpora – searchability and annotation of linguistic information – go hand-in-hand. Obviously, if a corpus is to be of any use, the researcher must be able to search the corpus for the linguistic element under study. Most corpora today are available on CDROM or, more commonly, through the Internet – a far cry from the first machine-readable corpus, Father Robert Busa's collection of sentences from the works of St. Thomas Aquinas that was housed entirely on punched cards (readable only by certain IBM computers in the 1950s). But it is not enough to simply have searchable text if all the linguistic data remains implicit – that is, if a researcher wanted to study strong versus weak verbs in English, how will s/he find the verbs in a corpus? The corpus must be **annotated** – each element within the corpus must be coded with relevant information, such as part of speech, semantic class, tense, case, number, syntactic role, etc., which makes searching the corpus along linguistic parameters possible. For instance,

in the sentence *John ran all the way home*, the word *ran* could be annotated as such: verb, past tense, 3rd person, singular, indicative mood, intransitive. Each annotation (or **tag**) is one possible **value** or **level** of an **ID tag**. For example, the tag "verb" is a value/level of the ID tag "part of speech", and the tag "past tense" is a value/level of the ID tag "tense". Each ID tag has multiple possible levels – the ID tag "tense" can be "past", "present", or "future". ID tags are foundational to corpus-based research, and the ID tags used in any study warrant careful discussion; I explain the ID tags employed in this dissertation in Section 3.3.3. Unfortunately, while some annotation (particularly of English corpora) has been automated, much must still be accomplished by hand.

At this point some assumptions underlying corpus linguistic methods become apparent. First, the use of corpora in research assumes that natural language production is a reliable indicator of the actual structure in a language. Thus contra Chomsky (1964), the object of linguistic study is not **competence** (a speaker's internalized linguistic knowledge) but rather **performance** (actual language production). Space does not permit me here to examine all of Chomsky's criticisms of corpus linguistic practice during the period up until and during the 1960s, so I will briefly summarize the position adopted in this dissertation: The ideas that Chomsky found most problematic are no longer accepted in corpus linguistic work – no one believes anymore, whether explicitly or implicitly, that language is finite, that all the sentences of a language are enumerable, or that corpus data alone suffices to explain language without reference to the inner workings of the mind. Instead, present-day corpus linguistics acknowledges the imperfect nature of corpus data, while simultaneously not dismissing it out of hand as an unreliable reflection of linguistic structure (see Labov 1969; McEnery & Wilson 2001). Furthermore, most large corpora

today take great pains to be as representative as possible of the population they claim to represent; the folly of making broad assumptions about English from a corpus comprised solely of excerpts from *Jane Eyre* is quite apparent to everyone. Chomsky's criticisms were influential in the shaping of modern-day corpus linguistics, and without them the field would not have progressed to its present state. His criticisms, however, were specific to a certain period in the development of corpus linguistics, and as such they do not negate the conclusions of balanced, methodologically sound corpus-based studies in the twenty-first century.

Secondly, and perhaps most importantly for this dissertation, corpus-based semantic studies rest on the idea that "distributional similarity reflects, or is indicative of, functional similarity" (Gries & Divjak 2008; see also Divjak 2010). In simpler terms, a word's (or prefix's, as the case may be) syntactic, pragmatic, and discourse behavior is intimately tied to its meaning, so much so that a given meaning of a word will have its own pattern of co-occurrences that distinguishes that sense from the other senses of the same word. These co-occurrences can be frequent lexical combinations, such as *different from* versus *different than* or *different to* – here the word following *different* does not differentiate between meanings of *different* but is rather a marker of dialect or region of origin. Such lexical co-occurrences are called **collocations**. Another type of combination of interest is the **colligation**, the co-occurrence of a word with specific grammatical phenomena. Colligations can also differentiate between word meanings, as is the case with the English verb *run*: In the concrete sense, *run* is an intransitive verb meaning 'move quickly on foot'. However, in the metaphorical sense 'execute, utilize (a computer program)', the verb *run* is transitive and takes an inanimate direct object: *My computer*

runs *Windows XP. He needs to* run *his debugging software*. Atkins' (1978) work on the word *danger* shows that an analysis of collocations can disambiguate the sense of *danger* in a given sentence, and that such an analysis can also provide an objective, non-introspective basis for determining how many senses *danger* possesses. In his work on *urge,* Hanks (1996:77) goes so far as to say that "the semantics of a verb are determined by the totality of its complementation [i.e. collocational and colligational] patterns", and McDonald (1997) verifies experimentally that collocational patterns are a reliable indicator of semantic similarity. A corpus, then, provides us with a source of collocational and colligational patterns. By examining many instances of a given word (or prefix, as in this dissertation) in context, the researcher can quantify distributional frequencies that, when subjected to statistical procedures, reveal much about a word's semantic structure. Section 3.5 details the statistical procedures used in this dissertation and Sections 4.2 – 4.3 discuss what those procedures reveal about the semantics of *po-*.

2.5 Tying it all together: Cognitive linguistics, corpus linguistics, and why this dissertation is important

In this section I connect three of the major themes discussed so far (the meanings of *po-*, cognitive linguistic analysis, and corpus linguistic analysis), and I further discuss the theoretical basis for my study on *po-*. I explain how (and why) I join ideas and methodologies from cognitive and corpus linguistics in this study, and how this relates to current trends in cognitive/corpus linguistic research. I re-iterate my specific goals for this dissertation and set the stage for the methodological discussion in Chapter 3.

Prior to widespread public access to (and acceptance of) large corpora, linguists had no choice but to rely on intuitive (and at times subjective) judgments when

investigating the semantic structure of polysemous words. And once corpora did become available, these corpora were often restricted in both size and number of sources (see Guiraud-Weber 1993[17], Veyrenc 1980:159-179 for examples) – too much so to be indicative of the language as a whole. The situation today is much different, and the availability of robust corpora (and the software to make use of them) makes more empirically verifiable analyses of polysemy possible. No aspersion is cast here on prior studies that did not use robust corpora; rather, this dissertation builds on those previous works and exploits recent advances in technology and linguistic thought to address an old, unresolved issue – namely, the relationship(s) among the meanings of *po-*.

Not surprisingly, recent trends in cognitive linguistics reflect an increasing interest in and adaptation of corpus methods for cognitive analyses, particularly with regard to questions of polysemy. Usually cognitive analyses of polysemous words aim to accomplish three goals: distinguish the prototypical sense(s) of a word, motivate relationships between the prototypical sense and other related word senses, and determine when a sense is sufficiently distinct from related senses to be called a separate meaning in its own right. Given these aims, three questions naturally follow: How is the prototypical sense determined? How are the relationships among meanings motivated (aside from relying on intuitive and subjective judgments of similarity)? On what grounds can one establish the existence of separate senses? A number of approaches to these questions have been proposed, sometimes leading to contradictory conclusions. As has been noted (Sandra & Rice 1995; Gries & Divjak 2008; see Gibbs & Perlman 2006 for similar

[17] Although Guiraud-Weber (1993) makes use of a 700-page corpus in her study on *po-*, nothing of the composition of the corpus is revealed, save that it included no spoken language. As a result, Guiraud-Weber's frequencies, while interesting, must be used with caution.

discussion), many of these approaches suffer from a lack of empirically verifiable methods. Corpus data, however, can provide the needed basis for an empirically sound study. By marrying the theoretical tools of cognitive linguistics with the statistical tools used in corpus analysis, we have a methodologically rigorous means of evaluating the hypotheses generated by an approach to polysemy grounded in our knowledge of human cognition. This dissertation adopts just such a cognitive-corpus linguistic approach, and I maintain that this approach yields important insights into the semantic structure of *po-*. Some of these insights would not be possible without the statistical analysis (and software for analysis) of corpus data. While the human mind is unparalleled in its creativity, computers are far better equipped to find patterns among thousands and thousands of data points. And because I rely on empirical methods to validate my hypotheses, this study is both falsifiable and repeatable by other researchers – one of the basic requirements of any scientific work.

Specifically, I will use corpus data to build a **behavioral profile** (Hanks 1996, Divjak & Gries 2006) of *po-*. Simply put, a behavioral profile reveals how frequently *po-* occurs with (or is *distributed* with) other bits of grammatical, syntactic, and semantic information (i.e. the collocations and colligations associated with *po-*). In other words, a behavioral profile allows us to see distributional characteristics of the senses of *po-* in a numerical format. Having extracted this kind of quantitative data from a corpus, we can answer the most important questions arising from a cognitive semantic analysis of *po-*: Which meaning is prototypical? What is the relationship among the meanings of *po-*? If we conceive of the meanings of *po-* as a radial network, what does that network look like? The answers to these questions can be obtained with the help of statistical analysis

of the data (hierarchical agglomerative cluster analysis and other investigative techniques; see Section 3.5), meaning that the conclusions of this study do not depend on subjective, introspective judgments of linguistic data. Rather, the results are empirically verifiable and falsifiable; other researchers could conduct similar studies using the same methodology and either support or refute my conclusions (cf. Leech 1992). For a more thorough explanation of the behavioral profile concept, see Sections 3.1 and 3.4, where I walk through the steps of creating and analyzing a behavioral profile of *po-*.

In sum, this dissertation takes advantage of recent ideas in linguistic research – namely, the combination of corpus methods with cognitive linguistic principles – to tackle a very old problem: how to structure the relationships among the disparate meanings of *po-*. As such my work here is part of a larger trend towards empirical, bottom-up approaches to linguistic analysis, and I believe this approach provides a satisfactory explanation of *po-*'s semantics. To my knowledge, this is the first work to use the behavioral profile concept to describe the polysemy not of a word, but of a prefix. In so doing I show that behavioral profile analysis can fruitfully be applied to a wider range of linguistic phenomena than it has been in the past.

3 Methodology

3.1 Introduction

In this dissertation I create a **behavioral profile** (Hanks 1996, Gries & Divjak 2008; see also Stefanowitsch & Gries 2003 and Janda & Solovyev 2009 for similar approaches) for the prefix *po-*, and I use that behavioral profile as the basis for statistical analysis. In simple terms, a behavioral profile shows the collocations and colligations associated with a certain linguistic unit in a corpus, and how frequently those collocations and colligations occur. A behavioral profile is usually formatted as a table – an excerpt from the behavioral profile of *po-* is given in Table 3.1-1, which shows just one of many ID tags (see 3.2, Step 2) and how its levels are distributed across the meanings of *po-* in my data. Note that the attenuative meaning is not listed in Table 3.1-1 because it was not attested in the randomly-selected data used in this study. See Section 3.3.2 for more information on how this table is organized:

Table 3.1-1

ID tag	ID tag level	delimitative	distributive	ingressive	intermittent-attenuative	resultative	Total
sentence	declarative	104	10	135	24	559	832
type	exclamation	0	1	0	0	10	11
	imperative	5	0	6	0	29	40
	interrogative	4	0	10	2	30	46
Total		113	11	151	26	628	929

Behavioral profiles are most often employed in investigating polysemy and synonymy (Atkins 1978; Hanks 1996; Gries 2006; Gries & Divjak 2008; Berez & Gries 2009), as they provide researchers an empirical basis for distinguishing between word senses or related words – the ability to compare how frequently certain collocations/colligations occur with the various items under investigation is what makes statistical analysis possible. Using these distributional frequencies, one can learn which senses of a polysemous word are most closely related, what is the internal semantic structure of a polysemous term, which characteristics distinguish between near synonyms, and so on. To this end, a number of statistical techniques for analyzing categorical data are employed. For a more complete description of the behavioral profile of *po-* used in this study, please see Section 3.4 and Tables 3.4.1 & 3.4.2.

In this section I will describe my behavioral profile analysis of *po-*. The construction and use of a behavioral profile consists of four steps (adapted from Berez & Gries 2009):

1. Retrieve a representative random sample of all occurrences of a word's lemma[18] from a corpus, along with the context of that lemma[19].

2. Analyze the properties of each use of that lemma in context. The properties are called ID tags and represent the morphological, syntactic, and semantic characteristics of that particular use of the word. Each ID tag has several levels or values. Ex: The ID tag *transitivity* has the levels/values *transitive* and *intransitive*. This stage of analysis often involves manual annotation of all instances of the lemma for the relevant properties.

---

[18] A *lemma* can be defined as the canonical, citation form of a word. Thus *run*, *runs*, *running*, *ran* are all forms of the word *run*. Here *run* is the lemma. From a psycholinguistic perspective, a lemma is the abstract, pre-phonetic conception of a word; the "idea" of a word without its phonetic "embodiment".

[19] In this case I am retrieving not occurrences of a word but of the prefix *po-*, yet the principle remains the same.

3. Generate a co-occurrence table that illustrates how frequently each possible level of an ID tag occurs with each meaning or lemma in question. In the narrow sense of the term, this table constitutes the behavioral profile of the lemma in question (see Hanks 1996 for the origin of this concept).

4. Analyze the data in the behavioral profile using descriptive statistical techniques, such as summary frequencies, correlational/distance measures, and cluster analysis.

I will discuss what each of these steps entails and how I executed them in the four sections that follow.

3.2 Step 1: Collecting the random sample: Selecting a corpus, searching for and retrieving data

The first question that one must answer when conducting a corpus-based study of linguistic phenomena is rather basic: Which corpus is best suited for the study? For this dissertation I chose to use the Russian National Corpus (hereafter RNC), from which I extracted all my data. The reasons for choosing this corpus are several: First, the RNC is by far the largest Russian-language corpus currently available – as of July 2008 it contained over 160,000,000 words. Secondly, the RNC is internet-based, free, and is easily searched via a detailed search interface, permitting the user to narrow searches by selecting grammatical, semantic, and syntactic features for the search term, in addition to allowing collocational and colligational searches in a variety of registers, genres and text types. All texts in the RNC are vetted by native speakers and represent authentic Russian; no such assurance of authenticity can be offered by data obtained from Google searches or searches on the widely-used Russian search engine Yandex (www.yandex.ru). And since the data in the RNC are relatively stable, other researchers have free access to the same data and can verify, refute, or amend my findings on the basis of that same data.

This cannot be said for internet data, which are constantly changing and subject to erasure, whether by the text authors, because of site closures, or by malicious hacking. Studies based solely on internet data are to some degree less repeatable than experiments that use corpus data, and a more scientific investigation of prefixal meaning depends on repeatability (to say nothing of falsifiability) of the study and its results. Finally, an overwhelming advantage of the RNC over internet data is that all words are already tagged for many semantic and syntactic parameters. These tags can be extracted from the search results and used in data analysis, saving the researcher from the insurmountable task of tagging thousands of examples by hand for those same parameters.

All the data used in the current study were collected during the period 2 July 2008 – 18 July 2008. To understand how my data are organized, one must first understand the RNC's architecture. The RNC is broken down into 3 broad subcorpora: fiction, nonfiction, and spoken "texts". Some texts in the RNC are dehomonymized[20], and so for each of the three corpora I conducted one search of the dehomonymized texts and one of all texts within that corpus, whether dehomonymized or not. In addition, each corpus gives additional options for narrowing one's search – for instance, within the spoken corpus one can search public speech, non-public speech, and film/TV transcripts. I searched each corpus using "*po**" as my search term, and naturally I limited the search to those items marked as verbs. Since the fall of the Soviet regime in 1991 had far-reaching effects on all areas of Russian life, I chose to limit my search to those texts created from

[20] In the dehomonymized portion of the corpus, the word *peč'* 'to bake (v.)' or 'oven (n.)' is disambiguated in its tagging, and thus a search for the term *peč'* would yield instances that were tagged clearly as nouns or verbs. In the non-dehomonymized portion of the corpus, all instances of *peč'* are tagged as both noun and verb, because the tagging is carried out automatically by a computer program. Dehomonymization often requires disambiguation of homonyms by hand, hence the smaller size of the dehomonymized corpus.

1992 to 2008, thus minimizing any confounding influence of historical change on my data. Seeking to maximize the number of results and produce data that were maximally useful, I gathered data using the following searches:

In the dehomonymized portion of the RNC:
1: Corpus of spoken Russian (includes the subcorpora in 2, 3, and 4)
2: Public spoken texts
3: Non-public spoken texts
4: Language of Film
5:Nonfiction (all functional spheres)
6:Fiction

All texts (dehomonymized and non-dehomonymized):
7: Corpus of spoken Russian (includes the subcorpora in 8, 9, and 10)
8: Public spoken texts
9: Non-public spoken texts
10: Language of Film
11: Nonfiction (all functional spheres)
12: Fiction

There were no documents in the dehomonymized portion of the Language of Film subcorpus that corresponded to my search period (1992-2008), and thus no results from that subcorpus are included in this study. In addition, the nonfiction texts are searchable as a whole or by functional sphere, type of text, and theme. Because of ongoing improvements to the RNC, some spoken texts are scattered throughout the nonfiction subcorpus. To avoid interference of these spoken texts in my analysis, I collected nonfiction texts using the only search option that excluded spoken texts: nonfiction, all functional spheres.

The search of the entire corpus (both dehomonymized and non-dehomonymized texts) yielded a large dataset of over several hundred thousand observations. Unfortunately, this amount of data would later prove too unwieldy for analysis. As a result, the remainder of this dissertation uses only the data in the dehomonymized portion

of the RNC. But the dehomonymized corpus is by no means small and still provides a large field from which to collect data. Table 3.1-1 gives a breakdown of subcorpora sizes for each data-collection search in the dehomonymized corpus at the time the data was collected (July 2008[21]):

Table 3.2-1

Dehomonymized Corpus			# of documents	# of words
	All spoken texts		49	217,163
		Public	29	212,542
		Non-public	20	4,621
		Film	0	0
	Non-fiction		1,488	1,767,674
	Fiction		52	630,912
Total			1,589	2,615,749

The spoken texts were drawn from 49 sources: 29 represent public or "official" speech, while 20 were private speeches not intended for an audience. The line between public and non-public is sometimes unclear, and in some cases the division depends on one's point of view. For instance, a number of records came from transcripts of the popular reality-TV show "*Dom 2*". While it is true that the show is meant for public consumption, the actors (purportedly) behave as though no one were watching. Thus their interactions resemble private conversations more than radio interviews, so I chose to group them with the non-public sources. As such my division of private and public spoken texts differs slightly from that used by the RNC.

Each search within the subcorpora of the RNC returned a great quantity of results, and I wanted to collect all results in order to have the most representative data set

[21] As of 2010, the size of the dehomonymized corpus has doubled to over five million words.

possible. To that end I collaborated with Adrian Ilie, a computer science graduate student at UNC - Chapel Hill, to develop the Extractor program, which automatically extracts all results of a given RNC search along with all the associated tags. The Extractor then combines these results into a format that is easily imported into any major spreadsheet software, making statistical analyses of the data possible.

Having searched the RNC's subcorpora using the Extractor, I had to remove two types of "noise" from the resulting data to ensure the relevance of my analysis: *pod*-prefixed verbs and certain verbs that begin with the letters *po* but are not *po*-prefixed verbs (ex: *pol'zovat'sja*[i] 'use'). At this point in time the RNC is unable to distinguish between *po*-prefixed verbs and *pod*-prefixed verbs, since both begin with the letters *po*. As a result, I constructed a list of all *pod*-prefixed verbs found in Volume 17 of the Academy Dictionary (1950-1965) and automatically purged them from my dataset using the Extractor.

After removing noise items, my final dataset consisted of 16,121 **observations** (or **records**) – that is, 16,121 instances of *po*-prefixed verbs in use, along with all the semantic, syntactic, and discourse information provided by the context of the verb. From these observations I would later take 1,000 randomly-selected observations to conduct my statistical analysis. The results of these analyses can be generalized to the larger population of the 16,121 observations. As noted earlier in Section 2.3, creating a representative corpus is a complex, challenging task that requires much foresight on the part of the corpus builders. Collecting all utterances of a language and then randomly selecting a sample is impossible; as such, corpora can only be approximations of language use. Nevertheless, by ensuring that a corpus draws from a variety of sources

and genres, one can reduce the degree of bias in the data. Since my data are drawn from fiction, nonfiction, and spoken "texts" from over 1,500 sources, I am confident that my data reflect many facets of current Russian language use reasonably well. The RNC, though not perfect, is the best source of empirical data on modern Russian currently available; a study that makes comprehensive and responsible use of that empirical data can yield significant results.

3.3 Step 2: Analyze and annotate

3.3.1 Verb types

Not all *po*-prefixed verbs are created equal. In Russian there are a number of verbs that have *po-* in both the Perfective and Imperfective forms, such as *polučit'*[p]/*polučat'*[i] 'to receive'. And among verbs with *po-* as a prefix in either the Perfective or Imperfective, there are several kinds of morphological relationships between the prefixed and un-prefixed forms. To simplify matters, I developed the following classification system that divides *po*-prefixed verbs into five categories based on their morphological (not semantic) behavior – namely, according to the existence of an aspectual "partner" for each verb, and whether that verb and its "partner" are prefixed or not. Classifying the data according to this five-class system is necessary because, as the reader will see, the semantic contribution of *po-* is clear in only one of the five types of verbs, Type I. The six meanings of *po-* discussed in 2.3.4 occur only with Type I verbs and thus Type I is the focus of the remainder of this dissertation. The other types, after being detailed here, are treated briefly in 5.2.2. For purposes of our discussion, a verb is defined as having a

Perfective and Imperfective form (for those verbs that exist in both aspects) or only a single form (for *perfectiva tantum* or *imperfectiva tantum* verbs, or for biaspectual verbs). Reflexive forms are considered separate verbs from their non-reflexive counterparts; whereas reflexive and non-reflexive forms may seem to differ only in transitivity, this is not always the case – compare *pytat'*^{i} 'torture' vs. *pytat'sja*^{i} 'try, attempt' and *polučit '*^{p} / *polučat'*^{i} 'receive' vs. *polučit'sja*^{p} / *polučat'sja*^{i} 'turn out'. Treating reflexive and non-reflexive forms separately also gives us the opportunity to investigate more subtle differences in usage later. All verbs classified under each type are listed in Appendix 3. A description of each of these morphological types now follows.

Type I: *po|blagodarit'* 'thank'

The base verb (to which the prefix is added) is usually a simplex Imperfective stem[22]. In Type I the prefix *po-* makes a clear contribution to the resulting prefixed verb, even if that contribution is only to change the verb's aspectual status from Imperfective to Perfective. This type includes all six meanings of *po-* listed in 2.3.4. Ex: *pogovorit'*^{p} 'talk (for a while)' < *govorit'*^{i} 'talk, speak'; *poblagodarit'*^{p} 'thank' < *blagodarit'*^{i} 'thank'. Of the 709 verbs in my dataset of 16,121 observations, the vast majority belong to Type I – 533 verbs, or 75.2%. However, these verbs represent a minority of the observations in the data: Type I verbs account for only 38.2% of the total records (6,152 out of 16,121).

Type I verbs are the easiest to analyze since any grammatical or semantic difference between the prefixed and unprefixed forms can (in most cases) be attributed to

[22] There are two exceptions: Imperfectives derived from prefixed Perfectives that take *po-* to express the delimitative or distributive meaning, such as *povyplačivat'*^{p} 'pay out all' (distributive) or 'pay out a while' (delimitative), and already prefixed Perfective verbs that take *po-* to express the attenuative meaning, such as *porazvleč'*^{p} 'entertain a little'. These verbs are included in Type I because *po-* makes a recognizable contribution to their semantics.

prefixation with *po-*. Consequently, these verbs are the focus of this study, and from the set of 6,152 observations containing Type I verbs I selected a random sample of 1,000 records for statistical analysis (see Sections 4.2 and 4.3 for discussion). Type I verbs present the clearest opportunity for investigating and understanding the polysemy of *po-*. All Type I verbs are either Natural Perfectives or Complex Acts (Janda 2007). Natural Perfectives are what is traditionally known as the "aspectual partner" verbs of Imperfectives and can express the logical completion of the action; verbs of the resultative meaning fall under this rubric. Complex Acts encode actions that are non-completable and have duration, yet are nevertheless temporally closed (i.e. Perfective) (Janda 2007:609). Verbs of the attenuative, delimitative, distributive, ingressive, and intermittent-attenuative meanings qualify as Complex Acts.

There are also verbs of Type I that exhibit prefix variation: The verb *po|žarit'* 'roast, fry', for instance, has an alternate Perfective form with another prefix: *sžarit'* 'roast, fry'. As with other Type I verbs, there is an unprefixed Imperfective associated with a *po*-prefixed Perfective (a Natural Perfective). The Perfective form *sžarit'*, according to Yandex's dictionary site, has a meaning virtually identical to *požarit'*[p]. In other similar verbs it is possible that the use of different prefixes may alter the meaning of the resultant verb somewhat more than in this case; how to tease apart the differences between *sžarit'*[p] and *požarit'*[p], however, is the subject of another study. For now we must be content with the dictionary assertion that the verbs *sžarit'*[p] and *požarit'*[p] can be synonymous[23].

---

[23] The form *požarit'*[p] is synomymous with *sžarit'*[p] only in the purely resultative meaning 'roast, fry'. *Požarit'*[p] can also mean 'roast, fry for a while' in certain contexts, and in that delimitative meaning it is of course not synonymous with *sžarit'*[p].

Type II: *polučit'*[p]/*polučat'*[i] 'receive'

In Type II the Perfective *po*-prefixed verb is not built from an unprefixed Imperfective stem; instead, both the Perfective and Imperfective are prefixed by ***po-:*** *polučit'*[p]/*polučat'*[i] 'receive'. In Type II the contribution of the prefix, if any, is much less clear. These verbs represent, as Townsend (1975) describes them, fused roots; there exists no un-prefixed verb **lučat'*[i][24], to which *po-* is affixed to form the pair *polučit'*[p]/*polučat'*[i] 'receive'. As a result the semantic contribution of *po-*, if any, is not apparent. Of the 709 verbs found in my RNC data, 119 (16.8%)[25] belong to Type II, and yet they account for 9,242 observations – over half (57.3%) of all records in my dataset.

Because the semantic contribution of *po-* in these verbs is unclear, I will return to them only briefly at the conclusion of this dissertation to suggest how future research might shed more light on the topic (Section 5.2.2). Although not completely uninformative, they do not offer as much insight into the semantics of *po-* as the Type I verbs.

Type III: *po|slat', posylat'*[i] 'send'

Type III verbs generally comprise an aspectual trio: a simplex Imperfective, a prefixed Perfective, and a derived (prefixed) Imperfective. Example: The simplex *slat'*[i] 'send' has a prefixed Perfective *poslat'*, from which a (still prefixed) Imperfective can be derived by suffixation, *posylat',* resulting in two apparently synonymous Imperfectives.

---

[24] There is a group of verbs that share the morphologically identical root *–lučat' /-lučit'*, but the root does not seem to provide a common semantic basis for its prefixed derivatives (at least not synchronically).
[25] For the purposes of the summary statistics introduced here, I consider pairs of verbs that differ only in aspect (the traditional "aspectual partners") to be two aspectually-related forms of the same verb. For instance, I treat *postroit'*[p] / *stroit'*[i] 'build' as a single verb and *polučit'*[p]/*polučat'*[i] 'receive' as a single verb. My definition of a verb here is semantic rather than purely morphological.

Like Type II verbs, the contribution of *po-* in Type III verbs is not clear. Of the 709 verbs represented in my data, only 30 (4.2%) belong to Type III, and they account for only 403 observations (2.5%) of the total 16,121 records in my dataset. Veyrenc (1980:159-179) discusses this sort of duplicate Imperfective, concluding that the meaning of the derived Imperfective is more metaphorical (see also Guiraud-Weber 1993: 58-59). However, Veyrenc's conclusions are based on citations of 30 such verbs from Pushkin's collected works, only four of which appear in my data. Using data from the works of a single author raises the possibility of individual bias in language use, and Pushkin's semantics may not reflect current usage. Veyrenc's results should thus be considered in light of these limitations. Unfortunately a full elucidation of this problem is beyond the bounds of this study. As a result, the semantic contribution of *po-* in Type III verbs is not investigated further.

Type IV: *poricat'*[i] 'reproach'

These verbs have only one morphological form, and that form is either Perfective or Imperfective, without a corresponding aspectual "partner", whether prefixed or not. Example: the verb *poricat'*[i] 'reproach' is an Imperfective that has no aspectual partner. From a historical perspective verbs like *poricat'*[i] 'reproach' were formed by attaching *po-* to a stem, but now seem to function as indivisible units when it comes to aspectual morphology. Whether or not speakers today analyze these verbs as prefixed cannot be answered here, and may vary from speaker to speaker. There are only 25 of these verbs in the data I originally collected (3.8% of the total verbs) and they account for 180 records (1.1%). Because they offer no insight into the semantics of *po-* in contemporary Russian, they will not be considered further.

Type V: *pokupat'*^{i} / *kupit'*^{p} 'buy'

This type contains only 3 verbs: *pokupat'*^{i} / *kupit'*^{p} 'buy', *pokupat'sja*^{i} / *kupit'sja*^{p} 'be bought', and *po|nukat'*^{i} 'urge on'. In the first two, the relationship of prefixed Perfective and simplex Imperfective is reversed: *pokupat'* 'buy' is Imperfective, while *kupit'* 'buy' is Perfective. In *po|nukat'*^{i} 'urge on', both forms are Imperfective. As such, these three verbs do not fit into any of the foregoing categories, and they are unrevealing with respect to the semantics of *po-*. They make up only 0.4% of the total verbs, and they account for 0.9% of all records (due largely to the high frequency of *pokupat'*^{i} 'buy'). These three verbs are excluded from further analysis.

Some verbs can belong to more than one type, depending on the meaning they express. For instance, *postavit'*^{p} 'put, place' is usually paired with the unprefixed *stavit'*^{i} 'put place'. However, *postavit'*^{p} can also be used figuratively to mean 'supply, provide (fuel, supplies, etc.)', and in this sense the corresponding Imperfective is the prefixed *postavljat'*^{i}. I considered those instances of *postavit'*^{p} in the sense of 'put, place' as Type I, and those instances in the sense 'supply, provide' as Type II. Separating the two in my data required painstaking reading of each instance to ensure proper classification.

To summarize: I divided the 709 verbs in my dataset according to the existence of an aspectual "partner" for that verb, and whether that "partner" is prefixed or not – the resulting verb types serve to highlight those verbs that could potentially reveal the most about the semantics of *po-*. Only in Type I verbs can the contribution of *po-* be reliably isolated from the semantics of the unprefixed stem. Type I verbs are thus the focus of my analysis, even though they are not the most frequently attested type (second to Type II). Types II, III, IV, and V present greater challenges for the semanticist since the

contribution of *po-*, if any, is much less obvious. Types III, IV and V are infrequent and play a marginal role at best. For these reasons the statistical analyses that follow (Sections 4.2 and 4.3) are based on a randomly-selected sample of Type I verbs. Types II, III, IV, and V will be visited again in Section 5.2.2. The absolute frequencies for each type in the dehomonymized portion of the RNC are summarized Table 3.3.1-1 (broken down by text type):

Table 3.3.1-1

Corpus			# of observations for each verb Type					Observations in subcorpus
			I	II	III	IV	V	
Dehomonymized corpus								
	All spoken texts		699	874	19	5	25	1,622
		Public	393	575	11	5	14	998
		Non-public	306	299	8	0	11	624
	Non-fiction		4,296	7,573	304	149	103	12,425
	Fiction		1,157	794	80	26	17	2,074
Total observations by Type			6,152	9,242	403	180	145	16,121

Table 3.3.1-2 (the same frequencies given as percentages[26])

Corpus			% of observations for each verb Type					% text type of total texts
			I	II	III	IV	V	
Dehomonymized corpus								
	All spoken texts		43.1%	53.9%	1.2%	0.3%	1.5%	10.1%
		Public	39.4%	57.6%	1.1%	0.5%	1.4%	6.2%
		Non-public	49.0%	47.9%	1.3%	0.0%	1.8%	3.9%
	Non-fiction		34.6%	60.9%	2.4%	1.2%	0.8%	77.1%
	Fiction		55.8%	38.2%	3.9%	1.3%	0.8%	12.9%
Total observations by Type			38.2%	57.3%	2.5%	1.1%	0.9%	100%

[26] Not all percentages add up exactly to 100% because of rounding. The rightmost column adds up to 100.1% because of rounding error.

Curiously, it seems that Type II verbs are more frequent in non-fiction and oral communications intended for public consumption than they are in fiction and private oral communications. Type I verbs exhibit the opposite trend. No clear explanation for this association is apparent. [tables now placed after description]

3.3.2 Organization of data: ID tags and ID tag levels

Here I will give a breakdown of how data is organized in the 1,000 randomly-selected observations[27] that are the subject of the statistical analyses in Chapter 4. Each item in my dataset (called an **observation** or **record**) is composed of a single instance of a *po-* prefixed verb in context, along with information about that verb and its context, such as tense, aspect, person, number, and gender of the verb (where applicable). Here is what a portion of a single observation in my dataset would look like:

Fig. 3.3.2-1

ID:	16778
Random #:	844
Text:	*Kogda v teatr prideš' / tam srazu* ***počuvstvueš'****...*
Passport*:*	*Biografija (beseda lingvista s informantom), Sankt-Peterburg // Arxiv Xel'sinkskogo universiteta* (1997)
Infinitive:	*počuvstvovat'*
Verb Type:	I
Transitivity:	transitive
Voice:	active
Semantic class[28]:	mental/psychological/emotional
Sentence type:	declarative
Text type:	spoken

[27] This sample size was chosen because, generally speaking, a sample size of 1,000 randomly-selected items can be considered representative of the population from which that sample was taken; it gives a relatively small margin of error for a number of statistical analyses.

[28] The semantic classes used in this study are condensed from the set used by the RNC. In cases where semantic class was not automatically provided by the RNC, the observations were manually annotated using the RNC's tagged observations as a guide.

Each observation is given a unique ID number (16778 here) that distinguishes it from all others, and each observation receives a randomly-generated number (844 in the example just given) by which I was able to select a random sample of 1,000 observations. The text containing the verb-in-context is found in the "Text" field. The field called "Passport" gives the source of the text – in this case, a conversation between a linguist and an informant. The remaining fields ("Text type", "Verb class", "Transitivity", etc.) are referred to as **ID tags**. ID tags encode relevant information about the *po*-prefixed verb form in each observation: whether the verb is transitive or intransitive, whether the verb is in past, present or future tense, whether the verb is Perfective or Imperfective, etc. Thus each ID tag has more than one possible value (or **level**): The ID tag "Transitivity" can have one of two levels, transitive or intransitive. A list of some of the ID tags and ID tag levels used in this study follows here. For the complete set of ID tags and ID tag levels, please see Appendix 5.

Table 3.3.2-2

Kind of ID tag	ID tag	Levels of ID tag
Verbs:		
morphological	transitivity	transitive, intransitive
	voice	active, middle, passive
	tense-mood	gerund, imperative, indicative-past, etc.
	aspect	perfective, imperfective
	gender	masculine, feminine, neuter
	number	singular, plural
	person	first, second, third
	case (for participles)	nominative, genitive, dative, accusative, instrumental, locative
semantic	semantic type	abstract action, change of state, existence/being, human qualities/behavior, impact/contact/support, location/placement, etc.
	prefix meaning	attenuative, delimitative, distributive, etc.

syntactic	negation	positive, verb negated, preceding verb negated, clause negated, *ne*-word
	sentence type	declarative, interrogative, imperative, exclamation (not imperative)
	clause type	dependent, independent
	(dependent clause type)	adjectival/appositive, gerundial, spatial, temporal, relative, etc.
Collocates:		
lexical	adverbial	duration (*dolgo*), manner (*bystro*, *legko*), futility (*zrja*), etc.
	particle	exhortation (*pust'*), restriction (*tol'ko*), intensification (*daže*), etc.
subject tags	syntactic type	nominative, implied nominative, dative to impersonal verb, dative to personal verb, impersonal construction (no subject), etc.
	number	singular, plural
	animacy	animate, inanimate
	countability of nouns	count, mass
	level of abstraction	concrete, abstract
	semantic type	human/supernatural, animal, plant, (social) events, mental/psychological/emotional, etc.
object tags	syntactic type	nominative, genitive, dative, accusative, instrumental, locative, *čto*-clause, etc.
	number	singular, plural
	animacy	animate, inanimate
	countability of nouns	count, mass
	level of abstraction	concrete, abstract
	semantic type	human/supernatural, animal, plant, (social) events, mental/psychological/emotional, etc.
Other:		
	text type	fiction, nonfiction, spoken

The list of ID tags is potentially endless, and one could choose to encode much more semantic, syntactic, or other grammatical information than I did here. The ID tags used in this study were chosen for at least two reasons: 1) Many ID tags were already embedded in the RNC and could be extracted automatically, and 2) the manually-annotated tags provide additional information about the *po*-prefixed verbs in question, or they provide information about syntactically-defined collocates, which can provide clues to meaning

(Atkins 1987; Hanks 1996; Gries & Divjak 2009). Tagging syntactically-defined collocates is less arbitrary and more comprehensive than tagging collocates within a certain distance (in number of words) from the *po*-prefixed verb (Stefanowitsch & Gries 2003), and syntactically-defined collocates have been shown to be significantly more informative in corpus-based semantic analysis than collocates from an arbitrarily defined word-distance window (Gries & Stefanowitsch forthcoming). More is not always better when it comes to ID tags; sometimes smaller sets of ID tags better discriminate between word meanings than do large sets of ID tags (Arppe 2008). My dataset contains the 29 ID tags listed in Appendix 5; similar sets of ID tags have been employed in other studies (Divjak & Gries 2006, Gries & Divjak 2008, Berez & Gries 2009), and these 29 were chosen as most pertinent to my investigation.

3.3.3 Assignment of ID tag values

Many ID tags and their values were provided by the RNC automatically, such as tense, aspect, number, gender – that is, the grammatical ID tags directly associated with the verb. Semantic type of the verb was also provided, but not for all verbs, and so manual annotation of that property was necessary. No ID tags were provided for collocates (subjects, objects, adverbs, prepositional phrases) of the verb – all of these were thus tagged manually.

For the most part, the assignment of manually-annotated ID tag values is an objective affair – the number, case, and animacy[29] of a direct object requires virtually no subjective judgment on the part of the researcher. Semantic ID tags, however, do come

[29] As Frairie (1992) notes, small creatures display variable animacy in Russian, and some nouns exhibit facultative animacy. These facts are considered when tagging for the property of animacy.

with an inherent level of subjectivity. Corpus linguists openly acknowledge this level of subjectivity as an inescapable part of a linguistic analysis that relies on the assignment of semantic labels. This does not make the findings of a corpus linguistic study any less valid than other types of semantic analyses, which often rely even more heavily on introspection. Indeed, semantic labels are not entirely subjective – when classifying verb subjects, it is obvious that *čelovek* 'person' is best classified as *human*, not *tool* or *psychological/emotional experience*. And though the differences between separate senses of a polysemous term are semantic in nature, non-semantic correlations are often just as effective in determining semantic category membership (Berez & Gries 2009:163).

Some problems arose when determining which meaning of *po-* was intended in a given instance – in some cases two different meanings seem equally plausible, leaving the meaning of the *po*-prefixed verb ambiguous. There are several reasons for this ambiguity. First, as Janda (2010b) notes, many verbs permit more than one construal and are ambiguous with respect to Completability (Janda 2007b)[30]. A good example of this is *podumat'*[p] 'think'. Compare the following:

> *Razve ty* ***podumal***[p]*, čto ja razljubil svoju stranu?*
> 'Did you really think that I quit loving my country?' [telic predicate]
>
> *On* ***podumal***[p] *nemnogo i soglasilsja.*
> 'He gave it some thought and agreed.' [atelic predicate]

In the first sentence *podumat'*[p] 'think' is construed as Completable – the verb has a direct object (the *čto* 'that' clause) and once that thought is completed, the action that the verb

[30] The dimension of Completability, as its name suggests, describes actions that can be construed as leading towards some sort of culmination (Janda 2008a), has a goal, and results in a change of state (Janda 2007b:615, 2008b). As such the dimension of Completability closely resembles the telic vs. atelic distinction (which goes by a number of names in the literature) and plays an important role in the Russian aspectual system; see Janda (2007b, 2008b) for discussion.

encodes has reached its terminus. Contrast this telic construal of thinking with the Non-completable construal in the second sentence, where the act of thinking has no inherent endpoint – the man mentioned could have thought for several hours more if he wanted to and we would still say *On* ***podumal***^{p} 'He thought'. This situation is not hard to understand given what we know about human thinking: Sometimes a full, complete thought crosses our minds in an instant of time ("I thought I saw him dart by"), and other times we think about something without a specific goal in mind ("I thought about her face for days and couldn't get the image out of my mind"). Since *po-* can express both delimitative and resultative meanings, it is the prefix used to encode both construals of *podumat'*^{p} 'think'. Anstatt (2002) comes to similar conclusions regarding the verb *poest'*^{p} 'eat', noting that this verb has a delimitative meaning when used with partitive genitive objects, but resultative meaning when used with accusative objects.

In a similar vein Dickey (2006:19-23)[31] notes that some typically delimitative verbs can express resultative meaning when combined with evaluative adverbs[32]:

> *Abxazskaja milicija* ***slavno porabotala***^{p}.
> 'The Abkhazi police **did a splendid job** [lit. worked splendidly].'

In this example the verb *porabotat'*^{p} 'work', which usually expresses the delimitative meaning 'work for a while', expresses resultative meaning when used with the evaluative

---

[31] Dickey refers to these usually-delimitative verbs with resultative meaning/function as belonging to the Delimitative Aktionsart (or procedural class), despite their resultative characteristics. In this dissertation I am investigating the semantics of *po-* and not the reality of Aktionsarten per se. The utility of the concept of Aktionsart/procedural verbs has been called into question (Krongauz 1998:128), and I will not attempt to resolve that issue here. Instead, I believe a semantic analysis of *po-* can be carried out regardless of one's stance on Aktionsarten. Given my non-committal stance, I classify verbs with resultative function as having resultative meaning; I do not complicate the analysis by analyzing those verbs as delimitatives masquerading as resultatives.

[32] I am grateful to a native speaker who pointed out that there may exist certain contexts when an evaluative adverb does not necessarily signal the resultative meaning in *po*-prefixed verbs that usually express delimitative meaning – this shift in meaning may not occur in instances of irony or sarcasm, for example. More discourse context is needed in order to determine the intended meaning of a *po*-prefixed verb in such cases.

adverb *slavno* 'splendidly'. Likewise, Zaliznjak & Šmelev (2000:112) note that "verbs of the delimitative Aktionsart have a tendency to turn into aspectual partner verbs" (translation mine). And a similar situation exists with regard to ingressive verbs (Isačenko 1960:231):

> In certain cases (for example, in the past tense, in the perfect meaning of the Perfective, in the imperative), the ingressive meaning of verbs such as *pojti*[p] 'go', *pobežat'*[p] 'run', *poletet'*[p] 'fly', etc., can be lost. The sentence *On pošel*[p] *v gorod* 'He walked to town' can mean "He set out on foot, began walking to town" (ingressive meaning), or rather "He left, he's not here" (the perfect meaning of the Perfective past). [translation and glosses mine]

These delimitative and ingressive verbs that behave as resultatives are instances of the PART-WHOLE metonymy mentioned in 2.3.4: Both the delimitative and ingressive meanings indicate a partial traversal of the verb's metaphorical path, and this partial traversal can in turn stand for the whole traversal of that path, allowing these verbs to acquire resultative meaning in some contexts.

What does this mean with regard to tagging the instances of verbs in my dataset? It is clear that ID tag values cannot be assigned to verbs simply because they are traditionally analyzed as having one particular meaning; verbs like *porabotat'*[p] 'work' may not always express delimitative meaning, and the familiar ingressive *pojti*[p] 'go' may not be ingressive in all contexts. Thus no automatic assignment of tags is possible; instead, careful attention must be paid to the meaning of the verb in each context, and the appropriate value can be assigned only after the meaning in context has been determined.

However, it is the process of determining the meaning in context that is most open to incorrect interpretation and holds the greatest potential for non-native speaker error. To counter both of these negative factors, I enlisted the help of an educated native speaker (a non-linguist) to decipher the meaning of all observations that were (potentially)

ambiguous. As a result of that collaboration, I was able to assign with confidence all but 71 observations to one meaning or another. The remaining 71 were then assigned to the following categories:

attenuative-delimitative	(1)
resultative-delimitative	(57)
resultative-distributive	(1)
resultative-ingressive	(12)

We would expect each of the four ambiguous classes to exist, given what is known about each meaning of *po-*: The attenuative is similar to the delimitative, except that intensity, not time, is delimited; delimitatives sometimes behave like resultatives; ingressives may lose their ingressivity in certain contexts; and the distributive is much like a multiplied or repeated resultative. By assigning a given instance of a verb to one of these categories, I am not claiming that the speaker did not have a specific construal of the event in mind, whether delimitative, or resultative, or ingressive, etc. Rather, I am saying that there is not sufficient evidence from the hearer's perspective to be certain which construal was intended. The existence of these "ambiguous" categories are for the purpose of analyzing data, and do not imply a theoretical commitment to (or proposal of) new, "in-between" meanings of *po-*. These ambiguous cases provide some clues about the semantic structure of *po-* and are discussed again in Section 4.4

Finally, it should be noted that of the six recognized meanings of *po-*, one did not occur conclusively in my data: the attenuative meaning. Although I was inclined to label one observation (out of the 1,000 in my sample) as attenuative, consultation with a native speaker revealed that a delimitative reading was also likely. I thus chose to label that observation as "attenuative-delimitative" to capture the ambiguity between those two readings; without access to the original speaker of the utterance, the intended meaning

remains uncertain. But even if this observation were best interpreted as an example of the attenuative, the impact of that classification on my study is close to nil; no conclusions can be drawn about the attenuative with a sample size of one. I can tentatively conclude, however, that the attenuative meaning is rare in modern Russian and unlikely to exert much influence on the semantic structure of *po-*. This makes sense, as attenuative verbs are frequently the result of adding *po-* to an already-prefixed verb (Isačenko 1960:238-239), and double prefixation in Russian is rare (Ludwig 1995).

3.3.4 Dataset management

A few notes about the software that supported the annotation process are in order. Once the data were extracted from the RNC web interface using the Extractor program, the resulting .txt files were imported into Microsoft Excel. However, MS Excel can be rather inefficient when handling large amounts of relational data. For instance, not every verb in my dataset has a direct object. For those verbs that do govern a direct object, I wished to record values for several ID tags: animacy, case, number, whether the object was a count or mass noun, etc. Because data in Excel exist only in two-dimensional grids, either some data rows would have to be repeated a number of times to account for these multiple ID tags, or the number of columns would have to be multiplied (meaning those observations lacking a direct object would have a multitude of empty column values associated with them). A relational database is much more suited to data of this nature, and so I decided to house my database in an MS Access database. Chuck Simmons, a professional MS Access developer and software engineer at Nortel in Raleigh, North Carolina, kindly assisted with the construction of the database framework. Design assistance was also

obtained from John Wrobel, a database developer for the North Carolina Institute for Public Health at the University of North Carolina. All data entry and annotation was carried out by myself.

The data harvested from the RNC were imported from MS Excel into MS Access. Separate tables in Access were constructed for various types of information. For instance, one table housed all the data extracted from the RNC and the automatically-generated ID tags (and levels) associated with that data, while other tables held information about the subjects associated with each verb, the objects of the verb, information about participles, and information about verb type and infinitive (all manually annotated). Appropriate relationships between tables were then established. In general, the Access database made manual annotation of the data much faster and less prone to typographical error, while at the same time reducing redundancy in data storage. Manually-entered information was double-checked for accuracy at the end of this process, and any errors were corrected.

3.4 Step 3: Generation of co-occurrence tables

Once all the data have been annotated, co-occurrence tables can be generated. Co-occurrence tables show the distribution of ID tag levels for each meaning under investigation, and thus these tables can reveal the most basic associations between ID tag levels and meanings – the heart of a behavioral profile analysis. The meanings of *po-* are given as column headings, and the ID tag levels as row headings (Tables 3.4-1 and 3.4-2). In a co-occurrence table showing the absolute frequencies of ID tag levels, the numbers in the cells represent the number of times an ID tag level occurred in

conjunction with a given meaning of *po-*. Table 3.4-1 shows the absolute frequencies of the levels of the ID tag *voice*:

Table 3.4-1

ID tag	ID tag level	delimitative	distributive	ingressive	intermittent-attenuative	resultative	Total
voice	act	92	7	139	22	453	713
	med	19	2	12	4	120	157
	pass	2	2	0	0	55	59
Total		113	11	151	26	628	929

So from Table 3.4-1 we see that 92 of the 113 delimitatives in the dataset are in the active voice, while 19 are middle voice (verbs ending in *–sja*), and only 2 are passives. Ingressives and intermittent-attenuatives never occur in the passive voice in my dataset, whereas 55 of the 628 resultatives are passive. However, since the absolute frequencies of each meaning vary widely (compare 11 distributives to 628 resultatives), the relative frequencies are used to compare the distribution of voice-types (active, middle, passive) within each meaning. Relative frequencies are found by dividing the absolute frequency by the total instances of that meaning. For example, 139 of the 151 ingressives are active, giving 139/151 = 0.9205298 x 100 = 92.1% of all ingressives are in the active voice. Relative frequencies for each ID tag level are given in Table 3.4-2:

Table 3.4-2

ID tag	ID tag level	delimitative	distributive	ingressive	intermittent-attenuative	resultative
voice	act	81.4%	63.6%	92.1%	84.6%	72.1%
	med	16.8%	18.2%	7.9%	15.4%	19.1%
	pass	01.9%	18.2%	0.0%	0.0%	8.8%
Total		100%	100%	100%	100%	100%

These percentages allow us to compare the distribution of ID tag levels for each meaning easily. For instance, we can now see that even though 55 of all resultatives are in the passive voice, this represents only a small fraction of resultative observations (8.8%), which patterns somewhere between distributives (18.2%) and delimitatives (1.9%). But since we are dealing with a sample of a population (and not the population itself), we must keep in mind the margin of error for each percentage. A small margin of error is desirable, since it means that the value in the table reflects the actual value fairly well. Because our sample consists of 1,000 observations[33], the margin of error is approximately ±3% – a small margin of error indeed. For instance, in Table 3.6 nearly 85% of intermittent-attenuatives are in the active voice. Taking the margin of error into consideration, we can say with confidence that between 82% and 88% of all intermittent-attenuatives in the dehomonymized portion of the RNC are in the active voice. If the margin of error were large, say ±10%, then we could only say that between 75% and 95% of all resultatives in dehomonymized portion of the RNC are in the active voice – a much less meaningful result.

3.5 Step 4: Statistical analysis

It is important to note that my data is categorical, not numerical: There is nothing numerical about whether a verb takes a genitive or an accusative complement, for instance. And while the frequency with which genitive or accusative complements occurs in the data may be a numerical value, the thing measured (complement type) is not,

[33] This number includes observations where the meaning of *po-* is ambiguous. The number of observations in which the meaning of *po-* is not ambiguous is 929, as indicated in Table 3.4-1.

which precludes the possibility of the data having a normal distribution[34]. This is an important point to make, as the nature of the data dictates which types of statistical analyses are permissible. Failing to recognize the limitations of the data can result in the choice of inappropriate statistical methods and thus questionable conclusions (see Gries 2006:80-81). For this dissertation, I use a hierarchical agglomerative cluster analysis (hereafter HAC, or simply cluster analysis) to determine relative similarity among meanings, and subsequently I use *t*-values and *z*-scores to investigate between- and within-cluster differences, respectively.

The purpose of a HAC analysis is to identify and understand the (dis)similarities between several related items, "clustering" similar items together in a dendrogram that resembles an upside-down tree. HAC groups items together so as to maximize the similarity of items within a cluster and to minimize the similarity between different clusters. This method has been frequently employed in linguistic analyses and has received robust support, both theoretical and empirical (see Manning & Schütze 2000; Gries 2006; Gries & Divjak 2008; Berez & Gries 2009; Gries & Stefanowitsch forthcoming). Generally speaking, a HAC involves three steps (adapted from the guidelines for behavioral profile analyses in Gries & Divjak 2008:65-67):

1. Once a behavioral profile has been generated for the lexical item(s) (or morphemes) in question, that behavioral profile must be converted to a similarity/dissimilarity matrix by means of an appropriate similarity/dissimilarity measure. The Canberra dissimilarity metric has worked well in similar linguistic studies (Kiss 1973, Gries 2006, Gries & Divjak 2008; see Deese 1965 for similar measures) and so was used here as well. After the dis/similarity matrix is

[34] In (over)simplified terms, a *normal distribution* refers to the following: If all the data points were plotted on a Cartesian-style graph, the majority would fall near the average value, with fewer and fewer points occurring as you move away from the average value – the graph depicts the well-known bell-curve. By their very nature, categorical (that is, non-numerical) data cannot have a normal distribution.

generated, then the meanings of *po-* must be clustered using an amalgamation algorithm – in this case Ward's rule was used, since the logic underlying Ward's rule is conceptually appropriate in this study. As stated before, step 1 of the analysis will result in a dendrogram that resembles an upside down tree – the shorter the "branches" connecting items in the dendrogram, the more similar those items are, while the longer the "branches", the less similar those items are. See Sections 4.1 and 4.2 for the dendrograms generated for the meanings of *po-* in my data.

2. Once items have been clustered according to step 1, the intra-cluster and inter-cluster similarities must be investigated. *T*-values can be used to determine which ID tags maximally distinguish clusters (i.e. to investigate inter-cluster differences).

3. Now that inter-cluster differences have been investigated, we can turn our attention to the evaluation of intra-cluster structure. Senses grouped together in a cluster by HAC are not necessarily highly related to one another; rather, the cluster analysis is simply saying that those senses are more similar to one another than they are to senses in other clusters. Standardized *z*-scores can be used to uncover more information about the internal structure of each cluster.

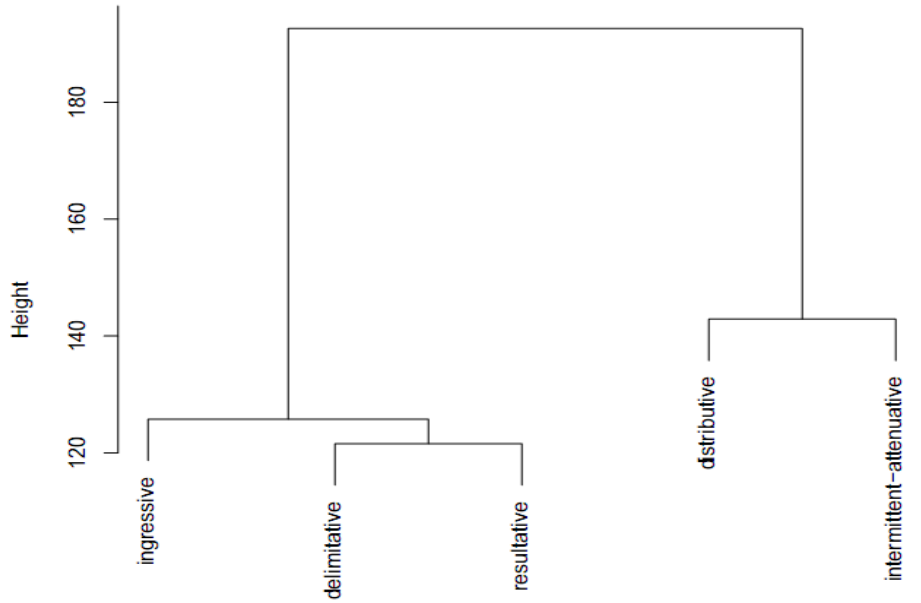
All of these steps are illustrated and explained in additional detail in Sections 4.1 – 4.3.

4 Results and Discussion

4.1 Introduction

This chapter focuses on the results of exploratory statistical analyses and how those results are interpreted to represent the cognitive network of *po-*, which consists of two clusters of meanings – [resultative + delimitative + attenuative + ingressive] and [distributive + intermittent-attenuative] – with the resultative as the prototype. In 4.2 a number of different variable sets are used to investigate the semantic structure of *po-*. While almost any combination of variables could hypothetically be used as the basis for a cluster analysis, in practice the number of useful variable sets is rather constrained. Since cluster analysis is a mathematical procedure, it is blind to the quality of the input; the old adage "garbage in, garbage out" applies. In the interest of linguistic validity, I use only those sets of variables that can be theoretically motivated: First I look at variables that pertain to the verb proper, then at syntactic variables within the sentence, semantic variables, variables associated only with the verbal complements, and finally clause- and sentence-level variables. Unmotivated sets of variables were excluded, since these would not produce meaningful results. The results support Hank's (1996) assertion that the meaning of a verb is reflected by the *totality* of its complementation patterns: The cluster structure produced by using all 29 variables coded in this study yields the following dendrogram, which is virtually identical to the several structures produced by subsets of variables from different levels of linguistic analysis (semantic, syntactic, discourse, etc.):

Fig. 4.1-1



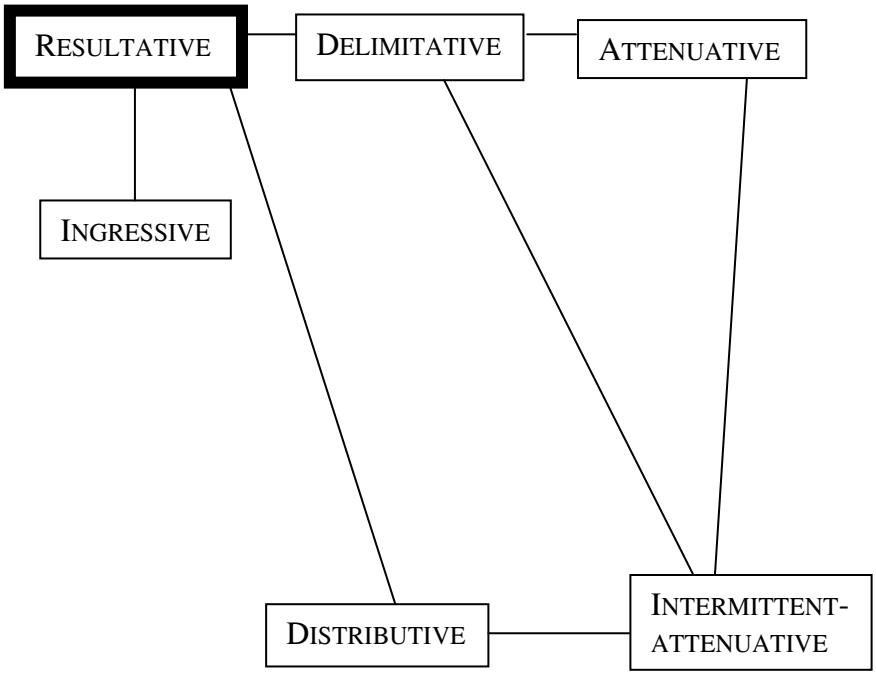
The two clusters are [ingressive + (delimitative + resultative)] and [distributive + intermittent-attenuative]. Hereafter I refer to the cluster [ingressive + (delimitative + resultative)] as cluster one and the cluster [distributive + intermittent-attenuative] I refer to as cluster two. The attenuative is omitted here because it is not attested in my data. The organization of *po-*'s meanings suggested by this dendrogram is new in the study of *po-* and is discussed in detail in 4.2.

In Section 4.3 I use *t*-values and *z*-scores to tease apart differences between and within the clusters, giving a very detailed list of similarities and differences among the meanings of *po-*, while pointing out some general semantic tendencies of each meaning: Resultatives express a wide variety of actions, more often involve inanimate actors, and they seem to be the meaning most common in verbs expressing a transfer from subject to object. Delimitatives tend to express a more specific scenario, centering on those actions performed by humans (and, more generally, living beings), and these activities quite

often have a cognitive component (though not always), and are volitional. Not surprisingly, the ingressive is distinguished by its focus on motion, both physical and metaphoric. Cluster two meanings (the distributive and intermittent-attenuative) are more common in the expressive language of fiction, and are more likely to be associated with the realm of the physically perceptible than their cluster one counterparts. The typical distributive scenario involves a plural subject and/or object involved in a transfer, which is not well-described by supplemental phrases; focus remains on the unmodified action expressed by the verb. Intermittent-attenuatives, on the other hand, involve singular subjects and/or objects, with more attention paid to exactly how the action was performed. Please note that these descriptions are stylized abstractions from the statistical data; I do not claim that all *po*-prefixed verbs of a given meaning must possess all (or in atypical cases, any) of the characteristics associated with that meaning; the statistical results merely capture dominant trends and do not constitute hard-and-fast rules for category membership.

The information from 4.2 and 4.3 is integrated in Section 4.4 with traditional (cf. Isačenko 1960; Zaliznjak & Šmelev 2000) and cognitive (Dickey 2006, 2007) analyses of *po-* to determine the prototypical meaning (the resultative; see 4.4.1) and to construct the following radial network representation of the relations among *po-*'s meanings:

Fig. 4.1-2



The resultative is the prototypical meaning (see 4.4.1 for discussion), hence the darker box surrounding it. Lines connecting meanings represent cognitive links between meanings (mostly metonymic; see 4.4.2 for full description), and the graphical distance between meanings is suggestive of the semantic distances revealed by the cluster analyses. Although the attenuative meaning is not attested in my data, it is included in this network on the basis of previous scholarship.

4.2 Hierarchical agglomerative cluster analysis of *po-*

As mentioned in Section 3.5, I used a hierarchical agglomerative cluster analysis (HAC analysis, or simply cluster analysis) to group the meanings of *po-* according to their relative (dis)similarity, and this grouping reflects the semantic relationships among the meanings of *po-* (hence the usefulness of cluster analysis). Here I will reiterate a few

points about HAC analysis that will help the reader understand the discussion of clustering results that follows. HAC is a family of statistical analyses that group items according to variables that characterize those items, or according to a (dis)similarity matrix based on those variables (Gries 2006; Divjak & Gries 2006). For this dissertation a (dis)similarity matrix was generated using the behavioral profile for each meaning of *po-*; each meaning's behavioral profile contains the relative frequencies of each attested value of the variables tagged in this study (refer to 3.3 and 3.4 for review). The Canberra dissimilarity metric has worked well in similar linguistic studies (Kiss 1973, Gries 2006, Gries & Divjak 2008; see Deese 1965 for similar measures) and so was used here as well to create the (dis)similarity matrix. Having created the matrix, the meanings of *po-* were clustered using Ward's amalgamation strategy (Ward 1963), since this strategy also has performed well in previous work and produces reasonably small clusters[35]. The result is a dendrogram resembling an inverted tree. Similar meanings amalgamate (or cluster) early – that is, near the bottom of the dendrogram – and successively dissimilar meanings or groups of meanings cluster later and later – that is, near the top of the dendrogram. How early or late meanings cluster in the dendrogram is a direct reflection of their (dis)similarity. Since HAC analyses are exploratory in nature, the results are not usually subjected to strict significance testing. Finally, it is important to bear two points in mind: First, the (dis)similarity implied in a dendrogram is relative among the items compared; even dissimilar meanings of *po-* can be more similar to each other than they are to

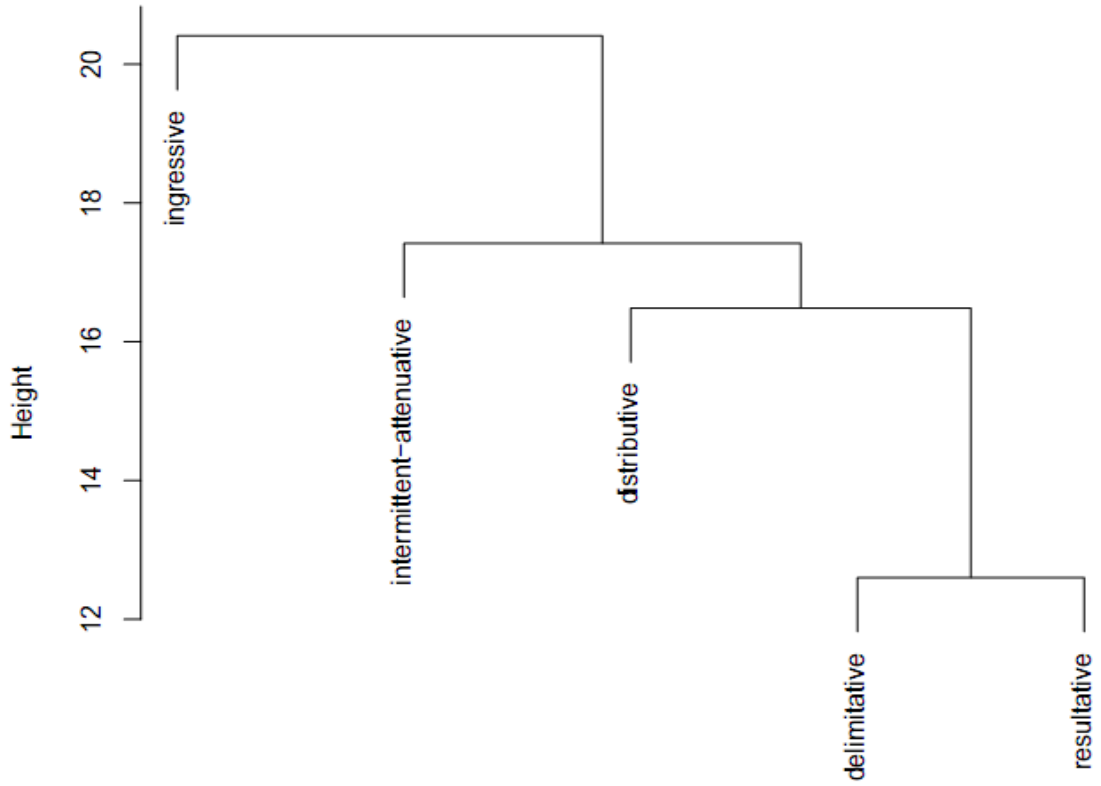
[35] It is possible that different (dis)similarity measures and amalgamation strategies may produce different clustering solutions. The two statistics employed here were chosen for their proven utility in other studies and for their conceptual similarity to the logic of this study. It is also worth noting that the number of clusters produced by an HAC analysis is not predetermined by the procedure but is rather determined by the nature of the data itself.

meanings of the prefix *vy*-, for instance. Thus there is no absolute scale of (dis)similarity implied. Secondly, the dendrograms (and, subsequently, the radial network derived from them) are abstractions that represent patterns of meaning/usage in a recent period of modern Russian (1992 – 2008) and aim to model structure in the language, not in the minds of specific individuals. This analysis is non-committal regarding the actual representation of *po*-'s semantics in the brain. It is possible that individual speakers may have mental representations differing somewhat from those presented here; idiosyncratic mental processes cannot be ruled out. Nevertheless, the very clear picture that emerges from this study is an empirically sound interpretation of the evidence and does much to resolve long-standing problems surrounding *po*-.

Disclaimers and caveats now aside, let us begin the investigation, progressing through sets of variables that belong to different linguistic dimensions. As has been pointed out (2.2.3), the meanings of *po*- are sublexical: *Po*-prefixed verbs are not semantically identical to their base verbs (except perhaps in the resultative meaning, which is often deemed "simple perfectivization"), but not different enough to warrant the formation of a derived Imperfective. Early studies on *po*- and Russian verbal prefixes in general assume a narrow scope of investigation – namely, the semantics of the prefix are understood with respect to certain characteristics of the prefixed verb, such as transitivity, temporal information, or even semantic type (Isačenko 1960; Flier 1975; Zaliznjak & Šmelev 2000). This sort of "verb-centric" information is encoded by several ID tags in my dataset: tense-mood, transitivity, voice, and semantic type of the verb. Following this earlier tradition of inquiry, I performed a cluster analysis using only these four variables

to see what, if anything, they might collectively reveal about the relationships among the meanings of *po-*. The results of that analysis produce the dendrogram in Fig. 4.2-1:

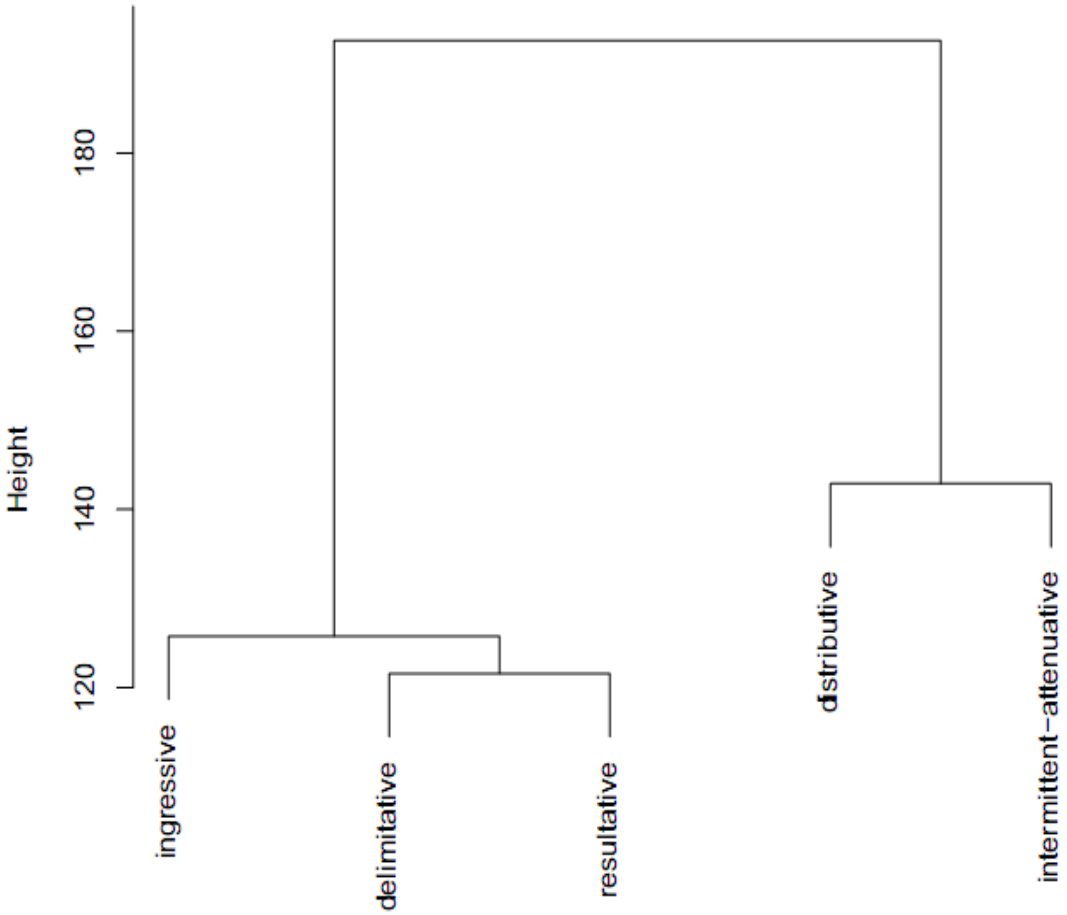
Fig. 4.2-1: Cluster dendrogram of "verb-centric" variables



The delimitative and resultative meanings amalgamate early, which is not surprising given the noted similarities between the two (Dickey 2006). On the basis of this dendrogram, one could potentially posit two clusters: one consisting of the delimitative and resultative, and the other consisting of the distributive, intermittent-attenuative, and ingressive. However, it is difficult to motivate such a division theoretically – ingressive verbs have long been known to express resultative meaning in many contexts, so it seems strange that it would be more similar to the distributive and intermittent-attenuative than to the resultative, as this diagram implies.

Using a wider range of variables, however, reveals more coherent relationships among the meanings of *po-*. Cognitive semantic analyses of Slavic verbal prefixes (Rudzka-Ostyn 1983a, Janda 1986, Tchizmarova 2006, Dickey 2007) look beyond the prefixed verb itself and consider other elements in the sentence, especially since complements frequently express trajectors and landmarks and may influence the meaning of the verb. However, the amount of "extra-verbal" information considered is often limited by the researcher's ability to recognize and process that information, and by the texts used as the basis for analysis. A behavioral profile of a prefix allows the researcher to overcome his or her human limitations by encoding large amounts of data in a statistically-analyzable format; subsequent analysis can detect patterns not readily apparent to the human eye. Such is the case with the HAC analysis of *po-*'s behavioral profile, which includes variables covering a broad range of linguistic information, drawing from both syntax and semantics, and encompasses collocates in several syntactically-defined slots (subject, direct object, indirect object, etc.). Using all 29 variables tagged in the behavioral profile, the dendrogram in Fig. 4.2-2 results (this dendrogram is identical to Fig. 4.1-1 presented earlier and is based on the same combination of variables):

Fig. 4.2-2

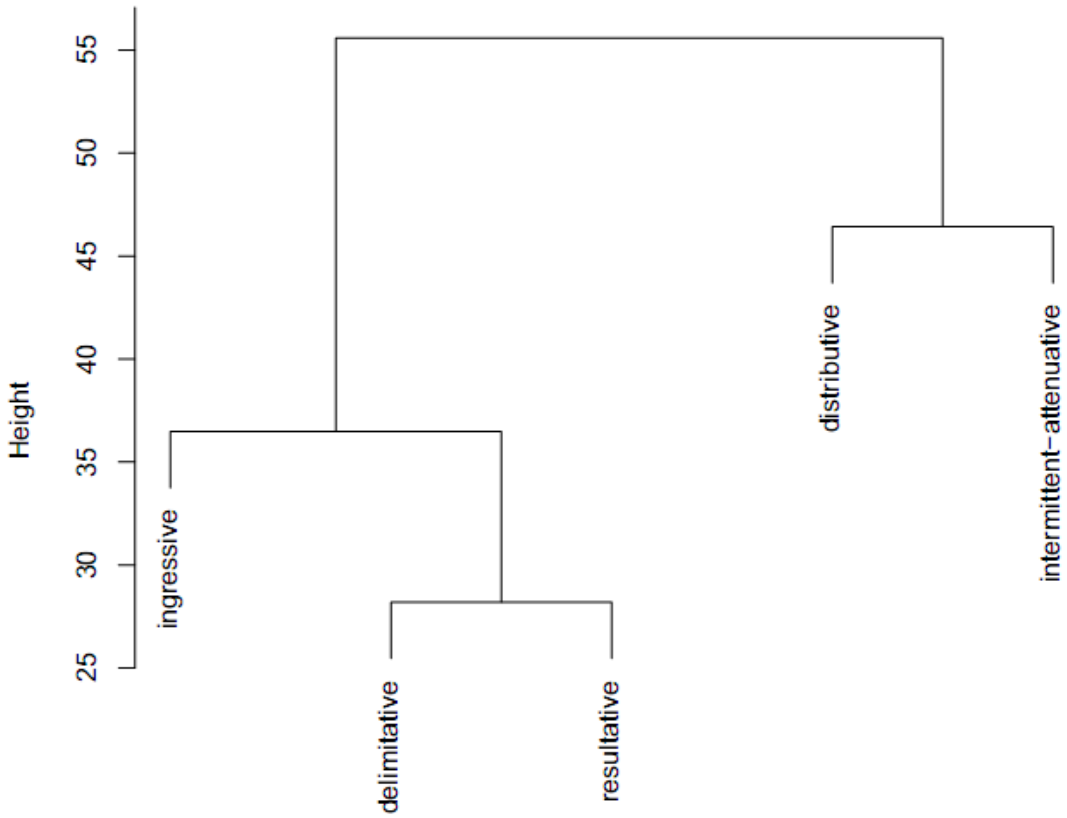


Here there are two distinct clusters: The delimitative and resultative meanings amalgamate first, followed by the ingressive; this is the first cluster. Then the distributive and intermittent-attenuative meanings amalgamate (the second cluster), to link with the first cluster much later. At first glance the attentive reader may be tempted to believe that the clustering algorithm simply grouped the least frequent senses together (distributive and intermittent-attenuative), while grouping the most frequent senses into another cluster (resultative, delimitative, ingressive) – that is, the clustering solution is based on raw frequency alone. However, this is not the case: Remember from Section 3.4 that raw tag frequencies in the behavioral profile were converted to *relative* frequencies for each

tag in each meaning, thus eliminating the effect of any difference in raw frequencies between the most and least frequent senses. The groupings here are thus not based on raw frequencies.

Next I investigate what syntactic variables might reveal about the relationships among the meanings of *po-*, and in particular whether a different picture might emerge when variables are restricted to only one level of linguistic analysis (syntax). To do so, I ran the cluster analysis using only the following variables: transitivity, voice, tense-mood, negation, sentence type, clause type, dependent clause type, object type, object number, participle number, subject type, and subject number (see Appendix 5 for a complete list of variables and values). The results strongly resemble those in 4.2-2:

Fig. 4.2-3

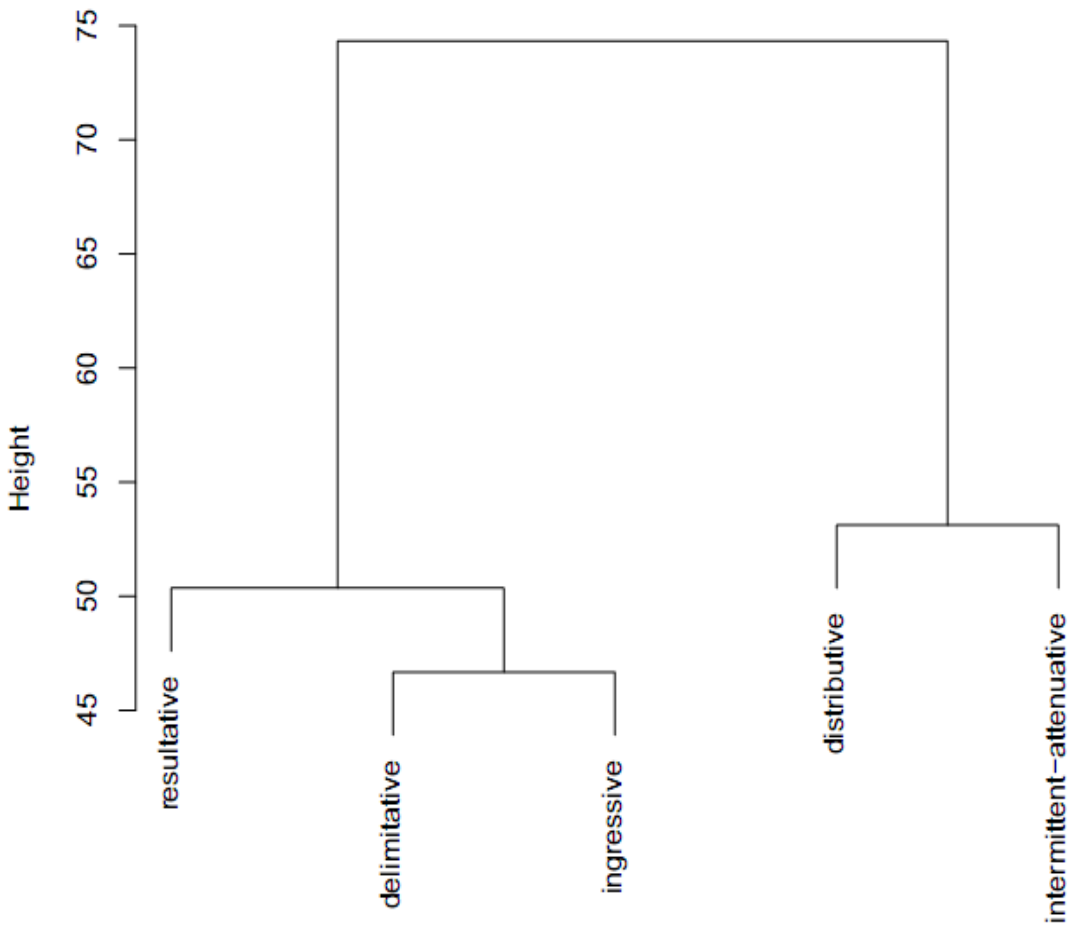


Although it appears that the ingressive meaning is amalgamating with the delimitative and resultative meanings later in 4.2-3 than in the preceding dendrogram, this is only because the scale in 4.2-3 is smaller than in 4.2-2; in both dendrograms the ingressive consistently amalgamates at roughly the same distance from the delimitative and resultative, meaning that we still have only two clusters.

To examine whether semantic variables[36] might produce different results than those presented thus far, I conducted two analyses that include only semantic variables and exclude all syntactic ID tags and the ID tag "text type". For the first analysis, I include only four variables representing the semantic type of the verb, of the subject (when present), of the object(s), and in the case of participles, the semantic type of the participle's headword. The results are remarkably similar to those of an analysis that encompasses all 29 ID tags:

[36]Cognitive linguistics differs from other linguistic paradigms in that it views semantics and syntax as ends of a continuum, rather than independent levels (as in the more traditional view). When dividing variables into "semantic" and "syntactic" sets, I include those variables that would traditionally belong to one class or the other, realizing that at the same time I am grouping variables that belong to one end of the semantic/syntactic continuum. While case in Russian has been shown to have a strong semantic component (Janda 2000, 2002a, 2002b; Janda & Clancy 2002, 2006), I chose to group case (represented by the ID tags "subject/object syntactic type") with the syntactic variables so that the results would be relevant to those working in more traditional paradigms as well. Test analyses grouping case variables with semantic variables (not discussed here) show that this alternative grouping did not produce any significant change in the results, probably due to the stability of *po-*'s meaning structure across linguistic levels/at both ends of the continuum.

Fig. 4.2-4

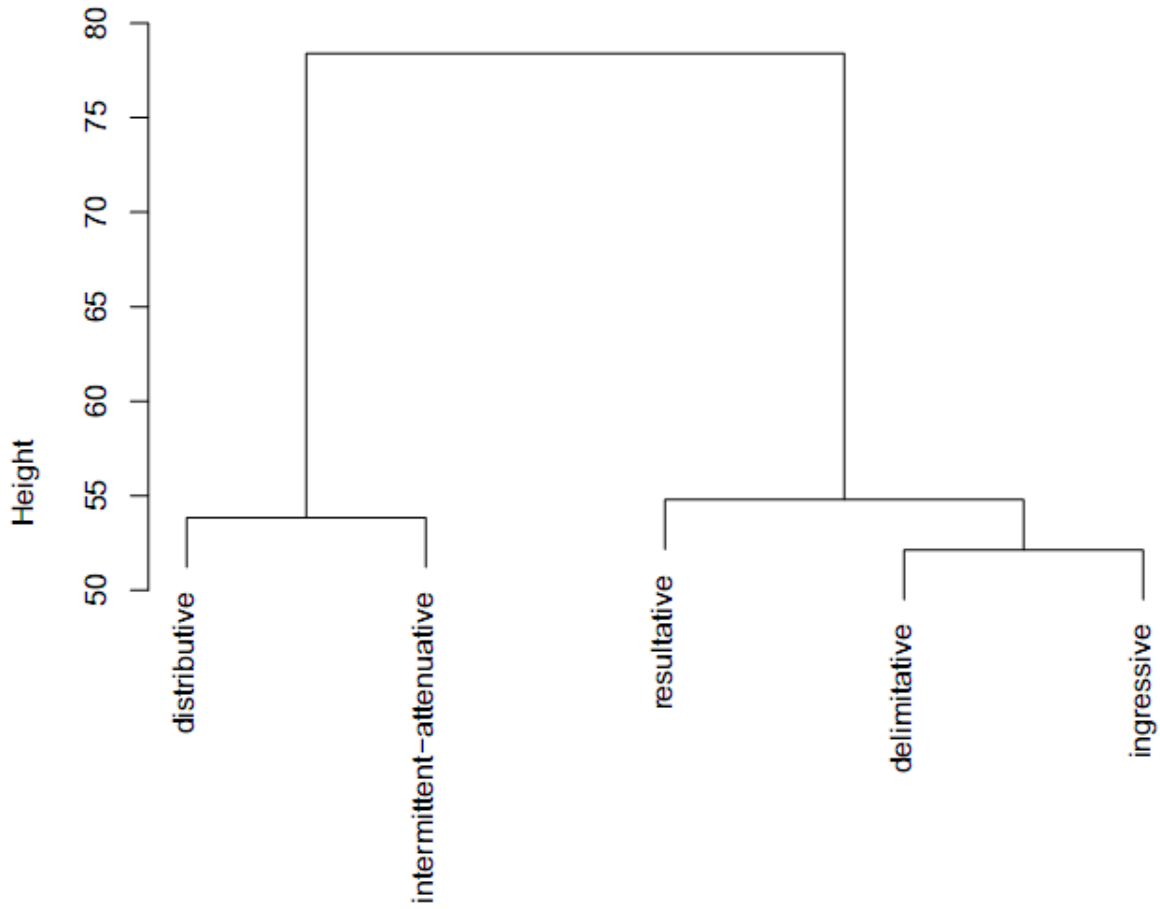


There is a minor difference between 4.2-4 and the earlier cluster dendrogram 4.2-2 (which uses all 29 variables) in that in 4.2-4 the delimitative amalgamates first with the ingressive, to be joined by the resultative shortly thereafter. However, the scale on the left shows that these three verbs still amalgamate very early, indicating that the difference between 4.2-2 and 4.2-4 is inconsequential.

The second analysis adds several variables to the set used in Fig. 4.2-4 and includes other types of distinctions that represent either more coarse-grained semantic distinctions (concrete vs. abstract for object and subject types) or at least have a semantic

component to them[37] (animacy of the subject and/or object, and whether the subject or object is a count or mass noun). The results remain unchanged, showing notable consistency in *po-*'s semantic structure, regardless of how narrowly or how broadly one investigates the semantics of the verb and its complements:

Fig. 4.2-5



[37] While the animate vs. inanimate and the count vs. mass distinctions do affect the morphological behavior of nouns in Russian, these distinctions ultimately have their roots in the meanings of the nouns themselves – that is, the nature of the real world entities to which those nouns refer. These are not grammatical distinctions on the order of masculine vs. feminine articles in Romance languages, and so their inclusion with other semantic variables is justified.

Note that the placement of the cluster [distributive + intermittent-attenuative] to the left of the other cluster is insignificant; we could just as easily move it to the right side of the diagram without changing the results.

It is interesting that criticisms of corpus analyses sometimes cite the process of classifying linguistic entities into one semantic group or another as a source of subjectivity and potential researcher bias in the results (Raukko 1999:87; see Berez & Gries 2009 for discussion). Despite this potential for researcher-induced distortion of the results, the clusters in 4.2-4 and 4.2-5 are identical to those presented in 4.2-2. Had the subjectivity inherent in semantic classification of the verb and its complements skewed the results, we would expect the dendrograms in 4.2-4 and 4.2-5 to present different cluster solutions – different from each other, perhaps, and certainly different from 4.2-2. However, the consistency of the cluster structures presented thus far suggests that researcher bias or subjective classifications did not adversely affect the results of this study. As the reader has already seen and will see in the remainder of this section, the cluster structure remains remarkably stable, even when strongly different sets of variables are used as the basis for analysis – again suggesting that the structure represented in 4.2-4 and 4.2-5 is no anomaly.

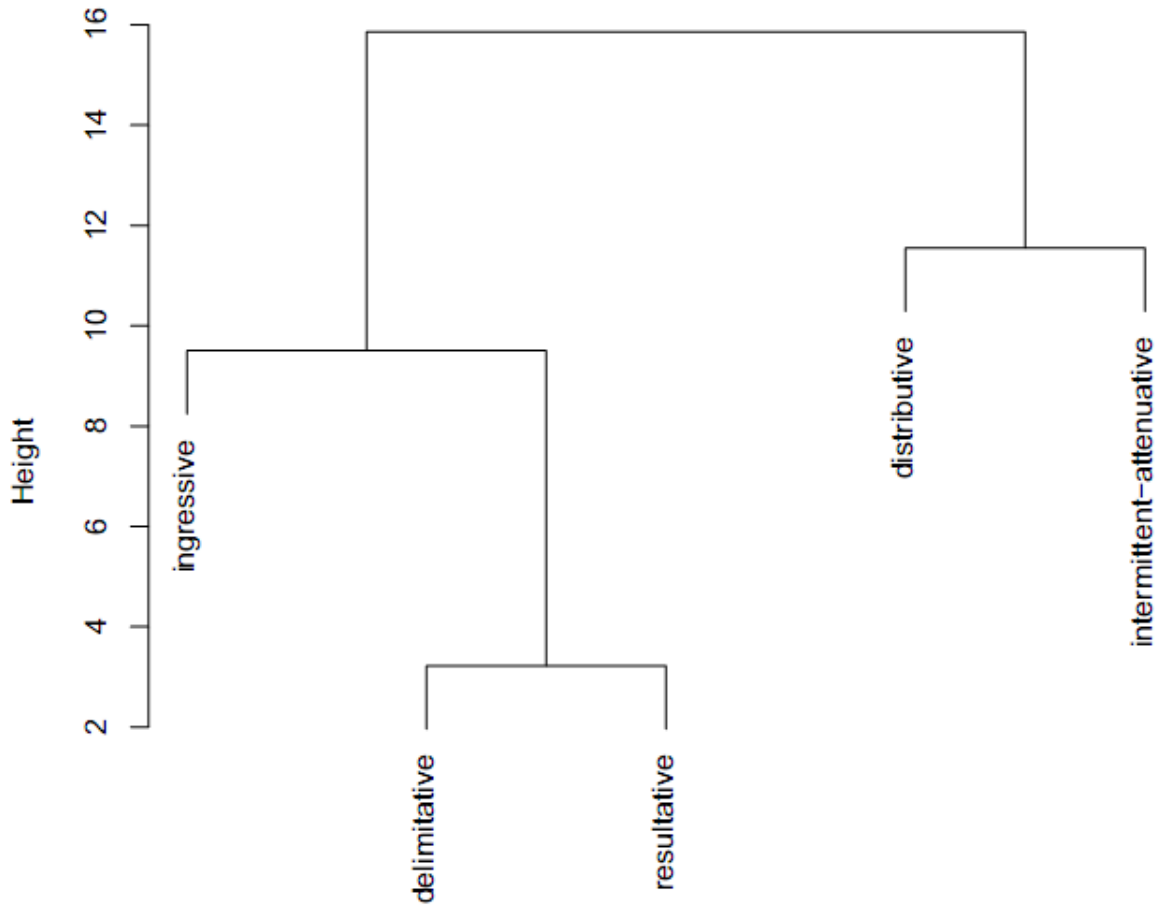
Given the stability of the relationships among the meanings of *po-* seen thus far, it is not surprising that if we consider only clause- and sentence-level variables (to the exclusion of all other syntactic and semantic information), the same structure remains. Fig. 4.2-6 displays the results of a cluster analysis based on only three variables[38]:

---

[38] A cluster analysis based on the pair of variables sentence type and clause type yields different results than an analysis based on sentence type and dependent clause type, with only the latter combination yielding a good match of the dendrograms in 4.2-2 through 4.2-6. However, the results of cluster analyses

sentence type (declarative, interrogative, exclamatory), clause type (dependent vs. independent), and type of dependent clause (spatial, temporal, relative, etc.):

Fig. 4.2-6



While at first glance the ingressive seems to amalgamate with the resultative and delimitative late, the scale here is much smaller than in many of the other analyses. The ingressive in Fig. 4.2-6 actually amalgamates at approximately the same distance from

---

based on one or two variables can be difficult to interpret. I can only say that the variables that make finer distinctions (sentence type and dependent clause type, each having a number of possible values) produce results consistent with what we have seen so far, while the very coarse-grained "clause type" (having only two values, dependent vs. independent) does not. It would seem that the coarse-grained distinction is insufficient to capture what is going on in the sentence.

(delimitative + resultative) as in 4.2-2 through 4.2-5, meaning that the cluster structure here in 4.2-6 is the same that we have seen so far.

One final note on the reliability of these cluster results: As with any statistical study based on data samples, there remains the possibility that the results are due to a random effect in the sample and do not reflect the characteristics of the population that the sample is intended to represent. In order to assess the chance that the cluster results obtained by the HAC analysis would not be found if I used other data samples, I ran the pvclust package for R (the statistical programming language used in this dissertation) on the complete behavioral profile of *po-*. The results indicate that both clusters, [resultative + delimitative + ingressive] and [distributive + intermittent-attenuative], are strongly supported by the data (Approximate Unbiased (AU) *p*-values exceed 99% for each cluster). The internal structure of cluster one was less strongly supported (AU *p*-value = 63%), but this is not surprising given the variation we have already seen in cluster analyses using different sets of variables and will not concern us further.

In sum, several conclusions can be drawn from the cluster analyses reviewed here. First, the semantic structure of *po-* is remarkably stable, cutting across several linguistic dimensions (semantic, syntactic, clause- and sentence-level distinctions). The five meanings examined belong to two clusters, the first consisting of [(delimitative + resultative) + ingressive], and the second consisting of [distributive + intermittent-attenuative]. Secondly, the level of subjectivity inherent in semantic classification of the verb and its complements produces no significant change in the cluster results. This suggests that the effects of subjective decisions on the researcher's part did not compromise the results of this study; there is no reason to exclude the semantic variables

from consideration. Thirdly, and perhaps most importantly, it appears that variables pertaining to the verb directly (tense-mood, transitivity, voice, semantic type) are insufficient in and of themselves to understand the semantic structure of *po-*. The cluster pattern produced by those four "verb-centric" variables (4.2-1) differs from the results of all other five analyses (which produce nearly identical results), despite the fact that those other five analyses employed widely dissimilar sets of variables as the basis for clustering. The study I present here underscores the fruitfulness of cognitive analyses based on behavioral profiles, especially in the study of Russian verbal prefixes: In this case the behavioral profile method allows the researcher to include more linguistic information (in an empirically sound manner) than was previously possible and reveals patterns of meaning undiscoverable by earlier methods. And given that language speakers and learners are exposed to the larger context in which the prefixed verb occurs, it only makes sense to include a broad range of variables in our analysis here, without imposing artificial restraints. Similar investigations of the remaining polysemous Russian verbal prefixes could prove equally enlightening.

4.3 Between- and within-cluster differences

In this section we will examine the two clusters in turn. For each cluster, we will first discover which parameters most strongly distinguish it from the other cluster, followed by a discussion of how the members of that cluster differ among themselves. To investigate between-cluster differences, *t*-values are a useful statistic (Gries & Divjak 2006, 2008) to determine which variables are strongly overrepresented in a cluster (resulting in high *t*-values, relatively speaking) and which variables are strongly

underrepresented in a cluster (yielding low *t*-values, relatively speaking). Since there are 223 ID tag levels attested in the data, full discussion of all results is neither feasible nor useful; I have chosen to restrict myself to the 30 ID tag levels with the highest *t*-values for each cluster (that is, just over 10% of all ID tag levels). High *t*-values (which indicate overrepresentation of a variable) generally indicate that the meanings in a cluster have a particular affinity for the variable in question, as compared to that variable's relationship to the meanings in the other cluster. To investigate difference among meanings *within* a single cluster, *z*-scores are employed. *Z*-scores are calculated for each meaning and each ID tag level using the normal approximation to the binomial distribution, which accounts for the discrepancies of sample sizes for each meaning of *po-*. If a given ID tag level yields a high *z*-score for a given meaning, then we can conclude that the ID tag level in question is strongly associated with that particular meaning, and low *z*-scores indicate strong dissociation with an ID tag level. Since the analysis produced a large number of *z*-scores (223 x 5 meanings), in the interest of space and maximum usefulness I restrict my discussion to most informative *z*-scores in each cluster, usually selected from the top and bottom 30 scores for each meaning of *po-*.

At this point a few notes to clarify the meanings of *t*-values and *z*-scores are in order: *T*-values do *not* reflect absolute frequency of an ID tag level within a cluster – that is, high *t*-values do not indicate that that ID tag level occurs frequently (in the absolute sense) in the cluster. A simple example will illustrate this difference. Let us assume the existence of two clusters (as in this study), and let us assume the existence of a variable "subject type" that has three possible levels/values: "human", "plant", and "animal". Let us then assume that "plant" has a high *t*-value in cluster one, and that "animal" has a high

t-value in cluster two. This means that in cluster one, "plant" is overrepresented – there is a stronger association between the subjects that are plants and cluster one than between plant subjects and cluster two. In cluster two, "animal" is overrepresented – there is a stronger association between subjects that are animals and cluster two than there is between animals and cluster one. However, this does *not* mean that plants are the most frequent type of subject in cluster one, nor does it mean that animals are the most frequent type of subject for verbs of cluster two. It is quite possible (even highly likely) that humans are the most frequent type of subject in both clusters. But since the relative frequency of human subjects in both cluster one and cluster two are about the same, we cannot say that humans are over- or underrepresented in either cluster; human subjects do not distinguish between the clusters, only animal and plant subjects do. Thus the *t*-value does not tell us anything about the most frequent type of subject. Its only purpose is to tell us which types of subjects *distinguish* between the two clusters. Thus the *t*-values given in 4.3.1 and 4.3.2 do not reflect the most frequent associations of each cluster – instead, *t*-values serve only to distinguish the clusters from one another, which makes them invaluable to the study of polysemy presented in this dissertation.

Secondly, *z*-scores tell us which ID tag levels are overrepresented in the observations for one member of a cluster as compared to the average expected representation of that ID tag level. *Z*-scores take into consideration the size of the sample, which is important since the samples for each meaning of *po-* in this study vary in size – for instance, the resultative meaning is found in 628 observations, whereas the delimitative is represented in only 113 observations. As was the case with *t*-values, *z*-scores do not tell us which levels of an ID tag are most frequent; instead, *z*-scores tell us

which levels are overrepresented. For instance, for the ID tag coding semantic type of the verb, the specific level coding "movement" returns the following *z*-scores for cluster one: resultative, z = -12.518; delimitative, z = -5.425; ingressive, z = 12.409. This means that verbs coding movement are strongly underrepresented in the resultative and delimitative; these two meanings seem to "avoid" combining with verbs that express motion. The ingressive, on the other hand, shows a strong association with verbs coding motion, and this semantic group of verbs is overrepresented in the ingressive. The reader will recall from the cluster dendrograms (Fig. 4.4-2) that the ingressive is separated from the primary resultative-delimitative pair. The difference in under- and overrepresentation of verbs coding movement is thus one of the variables that accounts for that branching, and we can say that the ingressive meaning is distinguished by its affinity for verbs expressing motion[39].

4.3.1 Cluster 1: [(delimitative + resultative) + ingressive]

The top 30 *t*-values for this cluster fall under 13 different ID tags. As the cluster analyses in 4.2 suggest, the top 30 *t*-values represent variables from several linguistic dimensions; for ease of discussion, I group these variables as follows:

- variables that pertain directly to the verb (tense-mood, semantic type)
- clause- and higher-level variables (sentence, dependent clause, and text type)
- specific collocates (adverbs, prepositions)
- variables pertaining to the object (animacy, count/mass, semantic type)
- variables pertaining to the subject (semantic type, presence/absence of subject)

[39] I say "verbs expressing motion" and not "motion verbs" because this ID tag did not code for the closed class of "motion verbs" in Russian that have both unidirectional and non-directional Imperfective forms. Instead, this is a semantic tag that, although it certainly does include the "motion verbs", is broader in scope.

Now for the breakdown: The three meanings in this cluster are more likely to be encountered in imperative verbs than the meanings of the other cluster ($t = 0.593$), and meanings of this cluster are more strongly associated with verbs expressing mental, psychological, or emotional states ($t = 0.624$). Meanings of this cluster are more frequently found with simple instances of verb negation than the other cluster ($t = 0.616$) and, not unexpectedly, occur more frequently in sentences whose sole purpose is to issue a command ($t = 0.727$). When cluster one meanings occur in dependent clauses, they are more likely to be found in temporal clauses ($t = 0.563$), *esli* 'if' clauses ($t = 0.648$), and purpose (*čtoby* 'so that/in order to') clauses ($t = 0.718$). Ingressive, delimitative, and resultative verbs are more robustly attested in nonfiction ($t = 0.586$) and spoken texts ($t = 0.668$) than are the distributives and intermittent-attenuatives. As for specific collocates, *nado/nužno* 'need to' and *možno* 'may' are more strongly tied to cluster one than to cluster two. Although one may not anticipate associations between clusters and specific prepositions, the *t*-values suggest that *s* 'with/down from', *po* 'along, by', and *za* 'behind/for' are more frequently found with cluster one verbs. However, this effect may be due to the frequency of specific verbs in this cluster (such as the ubiquitous verbs of motion) and should be interpreted with caution. With regard to properties of verbal objects, cluster one has a stronger affinity with animate objects ($t = 0.633$) that are countable ($t = 0.649$). The semantic types of verbal objects vary widely: social events ($t = 0.584$), human qualities or behavior ($t = 0.619$), perceptual objects ($t = 0.672$), even plants ($t = 0.584$). Though no clear picture emerges from the semantic types of the objects, this result is not surprising, given the vast variety of verbs that *po-* can prefix, and that cluster one includes the three most frequent meanings of *po-*. Turning to

properties of the verbal subject, we find that cluster one verbs are more strongly attracted to subjectless impersonal constructions ($t = 0.627$) and are more frequently used when the subject is omitted or understood ($t = 0.695$). The association of subjects from the psychological or emotional realm is stronger with cluster one verbs ($t = 0.702$). Cluster one verbs are also more strongly associated with subjects referring to groups of humans ($t = 0.698$).

Having seen how cluster one differs from cluster two, let us investigate how the members of cluster one – the ingressive, delimitative, and resultative meanings of *po-* – differ from one another. Since the resultative and delimitative amalgamate first, we will turn our attention to the similarities and differences between these two meanings before moving on to discuss the ingressive. Just as the two clusters differ from each other on several levels of analysis, so do the members of cluster one differ in terms of syntactic, semantic, and collocational variables. First, the resultative is more strongly represented in works of nonfiction ($z = 5.397$). In terms of syntax, the resultative exhibits a strong affinity for transitive ($z = 11.440$) and passive ($z = 7.148$) constructions, and displays a stronger association with past participles ($z = 7.542$), while the delimitative prefers intransitive ($z = 1.890$) and infinitival constructions ($z = 4.564$). The resultative is more likely to take genitive, accusative, or dative objects (z-scores ranging from 4.767 to 6.329), whereas prepositional phrases (including those with *čto*-clauses as objects) are overrepresented with the delimitative ($z = 1.964$ and $z = 1.727$, respectively). Resultative verbs are more likely to occur in subjectless constructions ($z = 7.457$), but not those where the "missing" subject is an implied pronoun ($z = -3.193$); to lack any related prepositional phrase ($z = 5.350$); and in the case of participles, to occur with singular

headwords (z = 27.423). Delimitatives show a dispreference for accusative and infinitival objects (z = -2.111 and z = -2.314, respectively) and for subjectless constructions (z = -1.470).

In terms of semantics, the resultative has a greater attraction to verbs that express location or placement (z = 8.717) but not movement (z = -12.518), and to abstract actions (z = 5.192). In both the delimitative and the resultative verbs encoding distinctly human qualities and behavior are overrepresented (z = 3.263 and z = 5.726). Verbs encoding perception are overrepresented in the delimitative (z = 4.112), but the resultative avoids these verbs (z = -3.409). The subjects of resultative verbs have a greater tendency to be inanimate (z = 4.151) and refer to ideas or facts already mentioned in the discourse (z = 4.927) or large abstractions involving humans (governments, institutions; z = 4.117). The delimitative prefers concrete, animate subjects (z = 2.214 and z = 2.932), while the resultative shows relative dispreference for these (z = -3.196 and -4.151). The objects of resultative verbs likewise display a greater affinity for large abstractions involving humans (z = 4.084). Exhortative participles (*davaj* 'let's', z = 1.693) and adverbial phrases indicating the space or time occupied by the action (z-scores ranging from 2.057 to 2.892) are overrepresented in the delimitative. Resultatives avoid the exhortative *davaj* 'let's' (z = -2.82) and are more likely to lack any adverbial complementation whatsoever (z = 6.585).

In the case of resultative participles, the headwords belong to a variety of semantic fields, such as spaces/places (z = 10.606), speech or texts (z = 8.595), social events (z = 6.984), large abstractions involving humans (z = 6.984), facts or ideas (z = 6.034), or sets/groupings of other (non-human) entities (z = 4.092). When delimitative

verbs occur as participles, they prefer inanimate (z = 6.332) concrete (z = 7.197) things (z = 13.689) in the plural (z = 5.816) as their headwords, which contrasts with their preference for animates in non-participial constructions.

From these patterns of over- and underrepresentation, a broad picture[40] emerges: Resultatives can express a wide variety of actions (often involving inanimate actors), and they seem to be the meaning most common in verbs expressing a transfer from subject to object (cf. the overrepresentation of the dative + accusative/genitive complementation pattern, an indicator of the ditransitive construction, with resultatives). Delimitatives seem to prefer a more specific scenario, centering on those actions performed by humans (and, more generally, living beings), and these activities quite often have a cognitive component (though not always), are volitional, and the spatial/temporal circumstances of the action play a more salient role.

Since the ingressive amalgamates early with the (delimitative + resultative) pair, we expect it to share many characteristics with both or either the delimitative and the resultative; z-scores show that this is indeed the case. Of greater interest are the ways in which the ingressive differs from the other two: The ingressive is distinguished by its focus on motion, whether physical or metaphoric. Verbs encoding movement (z = 2.410)[41] have a stronger association with the ingressive than with the delimitative or resultative. Ingressive verbs exhibit a greater affinity for adverbial phrases indicating the

---

[40] As a reminder to the reader, these statements are stylized abstractions from the data and do not imply that all instances of resultative or delimitative verbs must possess all (or any) of these traits. The numbers reflect generalized trends, not specific criteria for membership in the categories of resultative and delimitative verbs. We would expect that verbal predicates not corresponding to these characterizations should also exist.

[41] It is important to note that not all *po* + determinate motion stem were classified as ingressive. Observations in which contextual clues suggest a resultative reading were tagged as resultative, thus reducing the likelihood that these surprisingly large z-scores are a by-product of context-independent tagging.

intended destination of that motion (z = 6.868), and these destinations are often concrete (z = 2.849) spaces/places (z = 2.427). Lack of such clarifying expression is not preferred (z = -4.467). The (inception of) motion can be metaphorically transferred to the domain of mental activity, indicating the beginning of a psychological or emotional state (z = 8.991). Speech or text can also metaphorically begin to "move" (cf. the common Russian phrase *reč' idet* 'the speech goes = the matter is about…, being discussed is…'), indicated by the stronger affinity of the ingressive for subjects coding speech/text (z = 2.763).

4.3.2 Cluster 2: [distributive + intermittent-attenuative]

The top 30 *t*-values for this cluster fall under 17 different ID tags. Similarly to cluster one, the top 30 *t*-values represent variables from several linguistic dimensions. For convenience I group those variables thus:

- variables that pertain directly to the verb (tense-mood, semantic type)
- sentence- and higher-level variables (sentence and text type)
- specific collocates (adverbs, prepositions)
- variables pertaining to the object (animacy, count/mass, semantic type)
- variables pertaining to the subject (semantic type, presence/absence of subject)
- participial headword variables (animacy, number, and semantic type of the participle's headword)

The verbs expressing cluster two meanings (distributive or intermittent-attenuative) are more strongly associated with the past participial construction (t = 0.595) than are cluster one verbs. Cluster two also bears stronger affinity for the indicative present (t = 0.671), no doubt due to the fact that only the intermittent-attenuative can occur in the present tense. The distributive and intermittent-attenuative meanings are more attracted to verbs expressing physically perceptible events than their cluster one counterparts (t = 0.655). In

cases where the cluster two verb is an infinitive governed by another verb, that other verb is more often negated ($t = 0.610$). Sentences containing cluster two verbs are more likely to be declarative or non-imperative exclamations ($t = 0.585$ and $t = 0.610$, respectively), and cluster two is by far overrepresented in works of fiction ($t = 1.016$) – being peripheral members of the semantic network of *po-*, perhaps these meanings are more characteristic of the expressive, creative language of fiction. With regard to specific collocations, cluster two is more associated with adverbial phrases expressing cause/reason, source, and location ($t = 0.616$, $t = 0.885$, and $t = 0.923$, respectively). Curiously, the prepositions *čerez* 'through', *iz* 'out of', and *ot* 'from' show a preference for distributive and attenuative verbs ($t = 0.583$, $t = 0.616$, and $t = 0.659$, respectively). Whereas in cluster one the preference for specific prepositions may be due to the common collocational patterns associated with specific high-frequency verbs, this is not the case for cluster two: cluster two contains only 37 verbs, only three of which occur more than once, and none of which is repeated more than three times. Instead, this preference for *čerez* 'through', *iz* 'out of', and *ot* 'from', all three of which represent physical relationships, is more likely due to the greater affinity this cluster has for physically perceptible events, as already mentioned. As for properties of the verbal objects, cluster two exhibits a stronger affinity for genitive complements ($t = 0.749$) that are inanimate ($t = 0.949$) mass nouns ($t = 0.973$). Verbal objects encoding physiological processes ($t = 0.584$), natural phenomena ($t = 0.584$), and psychological or emotional states ($t = 0.764$) have a stronger association with cluster two. It is interesting that this group of object semantic types, though very diverse, excludes any reference to social interactions among humans (which were associated with cluster one). Three ID tag levels pertaining to

characteristics of the participle headword were also among the top 30 *t*-values for cluster two. However, only four participial phrases belong to cluster two, meaning that the sample size is too small to draw any conclusions[42].

Having established the distinguishing characteristics of cluster two, let us examine the differences between the distributive and the intermittent-attenuative. By comparison to cluster one, the distributive and intermittent-attenuative occur much more rarely and thus yield smaller sample sizes. Thus the *z*-scores for both of these meanings will tend to be smaller than those encountered in cluster one, simply because smaller sample size leads to less certainty about the level of over- or underrepresentation. In the distributive, transitive ($z = 1.784$) and passive constructions ($z = 1.049$) are overrepresented, along with past participles ($z = 1.072$). Greater affinity is shown for plural subjects ($z = 1.977$); singular objects are dispreferred ($z = -1.977$); adverbial phrases describing the manner of action are avoided ($z = -1.038$), as are prepositional phrases generally ($z = -1.565$). The intermittent-attenuative meaning is more strongly attracted to intransitive verbs ($z = 2.742$), prepositional phrases ($z = 2.406$) with singular objects ($z = 3.040$), adverbials describing how the action was accomplished ($z = 1.570$), and singular subjects ($z = 3.615$). These associations suggest that the typical distributive scenario involves a plural subject and/or object involved in a transfer, which is not well-described by supplemental phrases; focus remains on the unmodified action expressed by the verb. Intermittent-attenuatives, on the other hand, involve singular subjects and /or objects, with more attention paid to exactly how the action was performed. With reference to their clustering behavior (see Fig. 4.2-2), the distributive and intermittent-

[42] For the interested reader, the ID tag level preferences and their *t*-values: animate participle headwords ($t = 0.658$), singular headwords ($t = 0.804$), and headwords that refer to animals ($t = 0.918$).

attenuative amalgamate as a cluster later than the members of cluster one, meaning that this cluster is less homogeneous than cluster one.

4.4 Cognitive questions: Prototypicality and category structure

4.4.1 Determining the prototypical meaning

Up until now the semantic structure of *po-* has received only one treatment in the cognitive linguistic literature (Dickey 2007) that not only covers nearly all of the prefix's meanings, but also specifies a prototypical member: the delimitative. In this section I will discuss some major criteria used to determine sense prototypicality as they pertain to *po-*, and I examine my data in light of those criteria. In so doing, I demonstrate that the prototypical meaning of *po-* in modern Russian is not the delimitative, but rather the resultative.

Determining the prototypical sense is a long-standing challenge in cognitive linguistics, and a wide array of criteria for prototypicality has been proposed (Geeraerts 1988:222; Winters 1990; Rice 1996:145-146; Tyler & Evans 2001: Sec. 3.3; Evans 2005: Sec. 2.2.3; Gries 2006). Since not all criteria are particularly applicable to *po-* (or at least do not lend themselves to operationalized definitions within the bounds of this study), I focus on the following:

- point of amalgamation in the HAC analysis
- frequency of occurrence in the corpus
- diachronic primacy[43] of the sense
- shared family resemblances

[43] Although this dissertation is a synchronic study, my data sheds light on some diachronic phenomena outlined by Dickey (2007). As a result I will briefly explain how my findings fit neatly with the trends he observes, without attempting to offer an original diachronic analysis here.

First let us start with the amalgamation pattern in the cluster dendrogram. Divjak & Gries (2006:43) and Gries (2006) note that items that amalgamate early (i.e. at the bottom of the dendrogram) tend to fulfill a number of the requirements of prototypicality: They are often the most frequent senses, are both historically and developmentally primary, and are the combinatorially least constrained with respect to their associated ID tags. In 4.2-2 we saw that the delimitative and resultative have the shortest Canberra distance between them, meaning that they cluster first in the dendrogram, and the remaining three meanings attach to them at further points in the tree – the resultative and delimitative thus form the base of the hierarchy and are prime candidates for the prototypical meaning. However, other criteria must be employed to demonstrate which, if either, is more prototypical.

Frequency of use provides another piece of evidence that the resultative meaning is prototypical. Sense frequency is a strong (although admittedly not failsafe) indicator of prototypicality (Geeraerts 1988; Durkin & Manning 1989; Winters 1990). Frequency was originally seen as a direct correlate of psychological entrenchment (Langacker 1987:59-60; Bybee & Hopper 2001; Dąbrowska 2004:213, 223; Schmid 2000), which led researchers to identify the most frequent, and hence most psychologically entrenched, senses as prototypical. More recent works, however, suggest that the connection between frequency and entrenchment may be less direct (Gilquin 2006, 2007; Schmid forthcoming), and that in any case entrenchment is a difficult phenomenon to measure. These caveats notwithstanding, using sense frequency as an indicator of prototypicality leads to the same conclusion as the other pieces of evidence adduced in this study, and so I feel confident in using frequency data here. In my data the resultative is by far the most

frequent sense: In the 929 observations[44] that serve as the basis for the HAC analysis, 628 (or 67.6%) belong to the resultative meaning. The next most frequent sense is the ingressive, with 151 (or 16.3%) observations, followed by the delimitative (113 observations, or 12.2%), the intermittent-attenuative (26 observations, or 2.8%), and finally the distributive (11 observations, or 1.2%). It is interesting to note that Dickey (2007) favors the delimitative as the prototypical meaning on the basis of frequency as well. Dickey (2007) relies on dictionary data (drawing from Dmitrieva 1991) to determine the frequency of each meaning of *po-*. In dictionaries, however, a rare word will be listed the same number of times as a frequently-used word – namely, once. Dictionaries provide no information about frequency of use and thus can provide little information regarding the potential psychological entrenchment of senses; a rarely-used sense is usually less psychologically entrenched and wields less influence over category structure than a frequently-used sense. Random samples from a corpus address this problem: Rare words are less likely to occur in the sample, while common words are likely to occur frequently, and the disparity in frequency, as stated above, may reflect degrees of entrenchment (or at least give clues about category prototypes). The differences between corpus data and dictionary data can be striking: In Dickey's (2007) / Dmitrieva's (1991) data, delimitative verbs account for 31.8% of all dictionary entries, whereas resultatives account for only 26%. Contrast these figures with the number of delimitative and resultative verbs in my corpus data: In the 1,000 randomly-selected observations, 234 semantically unique verbs occur. Of these, only 23.5% (54 verbs) are delimitative, whereas resultatives occur twice as frequently at 56.4% (132 verbs). And as

[44] Randomly selected from the 16,121 instances of *po*-prefixed verbs originally harvested from the RNC.

we have seen earlier, the resultative occurs even more frequently if we consider not the number of unique verbs bearing resultative meaning, but the number of times the resultative meaning occurs (i.e. even if the same verb is repeated in several observations) – 67.6% for the resultative versus 12.2% for the delimitative, which is a five-fold difference[45]. Both Goldberg (2006:71) and Dąbrowska (2004:25-26, 32, 128) suggest that speakers are sensitive to statistical distributions provided by their linguistic environment, which indicates that speakers form cognitive representations of meaning based on the language they are exposed to (not frequency-blind catalogs of language, such as dictionaries), and thus corpus data can provide a clearer window into the structure of those representations.

Two asides are worth mentioning here: First, it is interesting to note that Isačenko (1962:391–392) states that delimitative *po-* "is so productive that even the most comprehensive dictionaries register only a small fraction of the delimitatives that actually occur" (translated and cited Dickey 2007:330). While this may indeed be true, corpus data do not substantiate the idea that these unrecorded delimitatives make up a large or frequently-encountered number of *po*-prefixed verbs. Leaving aside the data already adduced, we might expect that spoken Russian, often carrying a higher number of colloquial expressions, would demonstrate a stronger presence of the delimitative. But contrary to these expectations, delimitatives seem to occur even *less* frequently in the

---

[45] The percentages obtained from corpus data for the other meanings of *po-* also differ from Dickey's (2007) / Dmitrieva's (1991) figures, but less dramatically so. The following numbers consider the number of unique verbs, not observations: Corpus data: ingressive – 9% (21of 234); distributive – 4.3% (10 of 234); intermittent-attenuative – 9% (21 of 234); attenuative – no data available; and the ambiguous cases: resultative-delimitative – 4.7% (11 of 234); resultative-distributive – 0.4% (1 of 234); resultative-ingressive – 2.6% (6 of 234). The percentages total to slightly over 100% because some verbs occurred in more than one sense (*podumat'*[p] for instance, can be either resultative or delimitative, yet still means 'think'). Dickey's (2007) / Dmitrieva's (1991) figures: ingressive – 2.1%; distributive – 19%; attenuative – 11.9%; intermittent-attenuative – 11.9%.

spoken subcorpus[46] of the RNC than in the corpus as a whole: In the set of 929 randomly-selected observations, only 107 observations come from spoken "texts" (conversations recorded on the streets, scripted or semi-scripted radio or TV programs). Of these 107 observations, only 11 (10.3%) contained delimitative verbs, while 64 (59.8%) contained resultative verbs[47]. While this does not rule out the hypothesis that many new delimitatives are created colloquially on-the-fly, it does suggest that these spontaneous creations are not frequent.

Secondly, it is curious that Dmitrieva's (1991) account of the rise in the number of delimitatives versus resultatives over the historical development of *po-* does not consider the differences in size or scope between the two reference works surveyed: Sreznevskij's (1958) *Materialy dlja slovarja drevnerusskogo jazyka po pis'mennym pamjatnikam* [Materials for a Dictionary of Old Russian based on Written Texts], the source for Dmitrieva's delimitative:resultative ratio in Old Russian, only occupies three volumes. The source for her delimitative:resultative ratio in modern Russian is the *Slovar' sovremennogo russkogo literaturnogo jazyka* [Dictionary of Contemporary Literary Russian] (1950–65), which is roughly five times as large at 17 volumes. The larger, modern dictionary likely contains far more rare, infrequent, or specialized words

---

[46] As two reviewers point out, sometimes tone of voice or other extra-linguistic indicators, especially those associated with emotive or expressive utterances, can alter the meaning of an utterance so that the intended message cannot be directly inferred from the words alone. Unfortunately the RNC does not encode such information, and thus we are left to work with the bare transcriptions of spoken "texts". Tone of voice can often be inferred by referencing the discourse context of the utterance (that is, beyond the target sentence), but unfortunately the number of tags such an analysis would require exceeds the resources available for this study. As corpus research in Russian continues to develop, perhaps more auditory information will be included in future corpora. Although this limitation does not automatically negate the findings of a study based on spoken "texts" (after all, only a minority of utterances contain hidden or ironic meaning), the reader should remain aware of this fact.

[47] Figures for the remaining records and meanings attested in the spoken texts of the random, 1000-observation sample: ingressive – 21 (19.6%); intermittent-attenuative – 1 (0.9%); the remaining 10 (9.3%) were ambiguous.

than the smaller reference work, thus creating difficulties when comparing the two. One must also consider the different cultural contexts of writing reflected in these two dictionaries: If the cultural scope of writing reflected in the dictionary of Old Russian is narrower than that in the modern dictionary, then there exists another variable that complicates any comparison based on these two works. Nevertheless I do agree that the delimitative meaning has spread to more verbs in the modern period – Dmitrieva (1991) and Sigalov (1975) present convincing evidence that such an increase did occur, given that some originally resultative verbs now express only delimitative meaning. Furthermore, Dickey (2007) narrates a cognitively motivated path by which the semantics of *po-* has shifted over the past several centuries, making an expansion of the delimitative plausible. But neither of these facts necessarily entails that the delimitative has superseded the resultative as the prototypical sense.

Returning to the issue of sense prototypicality, there is a third reason why the resultative should be considered prototypical: diachronic primacy. Tyler & Evans (2002) note that in their study of over 20 English prepositions, the earliest attested sense(s) are still prototypical in modern English; Gries (2006) comes to a similar conclusion regarding English *run*; Tyler & Evans (2001) posit relative time of attestation as a criterion for determining the centrality (and hence prototypicality) of a word sense. Citing Němec (1953) and Shull (2000), Dickey (2007) notes that the prefix *po-* likely had three original, spatial meanings (paralleling the semantics of the preposition): ablativity, locativity, and allativity, or in cognitive linguistic terms, SOURCE, PATH, and GOAL. Early on the SOURCE meaning, present in verbs like *pojti*^{p} 'go, set out' was fairly weak, and these verbs acquired a more ingressive flavor, profiling the inception of the event in time

(Dickey 2007:335). Two spatial meanings remained, GOAL (Old Russian *postignuti*^{p} 'reach') and PATH (Old Russian *poplavati*^{p} 'roam/wander [an area]'). According to Dickey,

> [B]y historical times, PATH had given rise to a productive SURFACE-CONTACT meaning that was very resultative in nature, cf., e.g., ORus *posmoliti* 'cover with resin'. It should probably be assumed that in the original situation PATH and GOAL were distinct local prototypes in the network of *po-*. However, the PATH/SURFACE-CONTACT meaning was a very telic one, i.e., it tended to profile the complete affectedness of the surface in question, so that *poplavati* meant 'roam all over [an area]', and *posmoliti* meant 'cover [completely] with resin'. Thus, PATH/SURFACE-CONTACT could produce resultative verbs by metaphor on a par with the GOAL configuration. (2007:336)

Dickey (2007:336) also notes that the PATH/SURFACE-CONTACT meaning produced a large class of resultative verbs by metaphorical extension, and that this was the dominant meaning of *po-* in Old Russian; as the diachronically primary sense, the resultative is also likely to be prototypical in modern Russian as well.

The final piece of evidence suggesting that the resultative is the prototypical meaning of *po-* has to do with family resemblances, or attributes shared by members of the category (see 2.3.3 for review). Rosch & Mervis (1975:598-599) state that "the most prototypical members of…categories are those which bear the greatest family resemblance to other members of their own category" (see also Croft & Cruse 2004:78, 81; Janda 2010b; Divjak 2010:168-180). In this regard the ambiguous cases mentioned in 3.3.3 provide insight into the shared family resemblances among the meanings of *po-*: Of the 71 observations that I and my educated native speaker consultant could not confidently assign to one meaning or another, nearly all (70)[48] were ambiguous between

[48] The remaining case is an instance of ambiguity between the attenuative or delimitative meaning, which is not unexpected given the similarity between these two meanings (Isačenko 1960:238-239; Zaliznjak & Šmelev 2000:120).

the resultative and some other meaning (delimitative - 57; ingressive - 12; distributive - 1). As Dickey (2007) and Anstatt (2002) note, delimitative verbs can take on resultative meaning in certain contexts, and it is well-known that ingressive verbs often "lose" their ingressive focus and acquire purely resultative meaning (Isačenko 1960:23). These cases of ambiguity are only possible because these three meanings – delimitative, distributive, and ingressive – bear a strong similarity (i.e. family resemblance) to the resultative. The fact that there are no cases of delimitative–distributive, delimitative–ingressive, or ingressive–distributive ambiguity in the data indicates that these three meanings have less in common with each other than each of them does with the resultative. And as the reader will see in 4.4.2, the conceptual links between the resultative and the delimitative, distributive, and ingressive meanings are easily motivated by simple metonymies and metaphors. These ambiguous cases and the metonymic/metaphoric relations detailed in the next section (4.4.2) point toward the resultative as the meaning that "bear[s] the greatest family resemblance to other members of … [the] category" (Rosch & Mervis 1975:598-599). All the meanings of *po-* represent modifications of a journey along the metaphorical "path" expressed by the base verb; since the resultative expresses that path in its most basic, unmodified form, it is no surprise that the resultative bears more resemblance to all the other members of the category than does any other one meaning.

While none of these four pieces of evidence – amalgamation order[49] in the cluster analysis, the sheer frequency of the resultative in the corpus data, the primacy of the resultative in Old Russian, the strong family resemblances between the resultative and the

[49] While amalgamation order does not differentiate between the resultative and delimitative, it does narrow down the choices for strongest candidate for the prototypical meaning to only the resultative and delimitative. Additional criteria distinguish the resultative as more prototypical than the delimitative.

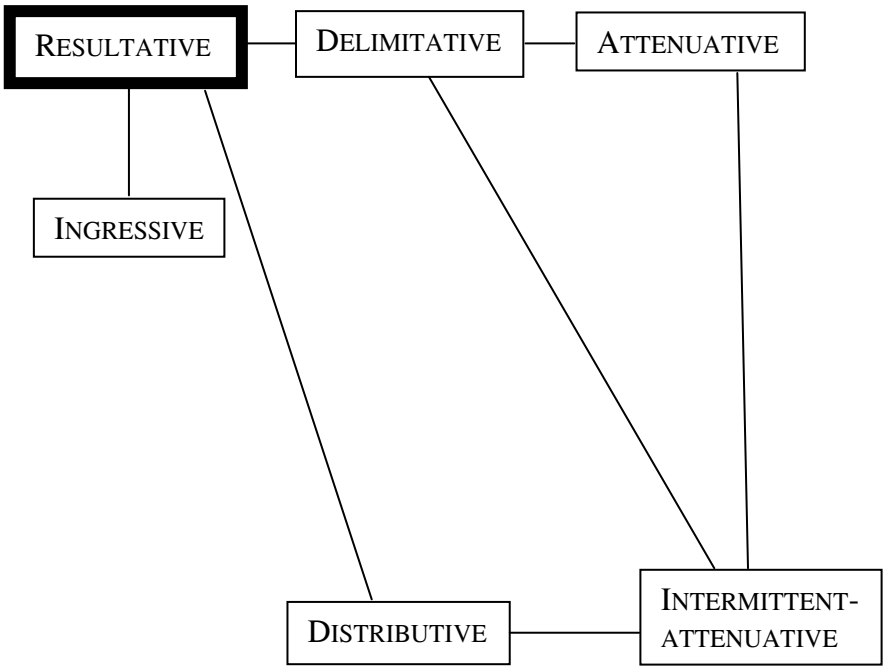
other meanings of *po-* – may be sufficient when considered separately, taken together they provide clear indication that the strongest candidate for prototypicality in the modern semantic network of *po-* is (still) the resultative.

4.4.2 Category structure: Image schemas + metaphoric and metonymic links

In this section I present the semantic structure of *po-* as a radial network of meanings, integrating new information obtained from the cluster analyses with the known semantic characteristics of each meaning of *po-*. I briefly recap the image-schematic representations of each meaning given in 2.3.4, and I discuss the cognitive links that connect these meanings in a sensible fashion. I offer a semantic motivation for the organization of the meanings into clusters, namely that cluster one meanings are metonymic extensions of the resultative, while cluster two meanings are multiplied variants of the cluster one meanings.

Based on the cluster analyses of 4.2, the semantic structure of *po-* can be represented as a radial network (2.2.3) of interconnected meanings:

Fig. 4.4.2-1

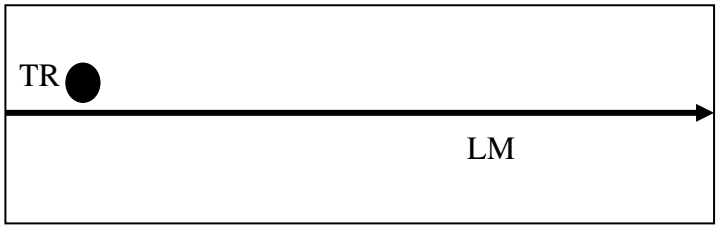


The resultative is highlighted by a thick-bordered box to signal that it is the prototypical meaning. Cluster one meanings (plus the attenuative meaning) are in the top portion of the diagram, and cluster two meanings are grouped together in the bottom. The lines connecting each meaning indicate (mostly metonymic) extension relations among the meanings. The relative distance between any two meanings is intended to suggest the degree of (dis)similarity between them, while not representing a precise scalar interpretation of the cluster dendrograms. Cluster one meanings are grouped more tightly together than cluster two meanings, since the cluster analyses show that cluster one amalgamates earlier than cluster two. The delimitative and resultative are most similar to one another (they amalgamate the earliest), with the ingressive ranking next in similarity, hence their relative distances from each other in the network. Since the attenuative was not represented in the random sample of 1000 observations, its location in the network is

based on Isačenko's (1960) and Zaliznjak & Šmelev's (2000) analyses of the attenuative as closely connected to the delimitative. In the cluster dendrograms the distributive and intermittent-attenuative consistently cluster together later than the cluster one meanings, so the distance between them is relatively greater than the distance separating cluster one meanings, and their location far from cluster one represents the distance/difference between the two clusters.

Each of these six meanings can be represented image-schematically. The resultative meaning implies that the action has been brought to its natural completion and is expressed in the Natural Perfective (Janda 2007b, 2008b) – that is, the so-called "empty" use of *po-*:

Fig. 4.4.2-2



The TR is the subject of the verb, the LM represents the canonical course of action encoded by the base verb, and the path marks a complete trajectory from inception of the action to termination, beyond which the action cannot naturally proceed. In Vendlerian (1957) terms, the resultative meaning combines with predicates expressing achievements or accomplishments. Prefixed Perfectives of the resultative meaning are related to their Imperfective base verbs by means of a contact metonymy (Peirsman & Geeraerts 2006): The resultative signals the culmination of an unbounded Imperfective action and is thus

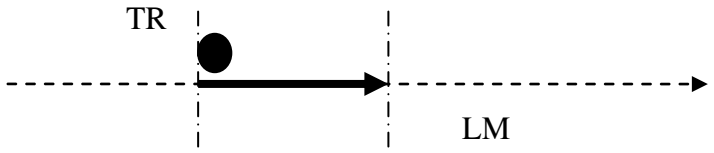
temporally contiguous with the action encoded by the Imperfective; the unprefixed Imperfective and the Perfective resultative share a temporal boundary (Janda 2008a).

Moving away from the prototypical meaning (the resultative), the remaining five meanings of *po-* belong to the class of Complex Acts (Janda 2007b, 2008a), in which "the prefix performs a more quantitative role, usually expressing a temporal limit on the action" (Janda & Nesset forthcoming). Each meaning is a metonymic extension from the resultative (or a metaphoric extension of another extension, as the attenuative is an extension of the delimitative) that expresses a unique TR-LM-path relationship compatible with the semantics of the base verb. The five Complex Act meanings of *po-* are related to their (usually Imperfective) base verbs by a PART-WHOLE metonymy (Peirsman & Geeraerts 2006) whereby the prefixed Perfective "describes a bounded portion of an unbounded Imperfective activity" (Janda 2008a). Like the resultative, these five meanings do not introduce a new path configuration to the base verb, but instead they modify the metaphoric paths expressed by the verb (Nesset forthcoming). And unlike the semantics of many other polysemous Russian prefixes which involve metaphoric/metonymic extensions from a spatial sense, the non-prototypical meanings of *po-* are all extensions from a temporal sense, the resultative.

First I examine the resultative's co-cluster members: the delimitative, attenuative, and ingressive. The delimitative can be defined as RELATIVE DELIMITATION (Dickey 2006). Unlike in the resultative, the prefix in the delimitative meaning profiles only a portion of the trajectory (that is, focus shifts from the whole to a part of the whole) and as such the delimitative represents a metonymic PART-WHOLE extension from the resultative. In the delimitative meaning *po-* does not make reference to inherent endpoints, and so it

is the perfectivizing prefix *par excellence* for atelic activity predicates (Vendler 1957; Janda 2007b) and for normally telic predicates that have been re-construed as atelic via quantification of the verbal argument (Mehlig 1996, 2004). Since the actions encoded by verbs in these predicates have no inherent endpoints, the delimitative meaning is completely compatible with them.

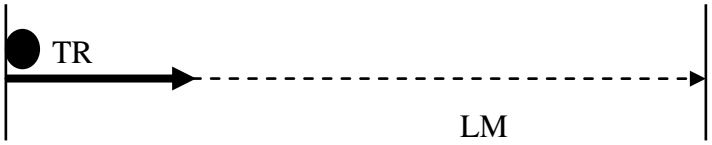
Fig. 4.4.2-3



Here again the TR is the subject of the verb, and the LM is the canonical course of action implied by the base verb. The path here is potentially limitless, but *po-* delimits a portion of the action relative to that trajectory.

Although the attenuative meaning does not occur in my data, Isačenko's (1960) analysis suggests that it can be schematicized similarly to the delimitative. With the attenuative, however, the meaning of RELATIVE DELIMITATION (see 2.3.4) is metaphorically transferred from the domain of TIME to the domain of INTENSITY, where the path no longer indicates a course through time but rather marks an imagined scale of potential intensity for the action. Instead of a full-strength performance, a *po*-prefixed verb in the attenuative meaning indicates that the action is performed at only a fraction of the usual intensity, covering only a portion of the potential path or scale of intensity. In this way the attenuative represents a metaphoric extension from the delimitative, via the metaphor TIME SPENT DOING AN ACTION IS THE INTENSITY OF THE ACTION.

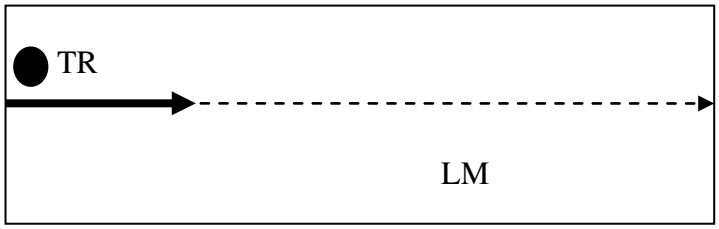
Fig. 4.4.2-4



What is relevant is that only a portion of the intensity scale is realized during the performance of the action, yielding the frequent translation 'do X slightly/a little/incrementally'.

The ingressive meaning stems from a meaning of PATH/PARTIAL-TRAJECTORY (Dickey 2007), in which the path and trajectory are understood metaphorically as aspects of an action's progression through time (via the TIME IS SPACE metaphor). Like the delimitative, the ingressive is a metonymic extension from the resultative meaning: Instead of focusing on the completion of the entire path, the ingressive denotes completion of only the first portion of the path (another PART-WHOLE relationship). We could represent this meaning image schematically like this:

Fig. 4.4.2-5



Here the TR is the subject of the *po*-prefixed verb, and the LM is the canonical course of action implied by the base verb (Shull 2003:152-153). The dashed horizontal line represents the full trajectory of the action encoded by the base verb, and the darkened arrow represents some subset of that trajectory – here *po*- only stipulates that some initial portion of the trajectory has been covered, and is non-committal regarding whether the

TR reaches the end of the trajectory. This schematic "lack of commitment" to covering the full possible trajectory accounts for the ambiguity of sentences like *Ivan pošel*^{p} *v bar*: The sentence can be interpreted as "Ivan set out for the bar", indicating that the speaker knows Ivan has left and nothing more, or the sentence can be interpreted as an assertion that Ivan indeed reached his destination – this latter interpretation is an instance of another metonymic relationship whereby AN INITIAL SUBEVENT (setting out for the bar) STANDS FOR A COMPLEX EVENT (setting out for, traveling to, and arriving at the bar). The sentential or discourse context usually disambiguates between the two interpretations.

The close relationships between the resultative and delimitative and between the resultative and ingressive are further substantiated by the ambiguous cases in the data. Of the 1000 observations (randomly-selected from the 16,121 collected from the RNC), 58 lacked sufficient context to disambiguate between the resultative or delimitative meanings, and 12 lacked enough context to disambiguate between the resultative and ingressive[50]. In the case of resultative-delimitative ambiguity, the majority of cases involved the verb *posmotret'*^{p} 'look, watch', which is used in both purely delimitative and resultative senses:

> *Pozže* ***posmotrel***^{p} *"Volosy" i ponjal, čto mne delat' v žizni.*
> 'Later I **watched** "Hair" and I understood what I was to do in life.'
> (Resultative meaning: The subject watched the film from start to finish.)

> *Ja* ***posmotrel***^{p} *na tebja minut desjat' i rešil – poprobuem.*
> 'I **looked** at you for about ten minutes and decided we'd try it.'
> (Delimitative meaning: The subject looked at the patient for an arbitrarily delimited period of time – he could have continued looking longer, since there is no necessary terminus to the action here.)

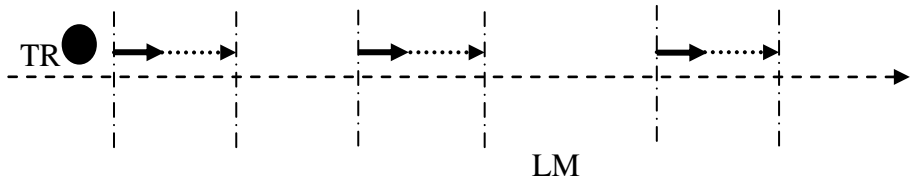
---

[50] The two remaining cases of ambiguity were between a resultative and a distributive reading of *porezat'*^{p} 'cut' and between an attenuative or delimitative reading of *podvigat'*^{p} 'move'.

Like many Russian verbs (Janda 2010b), *posmotret'*^{p} 'look, watch' permits more than one construal with regards to Completability, which in turn makes this verb compatible with more than one meaning of *po-*. In the above examples the distinction between a Completable action (like watching a movie, typically in its entirety; see Janda 2007) and a Non-completable action (like looking at someone) is pivotal to the choice of resultative or delimitative meaning, respectively. It is noteworthy that there were no cases of delimitative-ingressive ambiguity, and for this reason we can assume that those meanings are related primarily by their metonymic connections to the resultative.

The two meanings of cluster two [distributive + intermittent-attenuative] are related to the meanings of cluster one via another sort of metonymic extension whereby A SINGLE EVENT IS THE SOURCE FOR A COLLECTION OF SIMILAR EVENTS (see Kövecses 2002; Peirsman & Geeraerts 2006) – while the cluster one meanings signal the completion or delimitation of single events, those meanings are extended to refer to groups of similar completed or delimited events in cluster two. Simply stated, the intermittent-attenuative ('do X slightly, from time to time') generally refers to multiple instances of an action that would otherwise be encoded by a delimitative or attenuative verb. Of the 20 intermittent-attenuative verbs in the random sample, 17 also have related *po*-prefixed forms (sharing the same root) that express delimitative meaning. In addition to the obvious semantic characteristics shared by the intermittent-attenuative, delimitative, and attenuative, the existence of related delimitatives for the majority of intermittent-attenuative verbs in this study suggests that the intermittent-attenuative is connected to the delimitative (and, by extension, to the attenuative), as shown in the network (Fig. 4.4.2-1). The intermittent-attenuative can be represented image-schematically thus:

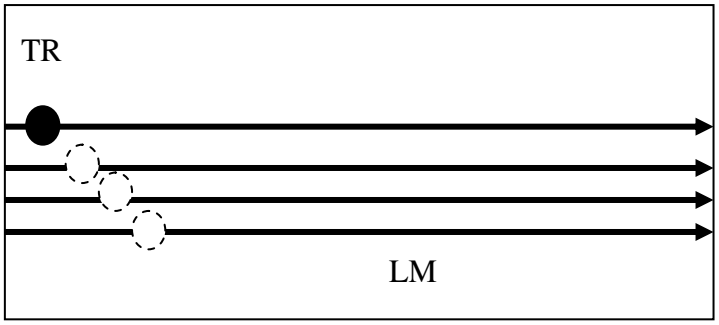
Fig. 4.4.2-6



In short, the intermittent-attenuative meaning denotes an action reduplicated along a temporal trajectory – an action is performed at less-than-full intensity and/or at irregularly spaced, relatively delimited intervals along a timeline. Once again the TR is the subject of the verb and the LM is the canonical course of action represented by the base verb. *Po*-prefixed verbs of the intermittent-attenuative meaning are typically atelic, so the metaphoric trajectory here is unbounded. Because the action encoded by the verb has no natural terminus, any period of performance is thus relatively delimited.

Just as the intermittent-attenuative indicates multiple performances of a delimitative/attenuative action, so the distributive is a sort of multiplied resultative, involving a number of completed actions carried out on a number of objects or by a multiplicity of subjects – the distributive is an extension from the resultative via the metonymy A SINGLE EVENT IS THE SOURCE FOR A COLLECTION OF SIMILAR EVENTS. Contra Dickey (2007), the corpus data do not support a connection between the ingressive and the distributive (see the dendrograms in 4.4.1). Instead my treatment resembles that of Isačenko (1960:288) and Zaliznjak & Šmelev (2000). The distributive meaning can be schematicized similarly to the resultative, only multiplied:

Fig. 4.4.2-7



The TR is the subject of the verb, and the LM is the canonical course of action encoded by the base verb. The TR can be plural or singular; the schema is non-committal regarding the subject's number. What is significant is that the action is performed multiple times (whether by multiple subjects or upon multiple objects by a single subject), and that the TR completes the possible trajectory from beginning to end. In the radial network (Fig. 4.2.2-1) the distributive is connected to both the resultative (based on known semantic similarities) and the intermittent-attenuative (based on distributional similarity, which reflects the shared semantic characteristic of multiplication).

At this point the reader understands how the HAC analysis partitions the meanings of *po-* into two groups, [(resultative + delimitative) + ingressive] and [distributive + intermittent-attenuative], based on the behavioral profile of each meaning. The reader has also seen how the behavioral profile of each meaning of *po-* can be further investigated by means of *t*-values and *z*-scores to uncover the distinguishing characteristics of each meaning. This information can then be used to motivate a radial network representation of the six senses of *po-*, which does much to address the long-standing difficulties in understanding the semantic structure of this prefix. In Chapter 5 I summarize these findings, highlight the contributions of this dissertation, and I point out

how future research on verbs of Types II, III, IV, and V (which are not investigated in this study) may shed further light on the historical development and synchronic structure of the semantics of *po-*.

5 Conclusions and Directions for Future Research

In this chapter I summarize the contributions of this dissertation (5.1), namely that by using the conceptual tools of cognitive linguistics combined with corpus linguistic methods, I have produced an empirically-based, psychologically plausible account of the semantic structure of the Russian verbal prefix *po-*. In so doing I have addressed the long-standing problems surrounding the relationships among the meanings of *po-,* and I have uncovered evidence that the resultative meaning is prototypical. After reviewing these findings and highlighting the contributions of this dissertation in 5.1, I point out directions for future research in 5.2.

5.1 Contributions to the study of *po-* and prefixal semantics in Russian verbs

As the reader saw in Chapter 2, the semantics of *po-* poses a number of theoretical difficulties: The first problem – determining what the meanings of *po-* actually are – appears to have already been resolved, while the second problem – determining how those meanings are related – finds an empirically-adduced solution in this dissertation. In contrast to the atomist period, when the catalog of meanings attributable to *po-* seemed to change with every successive reference work, today there is broad scholarly consensus (Isačenko 1960; Guiraud-Weber 1995; Zaliznjak & Šmelev 2000; see also Dickey 2007 and Nesset forthcoming) that *po-* can express five meanings on its own (attenuative, delimitative, distributive, ingressive, and resultative) and one additional meaning (the

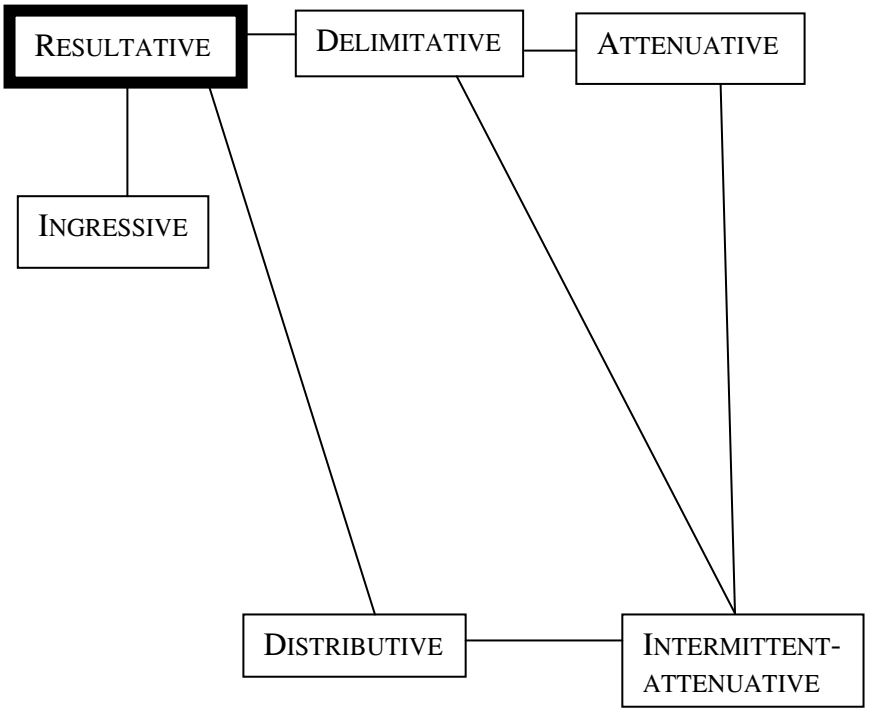
intermittent-attenuative) when it occurs in the prefix-suffix combination *po-…-yva-*. While *po-* does modify the meaning of the verb it prefixes, the modification is usually not the sort that leads to the derivation of a denotationally identical Imperfective. In the resultative meaning *po-* forms the Natural Perfective (Janda 2007b, 2008) of many verbs; this is the use often described as "empty prefixation" or as the "aspectual partner" of the related simplex Imperfective. The other five constitute Complex Act Perfectives (Janda 2007b, 2008) which describe not just the completion of an action, but give additional information about the temporal properties of that action.

The second problem – how the six meanings of *po-* are related – is the focus of this dissertation. Structuralists (Flier 1975, 1984; Gallant 1979; van Schooneveld 1978) first tackled the problem of Russian verbal prefix semantics by positing a single invariant meaning for each prefix, described in terms of the presence or absence of binary features; the various senses of a given prefix were seen as contextually derived from that invariant. However, the invariant meaning often had to be highly abstract in order to account for widely disparate senses of the same prefix, and as a result it became difficult to see exactly how these abstractions could account for the meanings in context. Cognitive linguists continued the search for an adequate theoretical account of prefixal semantics in Russian with much success (Rudzka-Ostyn 1983a; Janda 1985, 1986; Tchizmarova 2006; Janda & Nesset forthcoming; Nesset forthcoming). But while the diachronic development of *po-* has received some attention (Dickey 2007), a full network account of its semantic structure in synchronic terms was lacking.

Just such a synchronic account of *po-*'s semantic structure is the most immediate contribution of this dissertation. Using cognitive linguistic theory to integrate existing

knowledge about *po-* with statistical analysis of data from the Russian National Corpus, I discovered that the relationships among the senses of *po-* can be represented thus (see also Figs. 4.1-2 and 4.4.2-1):

Fig. 5.1-1



As this radial network indicates, the meanings of *po-* can be grouped into two clusters, the first consisting of the resultative, delimitative, attenuative, and ingressive, and the second consisting of the distributive and intermittent-attenuative. The first cluster contains the more central meanings of the prefix, clustered around the resultative as the prototype, while the second cluster contains the more peripheral meanings. The five non-prototypical meanings are metonymic (and in one case metaphoric) extensions of the resultative meaning (4.4.2). The resultative can be understood as the full traversal of the metaphoric "path" expressed by the base verb. Taken together, four criteria produce converging evidence that the resultative is the strongest candidate for the prototypical

meaning (4.4.1): The resultative is one of the first two meanings to amalgamate in the HAC analysis; the resultative is the most frequent meaning attested in the data (both in verb types and tokens); the resultative has strong family resemblances (i.e. metonymic and metaphoric links) to the other meanings in the category; and the resultative was the diachronically dominant meaning. With respect to the other meanings, the ingressive profiles the inceptive phase of that metaphoric path (thus related to the resultative by a PART-WHOLE metonymy), while the delimitative profiles a relatively-delimited portion of a potentially boundless path (related to the resultative by another PART-WHOLE metonymy). The attenuative denotes a metaphoric transfer of relative delimitation from the domain of TIME to the domain of INTENSITY. The two peripheral members of the category, the distributive and intermittent-attenuative meanings, are connected to the more central meanings by the metonymy A SINGLE EVENT IS THE SOURCE FOR A COLLECTION OF SIMILAR EVENTS.

In more general terms this dissertation shows that corpus-based studies can be just as fruitful in the investigation of the semantics of morphemes as they are in the study of independent lexemes (Gries 2006; Gries & Divjak 2008; Janda & Solovyev 2009) and abstract grammatical constructions (Stefanowitsch & Gries 2003), which is not surprising in light of the cognitive linguistic assumption that grammar and lexicon are not separate levels but rather opposite ends of a continuum. The results of this study of *po-*'s semantic structure are also empirically falsifiable; researchers using other data sets and similar procedures can either corroborate or refute the results presented here. As such this dissertation constitutes a strongly empirical approach to prefixal semantics (see Janda & Nesset forthcoming for another) and is part of the growing trend toward data-driven

analysis in cognitive linguistics. The need for data-driven, bottom-up analysis is exemplified by the fact that this dissertation structures the meanings of *po-* differently from another analysis not based on the same breadth of data (see Dickey 2007). Finally, this dissertation joins several other recent works (Janda 2008a, 2010b, and forthcoming; Nesset 2009) that illustrate the role of metonymy (see 2.3.4 and 4.4.2) in grammatical phenomena, in contrast to the earlier focus on metonymic relations among lexical items.

5.2 Directions for future research

5.2.1 Prefixal semantics: Further work

The most obvious extensions of the study presented in this dissertation involve the semantics of Russian verbal prefixes. With respect to *po-*, further corroboration of the resultative's prototypical status could be garnered by additional corpus work. Collecting equally-sized samples of each meaning of *po-* could allow the researcher to discover which meaning is combinatorially least constrained, which is yet another criterion for prototypicality (Divjak & Gries 2006; Gries 2006; Gries & Divjak 2008). Sample sizes for each of the six meanings of *po-* in this dissertation varied too widely to permit further investigation into combinatorial restrictions. Given the rarity of some meanings (the distributive and intermittent-attenuative) in the RNC, additional data would have to be mined judiciously from the Internet.

The results of this dissertation indicate that a corpus-based investigation of other Russian verbal prefixes could add to the growing body of cognitive linguistic knowledge on the subject (see Dickey 2006, 2007; Janda 2007b & 2010b; Dickey & Janda 2009;

Janda & Nesset forthcoming; Nesset forthcoming). In particular, the study of Aktionsarten formed by multiple prefixes could benefit from corpus-based investigations similar to the one presented in this dissertation. Section 4.3 shows how the results of an HAC analysis can be further analyzed by *t*-values and *z*-scores to tease apart differences between the various Aktionsart meanings expressed by *po-*. As Krongauz (1998:128) notes, some Aktionsarten are formed by multiple prefixes, and the semantic differences introduced by these prefixes are often obscured in the interest of maintaining the unity of each Aktionsart category. An HAC analysis, supplemented by the use of *t*-values and *z*-scores, of multiple-prefix Atkionsarten could highlight those fine differences in meaning that Krongauz (1998) fears are lost in our current understanding of Aktionsart in Russian.

5.2.2. Verbs of Type II, III, IV, and V: Diachronic & synchronic perspectives

The reader may recall from 3.3.1 that five different morphological types of verbs were uncovered during the data collection procedure, and that this dissertation only investigates Type I – those verbs where the semantic and grammatical contribution of *po-* can be determined with the greatest certainty. There remain four other verb types in which *po-*'s contribution, if any, is much less clear. I now turn my attention to these four sets of verbs, and I would like to speculate about the role of *po-* in each. I restrict my remarks to the realm of the hypothetical because only an extensive diachronic survey (or in some cases psycholinguistic research) could offer convincing proof of the relationships among the remaining verb types. Such additional research is beyond the scope of this dissertation; however, the possible links are interesting enough to warrant an outline here.

This section serves to point out intriguing directions for future research with these groups of verbs.

During the original data collection procedure, exactly 16,121instances of *po*-prefixed verbs were collected from the manually-annotated portion of the Russian National Corpus, and the verbs represented in those 16,121 observations were categorized into five verb types (3.3.1) according to their morphological behavior. Only Type I verbs (usually consisting of an unprefixed Imperfective and a *po*-prefixed Perfective) were used to generate the randomly-selected set of observations that are the basis of the HAC analysis in 4.2, primarily because only in Type I verbs is the contribution of *po*- clear. However, while the majority of *verbs* belong to Type I (75.2%, or 533 of 709), Type I actually represents a minority of the *observations* (38.2%, or 6,152 of 16,121). Type II verbs, on the other hand, represent the majority of observations (57.3%, or 9,242) while containing a minority of verbs (119 of 709, or 16.8%). In Type II verbs *po*- is found both in the Imperfective and the Perfective forms (ex: *polučit'*[p] / *polučat'*[i] 'receive'), and thus the semantic contribution of *po*-, if any, is more difficult to determine. I would like to hypothesize, however, that the meaning of *po*- in Type II verbs is perhaps a "bleached out" version of the resultative – if *po*- has any meaning at all in Type II verbs, then that meaning has transitioned from an original 'do X to completion' to simply 'do X'. The presence of *po*- in Type II Imperfectives would then be the result of a re-construal of an original resultative meaning, which could have followed at least one of two closely-related courses of development: creation of a prefixed derived Imperfective to preserve a secondarily acquired meaning of the prefixed Perfective, or a

metonymic re-construal of the focus of the resultative from successful completion of the action to simple performance of the action without reference to completion.

Let us first turn our attention to verbs like *po|stavit'* 'put, place' (Type I) and *postavit'*^{p} / *postavljat'*^{i} 'supply, provide' (Type II). In addition to its concrete, spatial sense, *postavit'*^{p} 'put, place' has acquired a metaphorical extension in some discourse contexts, namely 'supply, provide'. This new, metaphorical sense has led to the development of a derived Imperfective, much like the derived Imperfectives that are formed to preserve the lexical content introduced by prefixes in other verbs: *u- + verit'*^{i} 'believe' > *uverit'*^{p} 'assure' > *uverjat'*^{i} 'assure'. In the case of *postavit'*^{p} / *postavljat'*^{i}, however, the derived Imperfective does not preserve any new meaning added by the prefix *po-*; rather, the prefixed Imperfective *postavljat'*^{i} signals the new, metaphorically-derived meaning 'supply, provide' of *postavit'*^{p}, which is not generally associated with its unprefixed counterpart *stavit'*^{i}. The prefix *po-* remains in the newly derived Imperfective most likely because *po-* has undergone a sort of semantic bleaching, whereby the resultative meaning ("do X to completion") has weakened and has been re-construed as something akin to "do X". Otherwise, the (resultative) meaning of *po-* would be completely incompatible with the meaning of the Imperfective (cf. Nesset 2007; Janda & Nesset forthcoming). Historically this bleaching phenomenon is not new; see Dickey (2007:12-15, 26) for discussion on the bleaching out of *po-*'s spatial ablative meaning in conjunction with the stem *idti*^{i} 'go' to yield the modern ingressive meaning. In our case, the presence of *po-* in the Imperfective no longer signifies completion of the action, but instead *po-* serves simply to distinguish the 'supply, provide' meaning from 'place, put'; it has become a place-marker of other semantic content, while not expressing that content

itself. In terms of verb types (3.3.1), *postavit'*^{p} belongs to one of two types, depending on which sense is intended: The generalized *po|stavit'* 'put, place' falls under Type I, while the metaphoric *postavit'*^{p} / *postavljat'*^{i} 'supply, provide' is clearly Type II.

A similar course of development can be seen in Type III verbs, which consist of aspectual trios: an unprefixed Imperfective, a *po*-prefixed Perfective, and a derived *po*-prefixed Imperfective (ex: *slat'*^{i} 'send'> *poslat'*^{p} > *posylat'*^{i}). As in Type II, it is difficult to ascertain the contribution of *po-* in Type III verbs. There is some speculation that the derived Imperfective of a Type III verb is more metaphorical than the simplex Imperfective (Veyrenc 1980:159-179), but the data are so far inconclusive. The point of interest here is that in Type III verbs, it appears that the resultative meaning of *po-* is so closely identified with the successful performance of the action (see notes on "empty prefixation" in 2.2.1) that the prefix has been re-interpreted as part of the base verb itself, and so the prefix remains in the derived Imperfective (which is usually formed via suffixation). Indirect evidence of this process lies in the fact that in most of the 30 Type III verbs collected from the RNC, the presence of resultative meaning (as opposed to any other meaning of *po-*) is the primary difference between the *po*-prefixed Perfective and its simplex Imperfective (questions of style or concreteness/metaphoricity aside). It would be interesting to investigate whether a similar process may have been at work in the development of *podat'*^{p} / *podavat'*^{i} 'give (into someone's hand), serve', which is based on the more generalized pair *dat'*^{p} / *davat'*^{i} 'give'. But here the *po*-prefixed Perfective is not the aspectual "partner" of the simplex Imperfective; *podat'*^{p} / *podavat'*^{i} 'give (into someone's hand), serve' and *dat'*^{p} / *davat'*^{i} 'give' seem to occupy semantically different spheres.

Types IV and V present opportunity for further historical investigation. In the three verbs of Type V (*kupit'*^{p} / *pokupat'*^{i} 'buy', *kupit'sja*^{p} / *pokupat'sja*^{i} 'be bought, *po|nukat'*^{i} 'urge on'), the existence and role of the *po*-prefixed form can best be explained diachronically. Type IV contains a mixture of verbs, some of which have obvious connection to other types, and some which do not. A good example of the former case is *poživat'*^{i} 'live, get by/along', which is obviously related to the Type I pair *po|žit'* 'live (a while)'. The latter case (that is, those verbs not connected to similar roots in other types) can be illustrated by the verb *poricat'*^{p} 'blame, reproach' – although this verb is related to other verbs in Russian (*otricat'*^{i} 'negate, refute'), the presence of the prefix *po-* can again best be understood from a historical perspective. Given the heterogeneity of verbs in Types IV and V, it is uncertain what, if any, general trends further research may uncover.

Whether or not verbs of Type II, III, IV, and V can be considered "prefixed" in the same sense as Type I verbs is open to debate. It is obvious that the *po-* in these verbs was originally prefixal: Type III verbs demonstrate this fact by the existence of the unprefixed Imperfective, and many of the roots of Type II verbs form prefixed aspectual pairs using other prefixes – compare *polučit'*^{p} / *polučat'*^{i} 'receive' with the following:

zalučit'^{p} / *zalučat'*^{i} 'entice, lure'
razlučit'^{p} / *razlučat'*^{i} 'separate'
otlučit'^{p} / *otlučat'*^{i} 'separate, remove' [obsolete]
ulučit'^{p} / *ulučat'*^{i} 'find, seize'

Nevertheless the cognitive status of the prefix requires further study; the question remains whether all, or most, or at least some speakers still analyze the verbs of Type II, III, IV, and V as prefixed. Particular attention should be paid to Type II verbs – despite their ubiquity, the current literature has devoted far less attention to Type II verbs than to the somewhat less common Type I. Simple interviews with randomly selected, "linguistically

naïve" native speakers, in which speakers are asked to name some arbitrary number of *po*-prefixed verbs, might be a good place to start. If Type II verbs are among those listed, this could be interpreted as *prima facie* evidence that Type II verbs are analyzed as prefixed. Or a behavioral profile could be created for *po-* in Type II verbs using a randomly-selected sample from the RNC, and this behavioral profile could be compared to the behavioral profiles for the clearly-attested meanings of *po-* in Type I verbs. Any resulting similarities between the Type II behavioral profile and the behavioral profile of the resultative meaning (for instance) could be considered evidence for a semantic similarity. Behavioral profiles for verbs containing other prefixes (*raz-*, *za-*, or *pro-*, for example) could be used as a control group against which to judge degrees of similarity. The overwhelming frequency of Type II verbs in the data suggests that these verbs are deeply entrenched in the minds of speakers, and as such they may exert an as-yet-undetermined influence on the semantic structure of *po-*. Similar studies could be carried out for the much rarer Type IV and V verbs. Work on Type III, by contrast, could focus on the degree of semantic overlap between the unprefixed and prefixed Imperfective forms, increasing our understanding of the process by which the semantic content of *po-* "bleaches out".

APPENDIX 1: Meanings of *po-* culled from the reference works surveyed

This appendix contains a list of definitions of *po-* gleaned from the following nine reference works: *Грамматика русского языка* [Grammar of the Russian Language] (1960); *Prefiksacja czasownika we współczesnym języku* [Prefixation of the verb in the modern language] (Bogusławski 1963); *Словарь русского языка в четырех томах* [Russian Language Dictionary in 4 volumes] (1959); *Словарь современного русского литературного языка* [Dictionary of the contemporary Russian literary language] (1950-1965); *Russian Grammar* (Unbegaun 1967); *Толковый словарь русского языка* [Explanatory dictionary of the Russian language] (Ожегов & Шведова 2005); *Russian: A practical grammar with exercises* (Pulkina & Zakhava-Nekrasova 1974); *The Oxford Russian Dictionary* (1992); *Грамматический строй русского языка в сопоставлении с словацким* [Grammatical structure of Russian compared to Slovak] (Исаченко 1965); *Введение в русскую аспектологию* [Introduction to Russian aspectology] (Зализняк & Шмелев 2000); *Русский глагол: формы и их функции* [The Russian verb: Form and Function] (Andrews et. al 2004). Each meaning is followed by a number indicating the number of reference works that list it as one of the possible meanings of *po-*.

1. Distributivity: Either the action is performed by a number of subjects, or it is directed towards a number of objects. (10)
   Pobrosal *vse svoi magaziny, osobnjaki i jaxty, zajavilsja v kolledž.*
   'He **threw away all** his stores, mansions and yachts, and applied to college.'

2. Intermittent-attenuative: The action occurred with interruptions and with weakened intensity (several works do not include the stipulation "with weakened intensity"). This meaning arises only when the verb is both prefixed by *po-* and suffixed by –yva-. (9)
   Tol'ko vot muž vsjo čašče ***pogulival*** *na storone, ne udel'aja žene vnimanie…*
   'It's just that the husband **was cheating (a little, from time to time)** more and more often, not paying any attention to his wife….'

3. The action is completed gradually, incrementally, not all at once; this meaning usually arises in multiply-prefixed verbs. (1)
   A ja ko vsemu ***poprivyknul*** *i daže obradovalsja, čto vot opjat' ne odin budu.*
   'And I **got used to** everything (**little by little**) and was even glad that again I wouldn't be alone.'

4. Specification of the action/emphasis: The prefix *po-* indicates that the action is occurring at that very moment, as opposed to a usual or habitual occurrence. (1)
   Gružu i ***pogružaju.*** 'I'm **loading** right now'.

5. Change in spatial conditions or characteristics; covering the surface of an object. (2)
   Da-da, v obščem my im skazali, čto pošli nosik ***popudrit****, a sami na samom dele pošli k drugomu mužiku.*

'Yeah, generally we told them that we were going **to powder** our noses, but in reality we went to up to another guy.'

6. Acquisition of a quality, property. (1)
   *Stepan **poblednel**, pytalsja čto-to skazat'.*
   'Stepan **turned pale**, tried to say something.'

7. Completion of an action in one attempt/motion; short duration, momentaneousness. (3)
   *On **poblagodaril** menja.*
   'He **thanked** me.'

8. Completion of an action: The action has reached its (natural/expected) result. (5)
   *Milliarder **porval** kartinu Pikasso.*
   'The billionaire **ripped** a painting by Picasso.'
   *Buš **pozvonil** Putinu iz-za iranskogo krizisa.*
   'Bush **called** Putin because of the Iran crisis.'

9. Ingressivity: The po-prefixed verb indicates the inception of the action named by the verb. (10)
   *Nastja vdrug **pobežala** vpered.*
   'Nastja suddenly **took off running** straight ahead'.

10. Simple perfectivity (without additional connotations). (4)
    *Oni podarili mne cvety, **pocelovali** ruku i pri ètom oba gljadeli na menja, kak na ikonu, vlažnymi temno-karimi glazami.*
    'They gave me flowers, **kissed** my hand and in so doing they both looked at me, as at an icon, with teary, dark-brown eyes.'

11. Delimitativity[1]: The action is limited in time, without natural endpoint; often, the action has short duration. (10)
    ***Porabotala** nedolgo u nas v žurnale (mne èto bystro nadoelo), a potom rešila byt' vnučkoj-sekretarem bessmertnogo akademika.*
    'I **worked a short while** at our magazine company (I quickly got fed up with that), and then I decided to be the granddaughter-secretary of an immortal academic'.

12. Attenuative: The action occurs with less-than-usual intensity. (8)
    *Na vtoroj den' bolezni rodstvenniki priglasili vrača, kotoryj dal uspokoitel'nyx kapel' i velel čem-nibud' **porazvleč** bol'nogo, otvleč ego vnimanie ot slučajnoj bolezni.*
    'On the second day of the illness his relatives called the doctor, who administered calming droplets and ordered that they **entertain** the patient **a little**, distract his attention from the chance illness'.

[1] Three reference works recognize an overlap between the delimitative and the attenuative meanings of *po-*.

APPENDIX 2: List of tags used by the Russian National Corpus

This appendix lists all tags used by the RNC as of 2008-5-15. Tags are organized according to type. The purpose of this appendix is to show which tags were in use at the time of data collection.

Part of speech
S — существительное (*яблоня, лошадь, корпус, вечность*)
A — прилагательное (*коричневый, таинственный, морской*)
NUM — числительное (*четыре, десять, много*)
A-NUM — числительное-прилагательное (*один, седьмой, восьмидесятый*)
V — глагол (*пользоваться, обрабатывать*)
ADV — наречие (*сгоряча, очень*)
PRAEDIC — предикатив (*жаль, хорошо, пора*)
PARENTH — вводное слово (*кстати, по-моему*)
S-PRO — местоимение-существительное (*она, что*)
A-PRO — местоимение-прилагательное (*который, твой*)
ADV-PRO — местоименное наречие (*где, вот*)
PRAEDIC-PRO — местоимение-предикатив (*некого, нечего*)
PR — предлог (*под, напротив*)
CONJ — союз (*и, чтобы*)
PART — частица (*бы, же, пусть*)
INTJ — междометие (*увы, батюшки*)

Gender
m — мужской род (*работник, стол*)
f — женский род (*работница, табуретка*)
m-f — «общий род» (*задира, пьяница*)
n — средний род (*животное, озеро*)

Animacy
anim — одушевленность (*человек, ангел, утопленник*)
inan — неодушевленность (*рука, облако, культура*)

Number
sg — единственное число (*яблоко, гордость*)
pl — множественное число (*яблоки, ножницы, детишки*)

Case
nom — именительный падеж (*голова, сын, степь, сани, который*)
gen — родительный падеж (*головы, сына, степи, саней, которого*)
dat — дательный падеж (*голове, сыну, степи, саням, которому*)
acc — винительный падеж (*голову, сына, степь, сани, который/которого*)
ins — творительный падеж (*головой, сыном, степью, санями, которым*)
loc — предложный падеж (*[о] голове, сыне, степи, санях, котором*)
gen2 — второй родительный падеж (*чашка **чаю***)

acc2 — второй винительный падеж (*постричься в **монахи**; по **два** человека*)
loc2 — второй предложный падеж (*в **лесу**, на **осú***)
voc — звательная форма (*Господи, Серёж, ребят*)
adnum — счётная форма (*два **часá**, три **шарá***)

Form
brev — краткая форма (*высок, нежна, прочны, рад*)
plen — полная форма (*высокий, нежная, прочные, морской*)

Degree
comp — сравнительная степень (*глубже*)
comp2 — форма «*по*+сравнительная степень» (*поглубже*)
supr — превосходная степень (*глубочайший*)

Aspect
pf — совершенный вид (*пошёл, встречу*)
ipf — несовершенный вид (*ходил, встречаю*)

Transitivity
intr — непереходность (*ходить, вариться*)
tran — переходность (*вести, варить*)

Voice
act — действительный залог (*разрушил, разрушивший*)
pass — страдательный залог (только у причастий: *разрушаемый, разрушенный*)
med — медиальный, или средний залог (глагольные формы на *-ся*: *разрушился* и т.п.)

Non-finite verb form
inf — инфинитив (*украшать*)
partcp — причастие (*украшенный*)
ger — деепричастие (*украшая*)

Mood
indic — изъявительное наклонение (*украшаю, украшал, украшу*)
imper — повелительное наклонение (*украшай*)
imper2 — форма повелительного наклонения 1 л. мн. ч. на *-те* (*идемте*)

Tense
praet — прошедшее время (*украшали, украшавший, украсив*)
praes — настоящее время (*украшаем, украшающий, украшая*)
fut — будущее время (*украсим*)

Person
1p — первое лицо (*украшаю*)
2p — второе лицо (*украшаешь*)
3p — третье лицо (*украшает*)

Other grammatical
persn — личное имя (*Иван, Дарья, Леопольд, Эстер, Гомер, Маугли*)
patrn — отчество (*Иванович, Павловна*)
famn — фамилия (*Николаев, Волконская, Гумбольдт*)
0 — несклоняемое (*шоссе, Седых*)

Non-standard
anom («Аномальная форма») — различного рода морфологические аномалии, возможные у устаревших или просторечных нелитературных форм (*три дни* при нормативном *три дня*, *ляжь* при нормативном *ляг*)
distort («Искаженная форма») — орфографическое и/или фонетическое искажение слова, часто передающее различные особенности произношения (*дэвушка, това'ищи, про-хо-ди, низнаю*).
ciph («Цифровая запись») — запись числительного, числительного-прилагательного или прилагательного (полностью или частично) при помощи цифр (*73, LXXIII, 73-й, 22-летний*). Для этих словоформ в поле «Лексема» также употребляется цифровая запись; число и падеж указываются только в тех случаях, когда выписано окончание (типа *14-му*).
INIT («Инициал») — запись вида «заглавная буква с точкой» (*М., Р.*). В поле «Лексема» инициал не раскрывается; грамматические признаки не указываются.
abbr («Сокращение») — сокращенная запись (*тов., гг., ч.*). В поле «Лексема» сокращение (кроме инициалов) раскрывается, указывается грамматическая форма, соответствующая контексту. Специально отметим, что акронимы вроде *ООН, вуз* и усеченные слова вроде *зав, зам*, записываемые без точки и не раскрываемые при чтении, не получают пометы abbr и трактуются как обычные слова (склоняемые или несклоняемые).
bastard: Кроме того, в корпусе с неснятой грамматической омонимией используется особая помета (bastard) для несловарной формы (не входящей в словарь автоматического анализатора, а порожденной по аналогии, например, форма вроде *Махабхарата* получает несколько гипотетических разборов, в том числе от псевдолексем *махабхаронок*, *махабхарать* и т. п.); по мере пополнения словаря анализатора число таких форм будет уменьшаться.
С целью снижения «шума» при поиске по корпусу с неснятой грамматической омонимией иногда бывает целесообразно исключить поиск по подобным формам; для ряда задач, напротив, можно ограничить поиск именно ими.

Category
r:concr — предметные имена (*девочка, стол, молоко*)
r:abstr — непредметные имена (*вождение, яркость, время*)
r:propn — имена собственные (*Иван, Эйнштейн, Петроград*)
r:qual — качественные (*хороший, большой*)
r:rel — вопросительные/относительные (*деревянный, лунный; кто, который, когда*)
r:poss — притяжательные (*божий, отцов, мужнин; мой, его, свой*)
r:invar — неизменяемые (*беж, джерси*)
r:card — количественные (*два, пять, десять*)
r:card:pauc — числительные малого количества (*два, три, четыре, оба, пол, полтора*)
r:ord — порядковые (*первый, второй, десятый*)
r:pers — личные (*я, он*)
r:ref — возвратные (*себя*)
r:dem — указательные (*этот, такой*)
r:indet — неопределенные (*некоторый, некогда*)
r:neg — отрицательные (*никакой, ничей*)
r:spec — кванторные (определительные) (*всякий, каждый, любой*)

(Semantic) Taxonomy
t:action — мероприятие (*аукцион, вернисаж, вечеринка, выборы, именины, заседание, культпоход*)
t:animal — животные (*корова, жираф, сорока, ящерица, муравей*)
t:be — бытийная сфера (*жить, возникнуть, убить*)
t:be:appear — начало существования (*возникновение, рождение, формирование, учреждение, творение*; *возникнуть, родиться, сформировать, создать*)
t:be:disapp — прекращение существования (*смерть, казнь, ликвидация*; *умереть, убить, улетучиться, ликвидировать, искоренить*)
t:be:exist — существование (*жизнь, наличие, бытие*; *жить, происходить*)
t:behav — поведение и поступки человека (*разгильдяйство, подхалимаж, неповиновение, ребячество, предательство*; *куролесить, привередничать*)
t:changest — изменение состояния или признака (*взрослеть, богатеть, расширить, испачкать; укрепление, затвердение, осушение, конденсация, осложнение*)
t:color — цвет (*окраска, колорит, желтизна, прозелень*)
t:constr — здания и сооружения (*дом, шалаш, мост*)
t:contact — контакт и опора (*касаться, обнимать, облокотиться*; *прикосновение, объятие*)
t:dir — направление (*обратный, подветренный*; *туда, наверх*)
t:disease — болезнь (*ангина, диабет*)
t:dist — расстояние (*далекий, соседний*; *далеко, близко*)
t:dist:max — большое (*далеко, вдали, вдалеке*; *дальний, отдаленный*)
t:dist:min — малое (*близкий, недалекий*; *близко, вблизи*)
t:famn — фамилии (*Пушкин*)
t:food — еда и напитки (*пирог, каша, молоко*)
t:game — игра (*жмурки, покер, домино, волейбол*)
t:hum — лица (*человек, учитель, Людмила*)

t:hum:etn — этнонимы (*эфиоп, итальянка*)
t:hum:kin — имена родства (*брат, бабушка*)
t:hum:supernat — сверхъестественные существа (*русалка, инопланетянин, Черномор*)
t:humq — качества человека (*умный, верный, ловкий*; *порядочность, безволие, остроумие*)
t:impact — физическое воздействие (*бить, колоть, вытирать*; *удар, втирание, обмолот*)
t:impact:creat — создание физического объекта (*выковать, смастерить, сшить*; *лепка, отливка, плетение, сооружение, строительство*)
t:impact:destr — уничтожение (*взорвать, сжечь, зарезать*; *слом, сожжение*)
t:inter — взаимодействие и взаимоотношение (*взаимопомощь, вражда, схватка, драка*)
t:light — свет (*гаснуть, лучиться*; *луч, полумрак, светлынь, иллюминация*)
t:loc — местонахождение (*лежать, стоять, положить*; *местоположение*)
t:loc:body — положение тела в пространстве (*лежание*; *сидеть*)
t:ment — ментальная сфера (*знание, абстракция, воображение, воспоминание, догадка; знать, верить, догадаться, помнить, считать*)
t:move — движение (*беготня, вынос, качка*; *бежать, дергаться, бросить, нести*)
t:move:body — изменение положения тела, части тела (*поклон*; *согнуть, нагнуться, примоститься*)
t:param — параметр (*высота, грузоподъемность*)
t:patrn — отчества (*Сергеевич*)
t:perc — восприятие (*осязание, слух, видимость, взгляд, зрелище*; *смотреть, слышать, нюхать, чуять*)
t:persn — имена (*Александр*)
t:physiol — физиологическая сфера (*жажда, кровоизлияние, судорога, утомление, икота*; *кашлять, икать*)
t:physq — физические свойства (*мягкий, вязкий*)
t:physq:color — цвет (*красный, бесцветный*)
t:physq:form — форма (*кривой, круглый*)а
t:physq:smell — запах (*ароматный, тухлый*)
t:physq:taste — вкус (*кислый, приторный*)
t:physq:temper — температура (*горячий, ледяной*)
t:physq:weight — вес (*тяжелый, легкий*)
t:place — место (*здесь, посередине*; *левый, придорожный, теменной*)
t:plant — растения (*береза, роза, трава*)
t:poss — посессивная сфера (*иметь дать, подарить, приобрести, лишиться*; *обладание, приобретение, покупка, потеря, лишение*)
t:psych — психическая сфера (*апатия, безумие, вдохновение, спокойствие*; *гипнотизировать, сочувствовать, настроиться, терпеть*)
t:psych:emot — эмоция (*восторг, раскаяние, печаль*; *радоваться, обидеть*)
t:psych:volit — воля (*намерение, решение*; *решить*)
t:put — помещение объекта (*положить, вложить, спрятать*; *размещение, расстановка, погрузка, намотка*)

t:quant — количество (*большой, достаточный, трехкратный*; *столько, достаточно*)
t:quant:abs — абсолютное (*двухтысячный, восьмимилионный*)
t:quant:max — большое (*много, навалом*; *обильный, многочисленный*)
t:quant:min — малое (*мало, чуть-чуть*; *ничтожный, малочисленный*)
t:size — размер (*высокий, короткий*)
t:size:abs — абсолютный (*двухэтажный*)
t:size:max — большой (*высокий, длинный*)
t:size:min — малый (*низкий, короткий*)
t:smell — запах (*аромат, перегар*; *пахнуть, благоухать*)
t:sound — звук (*гудеть, шелестеть*; *шум, перезвон, хлопок, аплодисменты, диссонанс*)
t:space — пространство и место (*космос, город, тайга, овраг, вход*)
t:speech — речь (*говорить, советовать, спорить, каламбурить*; *дискуссия, молва, ахинея, реплика, подковырка*)
t:speed — скорость (*быстро, медленно*; *проворный*)
t:speed:max — большая (*быстро, мигом*; *скорый, быстрый*)
t:speed:min — малая (*медленно, неторопливо*; *медленный, тягучий*)
t:sport — спорт (*спартакиада, акробатика, баскетбол*)
t:stuff — вещества и материалы (*вода, песок, тесто, жесть, шелк*)
t:taste — вкус (*вкуснота, горчинка, кислятина*)
t:temper — температура (*прохлада, стужа, нагрев*)
t:text — тексты (*рассказ, книга, афиша*)
t:time — время (*весна, годовщина, минута, современность*; *прошлый, ночной*; *тогда, поздно*)
t:time:age — возраст (*детство, молодость, двадцатилетие*)
t:time:age — возраст (*зрелый*)
t:time:age:abs — абсолютный (*трехлетний*)
t:time:age:max — большой (*старый, древний*)
t:time:age:min — малый (*молодой, малолетний*)
t:time:dur — длительность (*вечно, недолго*; *долгий, краткий*)
t:time:dur:abs — абсолютная (*восьмичасовой*)
t:time:dur:max — большая (*вечно, подолгу, всегда*; *долгий, продолжительный*)
t:time:dur:min — малая (*временно, недолго*; *краткий, кратковременный*)
t:time:moment — момент (*миг, мгновение*)
t:time:month — месяц (*январь*)
t:time:period — период (*межсезонье, путина, сенокос, стаж*)
t:time:week — день недели (*понедельник*)
t:tool — инструменты и приспособления (*молоток, палка, пуговица, машина*)
t:tool:cloth — одежда и обувь (*платье, шляпа, ботинки*)
t:tool:device — механизмы и приборы (*телефон, сеялка, градусник*)
t:tool:dish — посуда (*чашка, кастрюля, фляжка*)
t:tool:furn — мебель (*стол, диван, шкаф*)
t:tool:instr — инструменты (*молоток, штопор, игла, карандаш*)
t:tool:mus — музыкальные инструменты (*рояль, скрипка, колокол*)
t:tool:transp — транспортные средства (*автобус, поезд, сани*)

t:tool:weapon — оружие (*сабля, пистолет, гаубица*)
t:topon — топонимы (*Европа, Волга, Эльбрус, Москва, Преображенка*)
t:unit — единица измерения (*балл, килограмм, метр, минута*)
t:weather — природное явление (*бушевать, вьюжить*; *зарница, вьюга, зной*)

Mereology
hi:class — имена классов (*животное, ягода, инструмент*)
pt:part & pc:constr — части зданий и сооружений (*комната, дверь, арка*)
pt:part & pc:plant — части растений (*лист, ветка, корень*)
pt:part & pc:tool — части приспособлений (*деталь, лопасть, крышка*)
pt:part & pc:tool:cloth — части одежды и обуви (*рукав, каблук*)
pt:part & pc:tool:device — части механизмов и приборов (*дисплей, корпус, кнопка*)
pt:part & pc:tool:dish — части предметов посуды (*носик, горлышко*)
pt:part & pc:tool:furn — части предметов мебели (*сиденье, подлокотник*)
pt:part & pc:tool:instr — части инструментов (*топорище, лезвие*)
pt:part & pc:tool:mus — части музыкальных инструментов (*струна, гриф*)
pt:part & pc:tool:transp — части транспортных средств (*руль, колесо, капот*)
pt:part & pc:tool:weapon — части оружия (*дуло, курок, эфес*)
pt:part — части (*верхушка, кончик, половина; начало, финал*)
pt:partb & pc:animal — части тела и органы животных (*хвост, жало*)
pt:partb & pc:hum — части тела и органы человека (*голова, сердце, ноготь*)
pt:qtm — кванты и порции вещества (*капля, комок, порция; оборот, прыжок, кивок*)
pt:set — множество (*система, выборка, алгоритм*)
pt:set | pt:aggr — множества и совокупности объектов (*набор, букет, мебель, человечество*)

Topology
top:contain — вместилища (*кошелек, комната, озеро, ниша*)
top:horiz — горизонтальные поверхности (*пол, площадка*)

Evaluation
ev — оценка (неопределенная по признаку «положительная/отрицательная») (*беспечно, бойко*; *озорник, махина*; *толковый, мешковатый*)
ev:neg — отрицательная (*безвкусица, ахинея*; *бездарно, неловко*; *негодяй, вертихвостка*; *продажный, сварливый*)
ev:posit — положительная (*благоухание, загляденье, изюминка*; *бойко, безупречно*; *везучий, ладный*; *умница, светило*)

Word-formation
d:atten — аттенуативы (*рановато, суховато*; *угловатый, жуликоватый*)
d:aug — аугментативы (*детина, домище*; *здоровенный, злющий*)
d:carit — каритивы (*безглазый, бездыханный*)
d:dim — диминутивы (*зайчик, коробочка*; *немножко, быстренько*; *Саша, Женечка, Николаич*; *тихонький, крохотный*)
d:fem — nomina feminina (*немка, генеральша, доярка*)

d:habit — хабитивы (*глазастый, пузатый*)
d:impf — вторичные имперфективы (-ива-, -ва-, -а-) (*выпивать, вбивать, прогонять*)
d:impot — импоссибилитивы (*несоизмеримый, недееспособный*)
d:nag — nomina agentis (*писатель, создатель, докладчик*)
d:potent | d:impot — потенциальные (*плавучий, недееспособный*)
d:potent — поссибилитивы (*плавучий, плодородный, занимательный*)
d:pref — приставочные глаголы (*забегать, оглядеть*)
d:semelf — семельфактивы (*кивнуть, чихнуть, боднуть, качнуться*)
d:sing — сингулятивы (*пылинка, изюминка*)
der:a & dt:humq — качества человека (*внимательно, грубо*)
der:a & dt:physq — физические свойства (*твердо, плотно*)
der:a & dt:physq:color — цвет (*красно, добела*)
der:a & dt:physq:form — форма (*плоско, прямо*)
der:a & dt:physq:smell — запах (*смрадно, зловонно*)
der:a & dt:physq:taste — вкус (*горько, вкусно*)
der:a & dt:physq:temper — температура (*тепло, прохладно*)
der:a & dt:physq:weight — вес (*тяжело, легко*)
der:a & dt:size — размер (*высоко, коротко*)
der:a & dt:size:max — большой (*высоко, бесконечно*)
der:a & dt:size:min — малый (*коротко, низко*)
der:a — отадъективные имена, наречия (*краснота, жадность*; *быстро, обычно*)
der:adv — отадвербиальные прилагательные (*поздний, здешний*)
der:s — отыменные наречия, прилагательные (*вверху, дома*; *домашний, железный*)
der:v — отглагольные имена, наречия, прилагательные (*выбор, демонстрация*; *отродясь, стоймя*; *ковкий, навязчивый, кочевой*)

Causativity
ca:caus — каузативные глаголы (*показать, вертеть*)
ca:noncaus — некаузативные глаголы (*видеть, вертеться*)

Auxiliary Verbs
aux:phase — фазовые (*начать, продолжать, прекратить*)
aux:caus — служебные каузативные (*вызвать, привести (к)*)

APPENDIX 3: Verbs by Type

This appendix lists the 709 *po*-prefixed verbs that occurred in the 16,121 observations of *po*-prefixed verbs extracted from the dehomonymized portion of the RNC, grouped by Type (see Section 3.3.1 for an explanation of verb Types). Glosses are provided. Shades of Aktionsart meaning are generally only provided to disambiguate between similarly glossed verbs. Whenever a meaning is marked "do X (intermittently)", it should be understood that a sense of attenuation ("do X a little") is also often implied.

Type I Verbs (534)

po|agitirovat'	agitate
pobaivat'sja	be rather afraid
po|balovat'	treat; spoil
po|begat'	run (around)
po|bežat'	(start to) run
po|belit'	whitewash
po|besedovat'	chat, converse
po|bespokoit'sja	worry (about), trouble (about)
po|bit'	hit
po|blagodarit'	thank
po|blaženstvovat'	be blissful
po|blednet'	turn pale
po|bleknut'	fade; wither
pobleskivat'	gleam
po|božit'sja	swear
po|bolet'	cheer
po|boltat'	chat
po|bombit'	bomb
po|borot'	fight down, overcome
po|borot'sja	struggle, wrestle
po|bojat'sja	fear
po|bratat'sja	fraternize
po|brat'	take (in quantity)
po|brezgovat'	be squeamish about
po|bresti	(start to) stroll
pobrit'sja	shave
po|brodit'	stroll (around)
po|brosat'	throw
po|bryzgat'	sprinkle
po|buxat'	party
pobuxivat'	thunder, thud
po|valit'	fall/throw down; pile up
po|valit'sja	collapse
po|valjat'sja	wallow
po|varit'	roast, cook
po|vdovet'	widow, be widow(er)
po|vezti I	carry, transport
po|vezti II	be lucky
po|venčat'sja	get married
po|verit'	believe
po|vertet'	rotate, twist
po|veselet'	cheer up, become cheerful/merry
po|veselit'	cheer, amuse, divert
po|veselit'sja	enjoy oneself, have a good time
povesit' / vešat'	hang; pin, throw (on) (lit. & fig.)
povesit'sja / vešat'sja	hang oneself
po|vesti	begin to lead/conduct
po|vestis'	become the custom
po|vejat'	begin to blow; "one could sense" (impers.)
po|vzdyxat'	sigh, breathe; long for, yearn for
po|vzroslet'	grow up, mature
po|vidat'	see
povizgivat'	cry shrilly
po|viset'	hang
po|vleč'	draw, drag, pull, attract
po|vlijat'	influence
po|voevat'	wage war, fight; quarrel
po|vozit'sja	mess around, busy oneself
po|volnovat'sja	worry, be agitated
povoloč' / voločit'	drag
po|vonjat'	stink; fart; make a fuss
po|vorčat'	grumble
po|vremenit'	wait

povskakivat' jump up
po|vstrečat' meet by chance
po|vstrečat'sja meet by chance
po|vylezti come out, climb out
po|vysypat' empty, pour, spill out
po|vytaskivat' pull out, extract
po|vyt' wail, howl
po|vjazat' knit, tie, bind
po|gadat' tell a fortune
po|gladit' I iron, press
po|gladit' II stroke, caress
po|glazet' stare, gawk (at)
po|glumit'sja mock, jeer, scoff
po|gljadet' look; glance
pogljadyvat' look (intermittently)
po|gljanut'sja take a liking to
po|gnat' (start to) drive
po|gnat'sja start after, give pursuit
pogovarivat' talk a little, intermittently
po|govorit' talk
po|gonjat' drive on
po|goret' burn down/out; 'go up in flames'
po|gorjačit'sja get heated, worked up
po|gostevat' be a guest
po|gostit' stay at, visit, be a guest
po|gret'sja warm oneself
po|grozit' make a threatening hand gesture
po|groxotat' crash, roar, rumble
po|gruzit'sja be loaded, freighted
po|gubit' destroy, ruin
po|gudet' party, have a good time
po|guljat' take a walk; have a good time
po|davit' press, squeeze; run over
po|darit' give as a gift
po|dvigat' move
po|dvigat'sja be moved
po|devat'sja get (to); disappear
po|dejstvovat' have an effect on, act, work
po|delat' do
po|delit' divide, share
po|delit'sja divide, share
podergivat' pull, tug
po|deržat' hold
po|dernut' cover
po|deševet' become cheaper
po|dobret' become kinder
po|doit' milk, drain
po|dorožat' rise in price, go up
po|drat'sja fight (over, with)
po|dremat' doze
po|družit'sja make friends (with)
po|dumat' think
po|dumat'sja be thought
podumyvat' think a little, intermittently
po|durnet' lose one's good looks, look worse
po|dut' (begin to) blow
po|dyšat' breathe
poeživat'sja shudder (a little, intermittently)
po|ežit'sja shiver, shudder
po|ezdit' drive around
po|erošit' tousle, dishevel
po|exat' (start to) go by vehicle
po|žadničat' be greedy/mean
po|žalet' be sorry for, pity
po|žalovat' grant, bestow; visit
po|žalovat'sja complain
po|žarit' cook, roast
poževyvat' chew
po|želat' wish, desire
po|želtet' turn yellow
po|ženit'sja get married
po|žertvovat' donate, sacrifice
po|žit' live
po|žurit' rebuke
po|žuxnut' dry up, shrivel
po|zabavit' amuse
po|zabotit'sja look after, care for, be concerned about
po|zavidovat' envy
po|zavtrakat' have breakfast
po|zaimstvovat' borrow
po|zanimat'sja study, engage in
po|zarit'sja covet, have one's eye on
pozvanivat' call (intermittently)
po|zvat' call (for)

po|zvonit' call, phone
pozvjakivat' tinkle
po|zdorovat'sja greet
po|zloradstvovat' gloat
po|znakomit' introduce
po|znakomit'sja be introduced, get to know
po|zolotit' gild
po|igrat' play
po|izdevat'sja mock
po|imenovat' name
po|imet' get, acquire
po|interesovat'sja be interested
po|iskat' look, search
po|ispovedat'sja confess
po|isteret'sja grate, use up by rubbing
po|jmat' catch
po|jmat'sja be caught
po|jti set out
po|kakat' poop (bodily function)
po|kalečit' cripple, maim; destroy, ruin
po|kalečit'sja be crippled, maimed; be destroyed
po|kaljakat' have a chat
po|karat' punish, chastise
po|karaulit' guard, keep watch (on, for)
po|katat' roll (around); take for a ride
po|katat'sja go for a ride
po|katit' (start to) roll
po|katit'sja (start to) roll, drive
po|kačat' rock, shake, wave
pokačivat' rock (a little, intermittently)
po|kačivat'sja be rocked
po|kačnut'sja sway, give a lurch
pokašlivat' cough (intermittently)
po|kašljat' cough
po|kajat'sja repent, confess
po|kvitat'sja settle, get even (with)
po|kidat' throw, fling
po|klast' put, lay
po|klevat' peck
poklevyvat' peck (intermittently)
poklonit'sja / klanjat'sja bow (before)
po|kljast'sja swear, vow
po|kovyrjat' pick, peck at
pokolačivat' hit, beat up
po|koldovat' conjure, practice sorcery
po|kolebat' shake
po|kolebat'sja be shaken
po|kolotit' beat
po|komandovat' command, be boss
po|končit' put an end to, finish off
pokormit' feed
po|korobit' warp; make shudder/sick
po|kosit'sja become lopsided; glance sidelong
po|krasit' paint, dye; make-up
pokrasit'sja be made-up; paint, stain
po|krasnet' turn red
po|krestit' baptize, christen
po|krivit' bend, twist, distort
pokrikivat' cry, shout; shout at, chide
po|kritikovat' criticize
pokrošit' crumble
po|krutit' twist, turn, rotate
po|krutit'sja turn, rotate
pokrjakivat' quack
po|krjaxtet' groan
po|kuvyrkat'sja sommersault
po|kuražit'sja mock, jeer
pokurivat' smoke (intermittently)
po|kurit' smoke
po|kurolesit' play pranks
po|kusat' bite
pokusyvat' bite (intermittently)
po|kušat' eat
po|ladit' come to an understanding
po|lakomit'sja regale oneself with
po|legčat' lessen, abate
po|ležat' lie down
po|lezt' (start to) climb
po|lenit'sja be lazy
po|letat' fly (around)
po|letet' (start to) fly
po|lizat' lick
po|listat' leaf through, turn pages
položit' / klast' lay, place
po|lomat' break

po|lopat'sja burst
po|l'stit' flatter
po|l'stit'sja be tempted (with, by)
po|ljubit' (start to) love
po|ljubit'sja be loved; catch someone's fancy
po|ljubovat'sja admire, feast eyes on/upon
po|ljubopytstvovat' be curious, ask
po|mazat' daub, spread
pomalkivat' hold one's tongue
po|manit' beckon; lure
po|massirovat' massage
po|maxat' wave
pomaxivat' wave (intermittently)
po|medlit' linger
po|menjat' change
po|menjat'sja (be) change(d)
po|merit' try on (for fitting)
po|merit'sja compete/measure up against
po|merknut' fade; grow dim
po|mečtat' dream
po|mešat' bother, hinder, prevent
pomešivat' stir
po|milovat' show mercy, pardon
po|mirit'sja be reconciled
po|molit'sja pray
po|molodet' become younger
po|molčat' be quiet
po|morgat' wink, blink
po|morščit'sja wrinkle one's nose, knit one's brow
po|motat' shake
po|močit'sja urinate
po|mračnet' become gloomy, darken
po|mučit'sja suffer
po|mčat'sja dart, speed, tear
po|mylit' steal
po|myslit' think, conceive, imagine
po|myt' wash
po|myt'sja be washed
po|mjat' press, crush
po|nabežat' rush in, run in
po|nabljudat' observe; take care of
po|nabrat' gather, take
po|nadelat' do, make (in quantity)
po|nadejat'sja count on, rely on
po|naexat' come (in numbers)
po|nažimat' press
ponapadat' attack (in numbers)
po|napisat' write a lot
po|narisovat' draw a lot
po|nastavit' place in quantity
po|nastroit' build in quantity
po|natykat' stick in (in numbers)
po|naxvatat' get hold of (in numbers)
po|nežit' indulge, pamper
po|nervničat' be nervous
po|nesti carry, bear
po|nestis' rush after, dart after
po|nosit' wear
po|nostal'girovat' be nostalgic
po|nravit'sja like
po|njuxat' sniff, smell
po|obedat' have lunch
po|obeščat' promise
po|obžit'sja get used to
po|oblomat' break
po|obnimat'sja hug, embrace
po|obščat'sja communicate
po|orat' yell
po|ostereč'sja beware of
po|ostyt' cool down
po|oxotit'sja hunt
po|padat' fall
po|parit'sja take steam bath
popaxivat' smell
po|peret' push forward
po|perčit' pepper
po|pet' sing
popivat' drink (intermittently)
po|pirovat' feast
po|pisat' write
po|pit' drink
po|plavat' swim (around)
po|plakat' cry
po|plakat'sja complain, lament
po|platit'sja pay (for…with…)
po|plevat'sja spit
po|plestis' drag oneself along
po|plyt' (start to) swim

po|pljasat' dance
po|polemizirovat' polemicize
po|polzti (start to) crawl
po|pol'zovat'sja make use of; enjoy
po|pomnit' remember
po|portit' spoil
po|postit'sja fast
po|potet' sweat, work hard
po|potčevat' treat (to), entertain
po|prazdnovat' celebrate
po|praktikovat'sja practice
po|pribavit' add
po|privetstvovat' greet
po|prideržat' hold back
po|prisutstvovat' stay, attend
po|probovat' try
po|prosit' ask
po|prosit'sja ask; yearn for
po|prygat' jump
po|prjatat'sja hide
po|pustit' / puskat' let pass, allow
po|putat' get mixed up in
po|pytat' try, test
po|pytat'sja try, attempt, endeavor
po|pjatit'sja move backwards
po|rabotat' work
po|radovat' make happy
po|radovat'sja be glad, rejoice
po|razvešivat' weigh out
po|razvleč'sja have good time, fun
po|razmaxivat' swing
po|razmyslit', po|razmyšljat' think, consider
po|raskinut' stretch, spread
po|rassprosit' ask around
po|rassuždat' discuss, argue
po|rvat' tear
po|rvat'sja be torn
po|redet' thin out; be depleted
po|rezat' cut; (fig.) kill
po|rezvit'sja frolic, romp
po|rekomendovat' recommend
po|repetirovat' rehearse, coach
po|rešat' decide
po|rešit' decide
po|rubit' chop, hew
po|rugat'sja quarrel (with)
po|rulit' steer, be in charge
po|rušit' destroy
po|rybačit' fish
po|ryžet' turn reddish
po|ryt'sja dig, rummage
posadit' / sažat' seat, plant
po|sbivat' knock down
po|svetit' shine
po|svetlet' become lighter, visible
po|sedet' turn gray
po|serebrit' silver-plate
po|ser'eznet' become serious
po|setovat' lament
po|sejat' sow
po|sidet' sit
posižívat' sit (intermittently)
po|sinet' turn blue
po|skresti scratch
poskripyvat' scratch, creak
po|skučat' miss / be bored
po|sledit' look after, see to
po|sledovat' follow
po|služit' serve
po|slušat' listen
po|slušat'sja obey
po|slyšat'sja be heard
po|sljunit' wet with saliva
posmatrivat' look
posmeivat'sja laugh (a little, intermittently)
po|smet' dare
po|smejat'sja laugh
po|smotret' look
po|snimat' film
po|sovetovat' advise
po|sovetovat'sja consult
po|soveščat'sja deliberate
po|sodejstvovat' assist
po|sožalet' pity, regret
po|solit' salt; pickle
po|soperničat' compete, rival
po|sopet' breathe in heavily
po|sorevnovat'sja compete
po|sočuvstvovat' sympathize
po|spat' sleep

po|sporit' argue
po|sposobstvovat' assist, aid
po|ssorit'sja argue, fight (with)
po|stavit' put, place
postanyvat' moan
po|starat'sja try, endeavor
po|staret' grow old
po|stesnjat'sja be shy, ashamed
po|storonit'sja stand, step aside
po|stojat' stand
po|stradat' suffer
po|streljat' shoot
po|stroit' build
pos|troit'sja be built
po|stukivat' knock (lightly, intermittently)
po|stučat' knock
po|stučat'sja knock
po|sudit' judge
po|suetit'sja fuss, bustle
po|sulit' promise
po|sčastlivit'sja be lucky (have opportunity)
po|sčitat' count; think
po|sčitat'sja be considered
po|sypat'sja rain, pour
po|tait' hide, conceal, secret
po|tancevat' dance
po|taskat' drag
po|taščit' (start to) drag
po|taščit'sja trail along
po|temnet' grow darker
po|teplet' get warm
potereblivat' tug (at)
po|terpet' suffer, endure; bear
po|terjat' lose
po|terjat'sja be lost
po|tesnit' press, crowd
po|tesnit'sja be crowded, squeezed
po|teč' (start to) flow
po|tonirovat' tint
po|topat' hit the road
po|topit' sink
po|toptat'sja dawdle; shift feet
po|torgovat'sja bargain, haggle
po|toropit' hasten, hurry
po|toropit'sja be in a hurry
po|tratit' spend; waste
po|trebovat' require
po|trebovat'sja be required
po|trenirovat' train, exercise
po|trepat' pull about, tousle
po|trepyxat'sja flutter
po|treskat'sja crack, chap
potreskivat' crackle
po|trogat' touch
po|trudit'sja take trouble to, work
potrjasyvat' shake
po|tusknet' grow dull, lose luster
po|tusovat'sja party
po|tuxnut' go out, die out
po|tušit' put out
po|tykat' poke, prod
po|tjagat'sja contend with
potjagivat' sip
po|tjanut' pull
po|tjanut'sja move toward; stretch, lie (fig.)
po|ubavit'sja diminish
poubivat' kill
po|uvečit' maim, cripple
po|užinat' have dinner
po|ulučšat' improve
po|umnet' grow wiser
po|utixnut' fade away
po|uxaživat' nurse
po|učastvovat' participate
po|učit'sja study
poxaživat' walk
po|xvalit' praise
po|xvastat' brag, boast
po|xvastat'sja brag, boast
po|xixikat' giggle
po|xlopat' slap, pat
poxlopyvat' slap, pat
po|xmykat' make the sound "hmhm"
po|xodatajstvovat' solicit, intercede
po|xodit' I walk (around)
po|xolodet' grow cold
po|xoronit' bury
po|xorošet' get better
po|xoxotat' laugh

poxrapyvat' snore
poxripyvat' wheeze, speak hoarsely
po|xromat' limp
po|xudet' become thin
po|carapat' scratch
po|celovat' kiss
po|cokat' click, clatter
po|černet' turn black
po|česat' scratch
po|česat'sja scratch oneself
počesyvat'sja scratch oneself (intermittently)
po|čistit' scrub, brush, clean
po|čitat' read
počityvat' read (intermittently)
počityvat'sja be read
po|čuvstvovat' feel
po|čudit'sja seem, appear
po|čujat' smell; sense
pošalivat' play pranks
po|šalit' play pranks
po|šarit' fumble, rummage
po|šatnut' bend; shake
po|šatnut'sja be bent, shaken; stagger
pošatyvat'sja stagger; sway
po|ševelit'sja move
po|širet' become wider, widen
po|šutit' joke
po|šušukat' whisper
po|ščekotat', po|ščekotit' tickle
po|ščelkat' crack
poščelkivat' click, snap
po|ščipat' pluck, nibble
poščipyvat' pluck, nibble (intermittently)
po|ščupat' feel, touch, probe

Type II verbs (119)

pobedit' / pobeždat' conquer; overcome
pobudit' / pobuždat' arouse, incite
pobudit'sja / pobuždat'sja be aroused, incited
povedat' / povedyvat' inform
povergnut' / povergat' throw down, plunge
poverit' / poverjat' check, verify
povernut' / povoračivat' turn
povernut'sja / povorotit'sja / povoračivat'sja turn (oneself)
poverit'sja / poverjat'sja be checked, verified
povinovat'sja obey
povisnut' / povisat' hang (over), droop (over)
povesti / povodit' move (body part)
povredit' / povreždat' harm, damage, injure
povredit'sja / povreždat'sja be damaged, injured
povtorit' / povtorjat' repeat, review
povtorit'sja / povtorjat'sja repeat, review
povysit' / povyšat' raise; improve; heighten
povysit'sja / povyšat'sja rise
povjazat' / povjazyvat' tie; catch, nab, apprehend
pogasnut' / pogasat' go out, be extinguished
pogasit' / pogašat' pay off, cancel
pogibnut' / pogibat' die
poglotit' / pogloščat' absorb; gulp down; devour
pogresti / pogrebat' bury
pogruzit' / pogružat' submerse, immerse
pogruzit'sja / pogružat'sja be submersed, immersed
pogrjaznut' / pogrjazat' get stuck in, bogged down
podat' / podavat' submit, give
podat'sja / podavat'sja be submitted, given

podavit' / podavljat' suppress, repress, depress
podavit'sja / podavljat'sja be suppressed, repressed
podvinut' / podvigat' push, shove
podvinut'sja / podvigat'sja move forward, advance
podernut'sja / podergivat'sja be covered (with)
požat' / požinat' reap
pozabyt' / pozabyvat' forget (all about)
pozvolit' / pozvoljat' allow
pozvolit'sja / pozvoljat'sja be allowed
pozdravit' / pozdravljat' congratulate
poznat' / poznavat' get to know, learn
pokazat' / pokazyvat' show
pokazat'sja / pokazyvat'sja be shown
pokinut' / pokidat' leave, abandon
poklonit'sja / poklonjat'sja worship, bow
pokorit' / pokorjat' conquer, subjugate
pokorit'sja / pokorjat'sja be conquered, subjugated
pokryt'sja / pokryvat'sja be covered
pokusit'sja / pokušat'sja attempt; encroach
položit' / polagat' believe, think, suppose
položit'sja / polagat'sja pin (hopes) on; rely on
polučit' / polučat' receive, get
polučit'sja / polučat'sja turn out
polyxnut' / polyxat' blaze
pomeret' / pomirat' die
pomerznut' / pomerzat' be frostbitten
pomestit' / pomeščat' place, put
pomestit'sja / pomeščat'sja be housed; fit (into)
pometit' / pomečat' mark (with); date
pomjanut' / pominat' mention
pomjanut'sja / pominat'sja be mentioned
pomoč' / pomogat' help
pomyknut', pomknut' / pomykat' order about
pomyslit' / pomyšljat' think (about), dream (of)
ponizit' / ponižat' lower, reduce; demote
ponizit'sja / ponižat'sja fall, go down, drop
ponjat' / ponimat' understand
pooščrit' / pooščrjat' encourage
pooščrit'sja / pooščrjat'sja be encouraged
popast' / popadat' hit; get to; find oneself
popast'sja / popadat'sja be caught; come across
poprat' / popirat' trample, scorn
poprat'sja / popirat'sja be trampled, pressed
popolnit' / popolnjat' fill up; replenish
popolnit'sja / popolnjat'sja increase; be replenished
popravit' / popravljat' improve, set right
popravit'sja / popravljat'sja get well; gain weight
porabotit' / poraboščat' enslave, enthrall
porazit' / poražat' strike, hit
porazit'sja / poražat'sja be surprised
porasti / porastat' become overgrown
porvat' / poryvat' break off (with); desert
porodit' / porožđat' give birth to, engender, give rise to
porodit'sja / porožđat'sja be born, be produced
poručit' / poručat' entrust, commission
poručit'sja / poručat'sja be entrusted, charged with
posvjatit' / posvjaščat' devote, dedicate
posvjastit'sja / posvjaščat'sja be devoted to
posetit' / poseščat' visit, attend
poskol'znut'sja / poskal'zyvat'sja slip, trip
posobit' / posobljat' help, relieve
pospet' / pospevat' I ripen
pospet' / pospevat' II be in time; hurry

postavit' / postavljat' supply, provide
postavit'sja / postavljat'sja be supplied, purveyed
postanovit' / postanovljat' decree, enact, decide
postič', postignut' / postigat' understand; overtake, befall
postignut'sja, postič'sja / postigat'sja be understood; be overtaken
postrič'sja / postrigat'sja take monastic vows
postupit' / postupat' act, behave
postupit'sja / postupat'sja waive, forgo
posjagnut' / posjagat' encroach, infringe
potaknut' / potakat' indulge
potrafit' / potrafljat' please
potrebit' / potrebljat' consume, use
potrjasti / potrjasat' shake, brandish
potupit' / potupljat' cast down (gaze)
potupit'sja / potupljat'sja cast down one's gaze
potjanut'sja / potjagivat'sja stretch (one's body)
poxitit' / poxiščat' abduct
počat' / počinat' start
počerpnut' / počerpat' draw, take (fig.)
počit' / počivat' sleep
počest' / počitat' esteem, respect
počest'sja / počitat'sja be honored, esteemed
pojavit'sja / pojavljat'sja appear
pojasnit' / pojasnjat' explain
pojasnit'sja / pojasnjat'sja be explained
poleč' / polegat' lie down
ponudit' / ponuždat' force, compel

Type III verbs (30)

po|byt' / pobyvat' be (at), visit
po|velet' / povelevat' order, command
poglotit' / pogloščat', glotat' absorb; gulp down; devour
poglotit'sja / pogloščat'sja, glotat'sja be absorbed; be gulped down; be devoured
pogrešit' / pogrešat', grešit' sin
podoxnut' / podyxat', doxnut' die, croak
poest' / poedat', est' eat up / through
požat' / požimat', žat' press, squeeze
pokryt' / pokryvat', kryt' cover
polit' / polivat', lit' pour
polit'sja / polivat'sja, lit'sja be poured
pomnožit' / pomnožat', množit' multiply (by)
poniknut' / ponikat', niknut' droop; wilt
poprostit'sja / poproščat'sja, proščat'sja say good-bye, take leave
posvistet', svistet' / posvistat', svistat' whistle
poselit' / poseljat', selit' settle, lodge
poselit'sja / poseljat'sja, selit'sja settle, lodge
poslat' / posylat', slat' send
poslat'sja / posylat'sja, slat'sja be sent
pospešit' / pospešat', spešit' hurry
postelit' / postilat', stelit' spread, lay
postelit'sja / postilat'sja, stelit'sja be spread, laid
po|strič', postrigat' cut; mow
po|strič'sja, postrigat'sja get hair cut
posýpat' / posypát', sypat' strew, sprinkle
poteret' / potirat', teret' rub (delim)
potešit' / potešat', tešit' amuse, entertain
poxvalit'sja / poxvaljat'sja, xvalit'sja boast
počinit' / počinjat', činit' repair, mend
pobit' / pobivat', bit' stone to death

Type IV verbs (24)

pogodit'	wait (a while)
podvignut'	rouse (to)
podvizat'sja	work, act
po|divit'sja	be astonished, marvel
podirat'	frozen phrase: give goose bumps
podobat'	befit, become
požívat'	live
požívit'sja	profit; get a hold (of)
pokladat'	lay, put (?)
pokrovitel'stvovat'	protect
polagat'sja I	be supposed to, be appropriate
polagat'sja II	be thought, considered
ponadobit'sja	be(come) necessary
ponimat'sja	be understood
ponosit'	defame, slander
popustitel'stvovat'	wink (at), turn blind eye (to)
poricat'	blame for; reproach
poryvat'sja	try, endeavor
posredničat'	mediate
potvorstvovat'	pander
potemnjat'	make dark
poučat'	teach, lecture, preach
poxodit' II	resemble
počtit'	honor

Type V verbs (3)

kupit' / pokupat'	buy
kupit'sja / pokupat'sja	be bought
po|nukat'	drive on, urge on

APPENDIX 4: Verbs in sample with glosses & frequency counts

This appendix lists the 234 verbs that occur in the 1000-observation sample, on which the statistical analyses in this dissertation are based. Verbs are given in order of frequency, from most frequent to least frequent. A gloss and frequency count is given for each verb.

Verb	**Gloss**	**Frequency**
по|йти	(start to) go (on foot)	86
по|ставить	put, place	63
по|смотреть	look	54
по|строить	build	37
по|думать	think	31
по|терять	lose	28
по|просить	ask	24
по|ехать	(start to) drive	21
по|нравиться	like	19
по|пробовать	try	19
по|чувствовать	feel	19
по|звонить	call (telephone)	15
положить / класть	lay, place	15
по|менять	change	15
по|говорить	talk	14
посадить / сажать	seat, plant	14
по|дарить	give as a gift	13
по|пытаться	try, attempt, endeavor	13
по|стараться	try, endeavor	13
по|требовать	require	13
по|требоваться	be required	12
по|верить	believe	11
по|знакомиться	be introduced, get to know	9
по|любить	(fall in) love	9
по|следовать	follow	9
по|влиять	influence	8
по|делиться	divide, share	8
по|мешать	bother, hinder, prevent	8
по|обещать	promise	8
по|считать	count; think	8
по|интересоваться	be interested	7
по|страдать	suffer	7
по|тратить	spend; waste	7

Verb	Gloss	Frequency
повесить / вешать	hang; pin, throw (on) (lit. & fig.)	6
по|нести	carry, bear	6
по|работать	work	6
по|советовать	advise	6
по|везти I	carry, transport	5
по|жертвовать	donate, sacrifice	5
по|лететь	(start to) fly	5
по|служить	serve	5
по|слушать	listen	5
по|хоронить	bury	5
по|щупать	feel, touch, probe	5
по|бороться	struggle, wrestle	4
по|везти II	be lucky	4
по|вести	begin to lead/conduct	4
по|жаловать	grant, bestow; visit	4
по|завидовать	envy	4
по|качать	rock, shake, wave	4
по|лезть	(start to) climb	4
по|любоваться	admire, feast eyes on/upon	4
по|общаться	communicate	4
по|радовать	make happy	4
по|сыпаться	rain, pour	4
по|терпеть I	suffer	4
по|целовать	kiss	4
по|бежать	(start to) run	3
по|бить	hit	3
по|гладить	stroke, caress	3
по|гулять	take a walk; have a good time	3
по|делить	divide, share	3
по|играть	play	3
поймать / ловить	catch	3
по|катиться	(start to) roll, drive	3
по|кривить	bend, twist, distort	3
помалкивать	hold one's tongue	3
по|молчать	be quiet	3
по|радоваться	be glad, rejoice	3
посматривать	look (inter-atten)	3
по|стоять	stand	3
по|теснить	press, crowd	3

Verb	Gloss	Frequency
по|тянуться	move toward; stretch, lie (fig.)	3
по|участвовать	participate	3
по|читать	read	3
по|благодарить	thank	2
повеситься / вешаться	hang oneself	2
по|веять	begin to blow; "one could sense" (impers.)	2
повизгивать	cry shrilly	2
по|влечь	draw, drag, pull, attract	2
по|временить	wait	2
по|гнать	(start to) drive	2
по|гостить	stay at, visit, be a guest	2
по|губить	destroy, ruin	2
по|драться	fight (over, with)	2
по|дружиться	make friends (with)	2
по|жалеть	be sorry for, pity	2
по|желать	wish, desire	2
по|жениться	get married	2
по|заботиться	look after, care for, be concerned about	2
по|звать	call (for)	2
по|знакомить	introduce	2
по|искать	look, search	2
по|калечить	cripple, maim; destroy, ruin	2
по|клясться	swear, vow	2
по|колебаться	be shaken	2
по|курить	smoke	2
по|любиться	be loved; catch someone's fancy	2
по|молиться	pray	2
по|мрачнеть	become gloomy, darken	2
по|настроить	build in quantity	2
по|размыслить, по|размышлять	think, consider	2
по|сидеть	sit	2
по|советоваться	consult	2
по|солить	salt; pickle	2
по|спорить	argue	2
по|стареть	grow old	2
по|теряться	be lost	2

Verb	Gloss	Frequency
по|течь	(start to) flow	2
по|трепать	pull about, tousle	2
по|тянуть	pull	2
по|увечить	maim, cripple	2
по|шатнуть	bend; shake	2
по|шутить	joke	2
побаиваться	be rather afraid	1
по|бегать	run (around)	1
по|бледнеть	turn pale	1
по|болтать	chat	1
по|бороть	fight down, overcome	1
по|брезговать	be squeamish about	1
по|бриться	shave	1
по|бродить	stroll (around)	1
по|бросать	throw	1
побухивать	thunder, thud	1
по|валить	fall/throw down; pile up	1
по|валяться	wallow	1
по|варить	roast, cook	1
по|вдоветь	widow, be widow/er	1
по|венчаться	get married	1
по|вертеть	rotate, twist	1
по|веселеть	cheer up; become merry/cheerful	1
по|вздыхать	sigh, breathe; long for, yearn for	1
по|видать	see	1
по|висеть	hang	1
по|воевать	wage war, fight; quarrel, fight	1
по|высыпать	empty, pour, spill out	1
по|вязать	knit, tie, bind	1
по|гадать	tell a fortune	1
по|глазеть	stare, gawk (at)	1
по|глядеть	look, glance	1
поглядывать	look (intermittently)	1
поговаривать	talk a little, intermittently	1
по|грозить	make a threatening hand gesture	1
по|грохотать	crash, roar, rumble	1
по|двигать	move	1
по|держать	hold	1
по|дорожать	rise in price, go up	1

Verb	Gloss	Frequency
по|думаться	be thought	1
поеживаться	shiver, shudder	1
по|ездить	drive around	1
по|жарить	cook, roast	1
по|жить	live	1
по|завтракать	have breakfast	1
позвякивать	tinkle	1
по|здороваться	greet	1
по|исповедаться	confess	1
по|истереться	grate, use up by rubbing	1
по|караулить	guard, keep watch (on, for)	1
покачиваться	rock (intermittently)	1
по|каяться	repent, confess	1
по|класть	put, lay	1
поклониться / кланяться	bow (before)	1
по|коситься	become lopsided; glance sidelong	1
по|красить	paint, dye; make-up	1
по|краснеть	turn red	1
покрикивать	cry, shout; shout at, chide	1
по|критиковать	criticize	1
по|куражиться	mock, jeer	1
по|кусать	bite	1
по|кушать	eat	1
по|ладить	come to an understanding, get along	1
по|лакомиться	regale oneself with	1
по|лениться	be lazy	1
по|ломать	break	1
по|льститься	be tempted (with, by)	1
по|мазать	daub, spread	1
помахивать	wave	1
по|меняться	change	1
по|мериться	measure, be measured	1
по|меркнуть	fade; grow dim	1
по|мириться	be reconciled	1
по|морщиться	wrinkle one's nose, knit one's brow	1
по|мылить	steal	1
по|набежать	rush in, run in	1

Verb	Gloss	Frequency
по|написать	write a lot	1
по|наставить	place in quantity	1
по|нестись	rush after, dart after	1
попахивать	smell	1
по|петь	sing	1
попивать	drink (intermittently)	1
по|пить	drink	1
по|платиться	pay (for…with…)	1
по|ползти	(start to) crawl	1
по|портить	spoil	1
по|практиковаться	practice	1
по|проситься	ask; yearn for	1
по|прятаться	hide	1
по|пытать	try, test	1
поразмахивать	swing	1
по|расспросить	ask around	1
по|рваться	be torn	1
по|редеть	thin out; be depleted	1
по|резать I	cut	1
по|рекомендовать	recommend	1
по|решить	decide	1
по|рушить	destroy	1
по|светлеть	become lighter, visible	1
по|седеть	turn gray	1
по|сеять	sow	1
посиживать	sit (intermittently)	1
по|скучать	miss / be bored	1
по|слюнить	wet with saliva	1
по|смеяться	laugh	1
по|совещаться	deliberate	1
по|сочувствовать	sympathize	1
по|спать	sleep	1
по|способствовать	assist, aid	1
по|ссориться	argue, fight (with)	1
постанывать	moan	1
по|стесняться	be shy, ashamed	1
по|стрелять	shoot	1
по|топать	hit the road	1
по|торопиться	be in a hurry	1

Verb	Gloss	Frequency
по|трескаться	crack, chap	1
потрясывать	shake (inter-atten)	1
потягивать	sip	1
по|ужинать	have dinner	1
похаживать	walk (intermittently)	1
по|хвалить	praise	1
по|хвастать	brag, boast	1
по|хлопать	slap, pat	1
по|ходатайствовать	solicit, intercede	1
почитываться	be read	1
по|шалить	play pranks	1

APPENDIX 5: ID tags and ID tag levels

This appendix provides a full list of all ID tags and ID tag values/levels used in the analysis outlined in Chapters 3 and 4. Where appropriate, brief examples of ID tag levels are given.

Variable name	**Levels (variable values)**
Transitivity	intransitive
	transitive
Voice	active
	middle
	passive
Tense-mood	gerund
	imperative
	indicative-future
	indicative-past
	indicative-present
	infinitve
	participle-past
	subjunctive
Semantic type of the verb	abstract action
	change of state
	existence/being/relations
	human qualities/behavior
	impact/contact/support
	location/placement
	mental/psychological/emotional
	movement
	natural phenomena
	perception
	physiology
	possession
	speech/text
	things sensed: sounds, light, smells
Negation	clause negated
	positive
	preceding verb negated
	verb negated
	ne-word (*negde, nekogda*, etc.)

Sentence type	declarative
	exclamation (not imperative)
	imperative
	interrogative
Clause type	dependent
	independent
Dependent clause type	adjectival/appositive
	gerundial
	adverbial: manner
	NA
	adverbial: spatial
	adverbial: temporal
	esli clause
	relative
	čto-clause
	čtoby-clause
Text type	fiction
	nonfiction
	spoken
Adverbials & phrases with adverbial function	certainty
	duration (or accusative of time)
	futility (*zrja, naprasno*)
	intensity / degree (*ochen', stol'*)
	NA
	necessity (*objazatel'no*)
	other (those not subsumed under other levels)
	quantity of time (*dolgo, za pjat' minut*)
	restriction/limitation (*čut', edva*)
	location: adverbials answering *gde?*
	manner: adverbials answering *kak?*
	temporal: adverbials answering *kogda?*
	motion: adverbials answering *kuda? otkuda?*
	možno
	nado/nužno
	nel'zja
	causal: adverbials answering *počemu?*
Particles	exhortation (*pust'*, *davaj(te)*)
	intensification (*daže*)
	NA
	other particles not subsumed under other levels

Particles	restriction (*tol'ko*)
	untimely halt (*bylo*)
	xot'
	xotja
Objects: Syntactic type	accusative
	dative
	genitive
	infinitive
	instrumental
	NA
	other clause
	prepositional phrase
	prepositional phrase + *čto*-clause
	prep phrase + *čtoby*-clause
	clause without subordinator
	čto-clause
	čtoby-clause
Objects: Animacy	animate
	inanimate
Objects: Abstract vs. concrete	abstract
	concrete
Objects: Count vs. mass	count
	mass
Objects: Number	plural
	singular
Objects: Semantic type	(social) events
	abstract action
	animals
	change of state
	existence/being/relations
	facts/ideas
	human qualities/behavior
	humans or supernatural beings
	impact/contact/support
	large-scale abstractions involving humans (institutions, governments, nations, etc.)
	location/placement
	mental/psychological/emotional
	movement
	natural phenomena
	parameters, parts, measures
	perception

Objects: Semantic type	physiology plants possession qualities (small) groups of humans spaces/places speech/text things things sensed: sounds, light, smells time
Prepositions associated with the verb	NA *bez* *v* *v kačestve* *vdol'* *vsledstvie* *vmesto* *dlja* *do* *za* *iz* *iz-za* *k* *meždu* *na* *nad* *o* *okolo* *o* *pered* *po* *pod* *posle* *pri* *pro* *radi* *s* *sredi* *u* *čerez*
Participles: Animacy	animate inanimate

Participles: Abstract vs. concrete	abstract
	concrete
Participles: Count vs. mass	count
	mass
Participles: Number	plural
	singular
Participles: Semantic type	(social) events
	abstract action
	animals
	existence/being/relations
	facts/ideas
	humans or supernatural beings
	large-scale abstractions involving humans (institutions, governments, nations, etc.)
	mental/psychological/emotional
	possession
	sets, aggregates, or groups
	(small) groups of humans
	spaces/places
	speech/text
	things
Subjects: Syntactic type	accusative to an impersonal verb
	adjective + infinitive (*gotov sdelat' čto?*)
	adverbial + infinitive (*vredno delat' čto?*)
	dative to preceding verb or impersonal construction
	dative + impersonal verb
	dative + personal verb
	NA
	nominative implied
	nominative + infinitive
	nominative to the preceding verb
	nominative to the verb
	nado/nužno/možno/nel'zja + infinitive
	impersonal verb (no subject)
	infinitive to impersonal verb (no subject)
Subjects: Animacy	animate
	inanimate
Subjects: Abstract vs. concrete	abstract
	concrete

Subjects: Count vs. mass	count
	mass
Subjects: Number	plural
	singular
Subjects: Semantic type	(social) events
	abstract action
	animals
	existence/being/relations
	facts/ideas
	human qualities/behavior
	humans or supernatural beings
	impact/contact/support
	large-scale abstractions involving humans (institutions, governments, nations, etc.)
	mental/psychological/emotional
	NA
	natural phenomena
	possession
	qualities
	(small) groups of humans
	spaces/places
	speech/text
	things
	things sensed: sounds, light, smells
	time

REFERENCES

Agrell, Sigurd. 1908. *Aspektänderung und Aktionsartbildung beim polnischen Zeitworte: ein Beitrag zum Studium der indogermanischen Präverbia und ihrer Bedeutungsfunktionen*. Lund: Ohlsson.

Andrews, E., G. Averyanova, and G. Pyadusova. 2004. *The Russian Verb: Form and Function*. Moscow: Russkij Jazyk.

Anstatt, Tanja. 2002. Die Quantelung des zweiten Arguments im Russischen: Der Typus *s''est' jabloko – poest' supu*. In: Thomas Daiber (ed.). *Linguistische Beiträge zur Slavistik IX.* Munich: Kubon und Sagner. 7-30.

Arppe, Antti. 2008. *Univariate, bivariate, and multivariate methods in corpus-based lexicography – a study of synonymy.* Doctoral dissertation. University of Helsinki.

Atkins, B. T. S. 1978. Semantic ID tags: corpus evidence for dictionary senses. *Proceedings of the Third Annual Conference of the UW Centre for the New Oxford English Dictionary.* 17-36.

Axmanova, O. S. 1974. *Slovar' omonimov russkogo jazyka*. Moskva: Sovetskaja Enciklopedija.

Baayen, R. H. (in press) *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R*. Cambridge University Press.

Berez, Andrea L. and Stefan Th. Gries. 2009. In defense of corpus-based methods: a behavioral profile analysis of polysemous *get* in English. In: Steven Moran, Darren S. Tanner, and Michael Scanlon (eds.). *Proceedings of the 24th Northwest Linguistics Conference*. University of Washington Working Papers in Linguistics. Vol. 27. Seattle: Department of Linguistics. 157-66.

Bertinetto, Pier Marco and Denis Delfitto. 2000. Aspect vs. actionality: Why they should be kept apart. In: Östen Dahl (ed.). *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter. 189-225.

Biber, Douglas. 1993. Co-occurrence Patterns among Collocations: A Tool for Corpus-Based Lexical Knowledge Acquisition. *Computational Linguistics*. 19(3): 531-538.

Bierwisch, M. and R. Schreuder. 1992. From concepts to lexical items. *Cognition*. 42: 23-60.

Binnick, Robert I. 1991. *Time and the verb: A guide to tense and aspect*. New York/Oxford: Oxford University Press.

Bogusławski, Andrzej. 1963. *Prefiksacja czasownikowa we współczesnym języku rosyjskim*. Wrocław: Wydanctwo Polskiej Akademii Nauk.

Bondarko, Aleksandr V. 1983. *Principy funkcional'noj grammatiki i voprosy aspektologii*. Leningrad.

Brugman, Claudia. 1981. *The story of* over. M.A. thesis. University of California, Berkeley.

Buju-Bijunas (Vaitiekunaite), Genovaite. 1978. *The lexical meanings of the Lithuanian* per/pra- *and the Russian* pere-/pro- *verbal prefixes*. Doctoral dissertation. McGill University.

Bybee, Joan and Paul Hopper. 2001. Introduction to Frequency and the Emergence of Linguistic Structure. In: Joan Bybee and Paul Hopper (eds.). *Frequency and the Emergence of Linguistic Structure.* Amsterdam: John Benjamins. 1-26.

Bybee, Joan, William Pagliuca, and Revere Perkins. 1994. *The Evolution of Grammar*. University of Chicago.

Bybee, Joan and Carol Lynn Moder. 1983. Morphological Classes as Natural Categories. *Language*. Baltimore: Waverly Press. 15(2): 251-270.

Camus, Remi. 1998. Quelques considérations sur le préverbe *po-* en russe contemporain. *Revue des Études Slaves* LXX (1): 101-112.

Carston, R. 1996. The architecture of the mind: modularity and modularization. In: D. Green et al. (eds.). *Cognitive Science: An Introduction*. Cambridge: Blackwell. 53-83.

Čertkova, Marina. 1996. *Grammatičeskaja kategorija vida v sovremennom russkom jazyke*. Moskva: Moskovskij Gosudarstvennyj Universitet.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* MIT Press.

Chomsky, Noam. 1964. Formal Discussion. In: U. Bellugi and R. Brown (eds.). *The Acquisition of Language*. Monographs of the Society for Research in Child Development 29. University of Chicago. 37-39.

Coltheart, M. 1999. Modularity and cognition. *Trends in Cognitive Sciences.* 3: 115-120.

Comrie, Bernard. 1981. *Aspect*. New York: University of Cambridge.

Croft, William and D. Alan Cruise. 2004. *Cognitive Linguistics*. New York: Cambridge University Press.

Cuyckens, Hubert. 1993. The Dutch Spatial Preposition 'in': A Cognitive Semantic Analysis. In: Cornelia Zelinsky-Wibbelt (ed.). *The Semantics of Prepositions: From Mental Processing to Natural Language Processing.* Berlin/New York: Mouton de Gruyter. 27-71.

Dąbrowska, Eva. 2004. *Language, Mind, and Brain.* Washington: Georgetown University Press.

Dahl, Östen. 1985. *Tense and Aspect Systems*. New York: Basil Blackwell, Inc.

Deane, Paul D. 1993. Polysemy as the Consequence of Internal Conceptual Complexity: The Case of *Over*. *Proceedings of the Eastern States Conference on Linguistics*. 9: 32-43.

Deese, J. 1965. *The structure of associations in language and thought.* Baltimore: John Hopkins Press.

Dickey, Stephen. 2007. A Prototype Account of the Development of Delimitative *PO-* in Russian. In: Dagmar Divjak and Agata Kochanska (eds.).*Cognitive Paths into the Slavic Domain*. Berlin: Mouton de Gruyter. 326-371.

Dickey, Stephen. 2006. Aspectual Pairs, Goal Orientation, and *po-* Delimitatives in Russian. *Glossos* 7. http://www.seelrc.org/glossos/

Dickey, Stephen M. 2000. *Parameters of Slavic Aspect: A Cognitive Approach*. Stanford: CSLI Publications.

Dickey, Stephen M. and Julie Hutcheson. 2003. Delimitative Verbs in Russian, Czech, and Slavic. *American Contributions to the 13th International Congress of Slavists.* Vol. 1. Bloomington: Slavica. 23-36.

Dickey, Stephen M. and Laura A. Janda. 2009. *Xoxotnul*, *sxitril*: The relationship between semelfactives formed with *nu-* and *s-* in Russian. *Russian Linguistics.* 33(3): 229-248.

Divjak, Dagmar. 2010. *Structuring the Lexicon: a Clustered Model for Near-Synonymy*. Berlin/New York: Mouton de Gruyter.

Divjak, Dagmar. 2004. *Degrees of Verb Integration: Conceptualization and Categorization of Events in Russian*. Doctoral dissertation. Katholieke Universiteit – Leuven, Belgium.

Divjak, Dagmar and St. Th. Gries. 2008. Clusters in the Mind? Converging evidence from near-synonymy in Russian. *The Mental Lexicon.* 3(2): 188-213.

Divjak, Dagmar and St. Th. Gries. 2006 . Ways of Trying in Russian: Clustering Behavioral Profiles. *Corpus Linguistics and Linguistic Theory.* 2(1): 23- 60.

Dmitrieva, Ol'ga. 2000. Formirovanie sistemy russkix delimitativnyx glagolov. *Predloženie i slovo: paradigmatičeskij, tekstovyj i kommunikativnyj aspekty*. Saratov: Izdatel'stvo Saratovskogo Pedagogičeskogo Instituta. 28-33.

Dmitrieva, Ol'ga. 1991. Formirovanie semantičeskoj struktury russkogo glagol'nogo prefiksa *po-*. *Aktivnye processy v jazyke i reči.* Saratov: Izdatel'stvo Saratovskogo Universiteta. 68-74.

Durkin, Kevin and Jocelyn Manning. 1989. Polysemy and the subjective lexicon: semantic relatedness and the salience of intraword senses. *Journal of Psycholinguistic Research.* 18(6): 577-612.

Ervin, Gerard L., Donald K. Jarvis, Sophia Lubensky, and Larry McLellan. 2002. *Načalo*. 2nd edition. McGraw-Hill.

Evans-Romaine, Karen, Joanna Robin, Richard Robin, and Galina Shatalina. 2007. *Golosa: A Basic Course in Russian*. Upper Sadle River, New Jersey: Prentice Hall.

Feldman, Jerome A. 2008. *From Molecule to Metaphor: A Neural Theory of Language.* Cambridge, Massachusetts: MIT Press

Firth, John R. 1957. A synopsis of linguistic theory 1930 – 1955. In Philological Society (eds.). *Studies in Linguistic Analysis.* Oxford: Blackwell. 1-32.

Flank, Sharon. 1995. *Aspectualizers and scope*. Doctoral dissertation. Harvard University.

Flier, Michael S. 1985. The Scope of Prefixal Delimitation in Russian. In: Michael S. Flier and Alan Timberlake (eds.). *The Scope of Slavic Aspect*. Columbus, Ohio: Slavica. 41-58.

Flier, Michael S. 1984. Syntagmatic Constraints on the Russian Prefix *pere-*. In: Michael S. Flier and Richard D. Brecht (eds.). *Issues in Russian Morphosyntax. UCLA Slavic Studies.* Vol. 10. Columbus, Ohio: Slavica. 138-154.

Flier, Michael S. 1975. Remarks on Russian Verbal Prefixation. *Slavic and East European Journal.* 19: 218-229.

Flombaum, J. I., L. R. Santos, and M. D. Hauser. 2002. Neuroecology and psychological modularity. *Trends in Cognitive Sciences.* 6(3): 106-108.

Fodor, Jerry A. 1983. *Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, Massachusetts: MIT Press.

Fodor, J.D., J.A. Fodor, and M. Garrett. 1975. The unreality of semantic representations. *Linguistic Inquiry.* 4: 515-531.

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

Frairie, Susan. 1992. *Animacy in Czech and Russian*. Master's thesis. University of North Carolina at Chapel Hill.

Gallant, Charles James III. 1977. *Russian verbal prefixation and semantic features: An analysis of the prefix* vz-. Doctoral dissertation. University of California, Los Angeles.

Gavrilova, V. L. and R. M. Teremova. 2001. *Okno v mir russkoj reči.* Sankt-Peterburg: Izdatel'stvo "Borej".

Geeraerts, Dirk. 1997. *Diachronic Prototype Semantics.* Oxford: Clarendon Press.

Geeraerts, D. 1988. Where does prototypicality come from? In: B. Rudzka-Ostyn (ed.). *Topics in Cognitive Linguistics*. Amsterdam: John Benjamins. 207-229.

Geiger, Richard A. and Brygida Rudzka-Ostyn (eds.). *Conceptualizations and mental processing in language*. Berlin/New York: Mouton de Gruyter.

Gibbs, Raymond W. Jr. and Marcus Perlman. 2006. The contested impact of cognitive linguistic research on the psycholinguistics of metaphor understanding. In: Gitte Kristiansen, Michel Achard, René Dirven, and Francisco J. Ruiz (eds.). *Cognitive Linguistics:Current Applications and Future Perspectives.* Mouton de Gruyter. 211-228.

Gilquin, Gaëtanelle. 2007. To err is not all: What corpus and elicitation can reveal about the use of collocations by learners. *Zeitschrift für Anglistik und Amerikanistik.* 55(3): 273-291.

Gilquin, Gaëtanelle. 2006. The place of prototypicality in corpus linguistics: Causation in the hot seat. In: Stefan Th. Gries and Anatol Stefanowitsch (eds.). *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*. Berlin, Heidelberg, New York: Mouton de Gruyter. 159-191.

Goedsche, C. R. 1940. Aspect versus Aktionsart. *Journal of English and Germanic Philology*. 39: 189-197.

Goldberg, Adele. 2006. *Constructions at Work: The Nature of Generalizations in Language.* Oxford: Oxford University Press.

Goossens, Louis et al. 1995. *By word of mouth: metaphor, metonymy and linguistic action in a cognitive perspective*. Amsterdam/Philadelphia: John Benjamins.

Grammatika russkogo jazyka. 1960. Akademija Nauk SSSR, Institut Russkogo Jazyka. Moskva: Izdatel'stvo Akademii Nauk SSSR.

Gries, Stefan Th. 2006. Corpus-based methods and cognitive semantics: the many meanings of *to run*. In: Stefan Th. Gries and Anatol Stefanowitsch (eds.).*Corpora in cognitive linguistics: corpus-based approaches to syntax and lexis*. Berlin/New York: Mouton de Gruyter. 57-99.

Gries, Stefan Th. and Dagmar Divjak. 2008. Behavioral profiles: a corpus-based approach to cognitive semantic analysis. In: V. Evans and S. Pourcel (eds.). *New Directions in Cognitive Linguistics.* Amsterdam/Philadelphia: John Benjamins.

Gries, Stefan Th. and Anatol Stefanowitsch. Forthcoming. Cluster analysis and the identification of collexeme classes. In: John Newman and Sally Rice (eds.). *Experimental and empirical methods in the study of conceptual structure, discourse, and language*. Stanford: CSLI. 73-90.

Guiraud-Weber, Marguerite. 1993. Le préverbe *po-* en russe moderne. *La revue russe.* 5: 57-68.

Hanks, P. 1996. Contextual dependency and lexical sets. *International Journal of Corpus Linguistics.* 1(1): 75-98.

Hanson, David Adams. 1974. *Verbal Prefixation in Russian: Aktionsart meaning and the attenuative use of* pri-. Harvard University.

Haspelmath, M. 2006. Against Markedness (and what to replace it with). *Journal of Linguistics.* 1: 25-70.

Haspelmath, M. 1997. *From Space to Time. Temporal Adverbials in the World's Languages.* München: Lincom.

Hauser, M. D., N. Chomsky, and W. T. Fitch. 2002. The faculty of language: What is it, who has it, and how does it evolve? *Science.* 298: 1569-1579.

Hjelmslev, Louis. 1935. *La catégorie des cas: étude de grammaire générale*. Aarhus: Universitetsforlaget.

Holden, Kyril T. 1981. The role of metaphor in defining Russian aspect. *Canadian Slavonic Papers (Revue Canadienne des Slavistes).* 31: 25-43.

Isačenko, A. V. 1962. Die russische Sprache der Gegenwart. Halle, Saale: M. Niemeyer.

Isačenko, A. V. 1960. *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim. Časť vtoraja: morfologija*. Bratislava: Izdatel'stvo akademii nauk.

Jakobson, Roman. 1984. *Russian and Slavic Grammar: Studies 1931 – 1981*. Linda R. Waugh and Morris Halle (eds.). New York: Mouton.

Jakobson, Roman. 1971. *Selected writings* (Vol. II). The Hague: Mouton de Gruyter.

Jakobson, Roman. 1936. Beitrag zur allgemeinen Kasuslehre: Gesamtbedeutungen der russischen Kasus. *Travaux du cercle linguistique de Prague* VI.

Janda, Laura A. Forthcoming. The role of metonymy in grammar: a case study of Czech derivational morphemes. *Slovo a slovesnost*. 17 pp.

Janda, Laura A. 2010a. Cognitive Linguistics in the Year 2010. Forthcoming in the *International Journal of Cognitive Linguistics*. 32 pp.

Janda, Laura A. 2010b. Prefixed Perfectives in Non-determined Motion Verbs in Russian. In : Viktoria Hasko and Renee Perelmutter (eds.). *New Approaches to Slavic Verbs of Motion.* Studies in Language Companion Series 115. Amsterdam/Philadelphia: John Benjamins. 125-140.

Janda, Laura A. 2008a. Metonymy via Perfectivization of Russian Verbs. *Slavica Helsingiensia 35* (= *S ljubov'ju k slovu. Festschrift in Honour of Professor Arto Mustajoki on the Occasion of his 60th Birthday*). 77-85.

Janda, Laura A. 2008b. Semantic Motivations for Clusters of Russian Verbs. In: Christina Y. Bethin (ed.). *American Contributions to the 14th International Congress of Slavists, Ohrid, September 2008*. 181-196.

Janda, Laura A. 2007a. What makes Russian Bi-aspectual verbs special. In: Dagmar Divjak and Agata Kochanska (eds.). *Cognitive Paths into the Slavic Domain. Cognitive Linguistics Research.* Berlin/New York: Mouton de Gruyter. 83-109.

Janda, Laura A. 2007b. Aspectual clusters of Russian verbs. *Studies in Language.* 31(3): 607-648.

Janda, Laura. 2006. Cognitive Linguistics. *Glossos* 8. http://www.seelrc.org/glossos/

Janda, Laura A. 2004. A metaphor in a search of a domain: the categories of Slavic aspect. *Cognitive Linguistics.* 15(4): 471-527.

Janda, Laura A. 2002a. Cases in collision, cases in collusion: the semantic space of case in Czech and Russian. In: Laura A. Janda, Steven Franks, and Ronald Feldstein (eds.). *Where One's Tongue Rules Well: A Festschrift for Charles E. Townsend*. Columbus, Ohio: Slavica. 43-61.

Janda, Laura A. 2002b. Concepts of Case and Time in Slavic. *Glossos* 3. http://www.seelrc.org/glossos/

Janda, Laura A. 2002c. The conceptualization of events and their relationship to time in Russian. *Glossos* 2. http://www.seelrc.org/glossos/

Janda, Laura A. 2000. A cognitive model of the Russian accusative case. In: R. K. Potapova, V. D. Solov'ev, and V. N. Poljakov (eds.). *Trudy meždunarodnoj konferencii Kognitivnoe modelirovanie*, No. 4, Part I. Moscow: MISIS. 20-43.

Janda, Laura A. 1995. Unpacking markedness. In: Eugene Casad (ed.). *Linguistics in the Redwoods: The expansion of a new paradigm in Linguistics*. Berlin: Mouton de Gruyter. 207-233.

Janda, Laura A. 1993. Cognitive linguistics as a continuation of the Jakobsonian tradition: The semantics of Russian and Czech reflexives. In: Robert A. Maguire and Alan Timberlake (eds.). *American Contributions to the Eleventh International Congress of Slavists in Bratislava*. Columbus: Slavica. 310-319.

Janda, Laura A. 1988. The Mapping of Elements of Cognitive Space into Grammatical Relations: An Example from Russian Verbal Prefixation. In: Brygida Rudzka-Ostyn (ed.). *Topics in Cognitive Linguistics.* Amsterdam: John Benjamins. 327-343.

Janda, Laura A. 1986. *A Semantic Analysis of the Russian Verbal Prefixes* za-*,* pere-*,* do-*, and* ot-. Slavistische Beiträge, Band 192. München: Sagner.

Janda, Laura. 1985. The Meaning of Russian Verbal Prefixes: Semantics and Grammar . In: Michael Flier and Adam Timberlake (eds.). *The Scope of Slavic Aspect* (UCLA Slavic Studies, Vol. 12). Columbus, Ohio: Slavica. 26-40.

Janda, Laura A. and Steven J. Clancy. 2006. *The Case Book for Czech.* Bloomington, Indiana: Slavica.

Janda, Laura A. and Steven J. Clancy. 2002. *The Case Book for Russian*. Bloomington, Indiana: Slavica.

Janda, Laura A. and John Korba. 2008. Beyond the pair: Aspectual derivation for learners of Russian. *Slavic and East European Journal*. 52(2): 257-270.

Janda, Laura A. and Tore Nesset. Forthcoming. Taking Apart Russian *RAZ-*. *Slavic and East European Journal.* 28 pp.

Janda, Laura A. and Valery Solovyev. 2009. What Constructional Profiles Reveal About Synonymy: A Case Study of Russian Words for SADNESS and HAPPINESS. *Cognitive Linguistics* 20(2): 367-393.

Johnson, Keith. 2008. *Quantitative Methods in Linguistics*. Malden, Massachusetts: Blackwell.

Jonas, Libuse. 1976. *The semantics of Russian and Czech verbal prefixes: A comparison.* Doctoral dissertation. University of Alberta.

Kagan, Olga, Ganna Kudyma, and Frank Miller. 2006. *V puti: Russian Grammar in Context.* 2nd edition. Upper Saddle River, New Jersey: Prentice Hall.

Kahn, H. J., and Whitaker, H. A. 1991. Acalculia: an historical review of localization. *Brain Cognition*. 17: 102-115.

Kaufman, Leonard and Peter J. Rousseeuw. 1990. *Finding Groups in Data*. New York: John Wiley.

Kilgarriff, Adam and Gregory Grefenstette. 2003. Introduction to the Special Edition on the Web as a Corpus. *Computational Linguistics.* 29(3): 333-347.

Kiss, George R. 1973. Grammatical Word Classes: A Learning Process and its Simulation. *Psychology of Learning and Motivation*. 7: 1-41.

Kisseleva, Xenia and Sergei Tatevosov. 2004. Incrementality and the Delimitative. Presentation at the Perspectives on Slavistics conference. K. U. Leuven, Belgium. 18 September 2004.

Kolesnikov, N. P. 1995. *Slovar' omonimov russkogo jazyka*. Rostov-na-Donu: Feniks.

Komárek, M. 1984. Prefixace a slovesný vid: K prefixům prostě vidovým a subsumpci. *Slovo a slovesnost.* 45(4): 257-67.

Kövecses, Zoltán. 2002. *Metaphor: A practical introduction.* Oxford University Press.

Krifka, M. 1998. The origins of telicity. In: S. Rothstein (ed.). *Events and Grammar*. Dordrecht: Kluwer Academic Publishers. 197-235.

Krongauz, M. A. 1998. *Pristavki i glagoly v russkom jazyke: semantičeskaja grammatika*. Moskva: Jazyki russkoj kul'tury.

Labov, W. 1973. The boundaries of words and their meanings. In: C.-J. N. Bailey and R. W. Shuy (eds.). *New Ways of Analysing Variation in English.* Washington: Georgetown University Press. 340-373.

Labov, W.1969 (reprinted 1992). The logic of non-standard English. *Georgetown Monographs on Language and Linguistics 22*. Reprinted in P.P. Giglioli (ed.). *Language and Social Context*. London: Penguin.

Lakoff, George. 1990. The invariance hypothesis: Is abstract reason based on image schemas? *Cognitive Linguistics* 1: 39-74.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. University of Chicago Press.

Lakoff, George and Mark Johnson. 1980. *Metaphors We Live By.* University of Chicago Press.

Langacker, Ronald W. 2006. On the continuous debate about discreteness. *Cognitive Linguistics.* 17(1): 107-151.

Langacker, Ronald W. 2003. Constructional Integration, Grammaticization, and Serial Verb Constructions. *Language and Linguistics.* 4: 251-278.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar. Vol. I. Theoretical Prerequisites.* Stanford University Press.

LeBlanc, Nicholas. 2007. The Polysemy of an 'Empty' Prefix: Towards a cognitive network of Russian *po-*. Presentation at the 10th International Cognitive Linguistics Conference. Krakow, Poland. July 2007.

LeBlanc, Nicholas. 2006. Iterativity: Re-visiting and rethinking secondary homogeneity in Russian delimitatives. *Glossos* 7. http://www.seelrc.org/glossos/

Leech, G. 1992. Corpora and theories of linguistic performance. In: J. Svartik (ed.). *Directions in Corpus Linguistics.* Berlin: Mouton de Gruyter. 105-122.

Leed, Richard and Alexander Nakhimovsky. 1980. *Advanced Russian*. Columbus, Ohio: Slavica.

Lehečková, Helena. 2003. Recovery from aphasia after polytrauma in a Czech child: What is lost and what is left. *Brain and Language*. 87: 163-164.

Lehečková, Helena. 2002. Broca's aphasia in Czech: A linguistic analysis. *Brain and Language*. 83 14-17.

Lehečková, Helena. 2001. Manifestation of aphasic symptoms in Czech. *Journal of Neurolinguistics*. 14: 179-208.

Lifschitz, Elena D. and Sandra F. Rosengrant. 1996. *Focus on Russian*. New York: John Wiley and Sons.

Lindner, Susan. 1981. *A Lexico-Semantic Analysis of Verb-Particle Constructions with* Up *and* Out. Ph.D. dissertation. University of California - San Diego.

Lozinska, Monika A. 1992. *A comparison of a contextual usage of selected secondary and simplex imperfectives in Russian: An experimental study.* Doctoral dissertation. University of Alberta.

Ludwig, Jonathan Zachary. 1995. *Multiply-Prefixed Verbs in Russian, Polish, and Ukrainian*. Doctoral dissertation. Indiana University.

Manning, C. D. and H. Schütze. 2000. *Foundations of Statistical Natural Language Processing.* Cambridge, Massachusetts: MIT Press.

Maslov, Ju. S. 1959. Glagol'nyj vid v sovremmenom bolgarskom literaturnom jazyke. In: S. B. Bernštejn (ed.). *Voprosy grammatiki bolgarskogo literurnogo jazyka*. Moscow: Izdatel'stvo Akademii nauk SSSR. 157-378.

McDonald, Scott. 1997. Exploring the Validity of Corpus-derived Measures of Semantic Similarity. Paper presented at the 9th Annual CCS/HCRC.

McEnery, Tony and Andrew Wilson. 2001. *Corpus Linguistics: An Introduction.* Edinburgh University Press.

Mehlig, Hans Robert. 2004. Glagol'nyj vid i vtoričnaja gomogenizacija oboznačaemoj situacii posredstvom kvantifikacii: K upotrebleniju delimitativnogo sposoba dejstvija v russkom jazyke. In: V. Lehmann (ed.). *Slavjanskij vid i leksografija*. *Slavolinguistica 4*. München: Sagner.

Mehlig, Hans Robert. 1996. Some analogies between the morphology of nouns and the morphology of aspect in Russian. *Folia Linguistica*. 30: 87-109.

Młynarczyk, Anna. 2004. *Aspectual Pairing in Polish*. Doctoral dissertation. Utrecht: Landelijke Onderzoekschool Taalwetenschap.

Moravcsik, Edith A. and Jessica Wirth. 1986. Markedness – an overview. In: F.R. Eckman, E.A. Moravcsik, and J.R. Wirth (eds.). *Markedness*. New York: Plenum Press. 1-11.

Němec, Igor. 1954. O slovanské přepodně *po-* slovesné. *Slavia.* 23: 1-22.

Nesset, Tore. Forthcoming. Is the choice of prefix arbitrary? Aspectual prefixation and Russian verbs of perception. *Slavic and East European Journal*. 52 pp.

Nesset, Tore. 2009. Metonymy of Aspect/Aspects of Metonymy. *Scando-Slavica.* 55: 65-77.

Nesset, Tore. 2007. The Path to Neutralization: Image Schemas and Prefixed Motion Verbs. *Poljarnyj Vestnik.* 10: 61-71.

Ožegov, S. I. and N. Ju. Švedova. 2003. *Tolkovyj slovar' russkogo jazyka*. 4-e izdanie. Moskva: Rossijskaja akademija nauk. Institut russkogo jazyka im. V. V. Vinogradova.

Paprotté, Wolf and René Dirven (eds.). 1985. *The Ubiquity of metaphor: Metaphor in language and thought*. Amsterdam/Philadelphia: John Benjamins.

Peirsman, Yves and Dirk Geeraerts. 2006. Metonomy as a prototypical category. *Cognitive Linguistics*. 17(3): 269-316.

Petruxina, E. V. 2000. *Aspektual'nye kategorii glagola v russkom jazyke: v sopostovlenii s češskim, slovackim, pol'skim i bolgarskim jazykami*. Moscow: Izdatel'stvo Moskovskogo Universiteta.

Pinker, S. and R. Jackendoff. 2005. The faculty of language: What's special about it? *Cognition.* 95: 201-236.

Pinker, S. 1994. *The language instinct: How the mind creates language*. New York: W. Morrow.

Pulkina, I. and E. Zakhava-Nekrasova. 1974. *Russian: A Practical Grammar with Exercises*. trans. V. Korotky, ed. by R. Dixon and G. Konorova. 2nd ed. Moscow: Russkij Jazyk.

Pulvermuller, F. 2003. *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. UK, Cambridge: Cambridge University Press.

Raukko, Jarno. 1999. An "intersubjective" method for cognitive semantic research on polysemy: the case of *get*. In: Hiraga, Masako K., Chris Sinha, and Sherman Wilcox (eds.). *Cultural, psychological and typological issues in cognitive linguistics. Selected papers of the bi-annual ICLA meeting in Albuquerque, July 1995.* Amsterdam/Philadelphia: John Benjamins. 87-105.

Rosch, Eleanor. 1978. Principles of categorization. In: E. Rosch and B. B. Lloyd (eds.). *Cognition and Categorization.* 27-48.

Rosch, Eleanor. 1973a. Natural Categories. *Cognitive Psychology*. 4: 328-350.

Rosch, Eleanor. 1973b. On the internal structure of perceptual and semantic categories. In: T. E. Moore (ed.). *Cognitive development and the acquisition of language*. New York: Academic Press. 111-144.

Rosch, Eleanor and Carolyn B. Mervis. 1975. Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology.* 7: 573-605.

Rudzka-Ostyn, Brygida. 1983a. Cognitive Grammar and the Structure of Dutch UIT and Polish WY. Unpublished manuscript. University of Belgium.

Rudzka-Ostyn, Brygida. 1983b. Metaphoric Processes in Word Fromation: The Case of Prefixed Verbs. Unpublished manuscript. University of Belgium.

Russell, Pamela. 1985. Aspectual Properties of the Russian Verbal Prefix *na-*. In: Michael S. Flier and Alan Timberlake (eds.). *The Scope of Slavic Aspect*. Columbus, Ohio: Slavica. 59-75.

Russkaja Grammatika. 1982. Moscow: Izdatel'stvo Nauka.

Sandra, Dominiek and Sally Rice. 1995. Network analyses of prepositional meaning: Mirroring whose mind – the linguist's or the language user's? *Cognitive Linguistics.* 6(1): 89-130.

Sasse, Hans-Jürgen. 2002. Recent activity in the theory of aspect: Accomplishments, achievements, or just non-progressive state? *Linguistic Typology.* 6: 199-271.

Šaxmatov, A. A. 1941. *Sintaksis russkogo jazyka*. Leningrad.

Schmid, Hans-Jörg. Forthcoming. Does frequency in text instantiate entrenchment in the cognitive system? 32 pp.

Schmid, Hans-Jörg. 2000. *English Abstract Nouns as Conceptual Shells.* (= *Topics in English Linguistics 34*.) Berlin: Mouton de Gruyter.

Shull, Sarah Ellen. 2003. *The experience of space: The privileged role of spatial prefixation in Czech and Russian.* Slavistische Beiträge. München: Sagner.

Sigalov, Pavel S. 1975. Istorija russkix ograničitel'nyx glagolov. *Trudy po russkoj i slavjanskoj filologii: Serija lingvističeskaja*. 24: 141-181.

Sinha, Chris and Tania Kuteva. 1995. Distributed Spatial Semantics. *Nordic Journal of Linguistics.* 18: 167-199.

Slovar' sovremennogo russkogo literaturnogo jazyka. 1950–65. Moscow: Izdatel'stvo Akademii nauk SSSR.

Stefanowitsch, Anatol and Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics.* 8(2): 209-243.

Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, Julie C. Sedivy. 1995. Integration of Visual and Linguistic Information in Spoken Language Comprehension. *Science*, New Series. 268(5217): 1632-1634.

Tatevosov, Sergej. 2002. The parameter of actionality. *Linguistic Typology.* 6: 317-401.

Taylor, John R. 1995. *Linguistic Categorization: Prototypes in Linguistic Theory*. New York: Oxford University Press.

Tchizmarova, Ivelina K. 2006. A Cognitive Linguistic Analysis of the Bulgarian Verbal Prefix *Pre-* 'across, through, over'. *Glossos* 7. http://www.seelrc.org/glosos/

Tchizmarova, Ivelina. 2005. *Verbal Prefixes in Bulgarian and Their Correspondences in American English: A Cognitive Linguistic Analysis.* Doctoral dissertation. Ball State University.

Timberlake, Alan. 1984. The Temporal Schemata of Russian Predicates. In: Michael S. Flier and Richard D. Brecht (eds.). *Issues in Russian Morphosyntax*. Columbus, Ohio: UCLA Slavic Studies, Vol. 10. 35-57.

Tixonov, A. N. 1998. *Russkij glagol: Problemy teorii i leksikografirovanija.* Moscow: Akademia.

Tixonov, A. N. 1962. K voprosu o čistovidovyx pristavkax v russkom jazyke. *Trudy Samarkandskogo gosudarstvennogo universiteta im. A. Navoi*. Novaja serija, No. 118: *Issledovanija po russkomu jazyku*. Samarkand. 31-57.

Tixonov, A. N. 1961. Grammatikalizacija glagol'nyx pristavok v russkom jazyke. *Materialy tret'ej objedinjonnoj naučnoj konferencii učonyx goroda Samarkanda. Serija gumanitarnyx i estestvennyx nauk*. Samarkand. 173-176.

Tixonov, A. N. 1958. Zametki o vtoričnoj imperfektivacii glagolov s čistovidovymi pristavkami. *Trudy Uzbekskogo gosudarstvennogo universiteta im. A. Navoi.* Novaja serija, No. 92. Samarkand. 59-81.

Townsend, Charles E. 1975. *Russian Word-Formation*. Columbus, Ohio: Slavica.

Tyler, Andrea and Vivian Evans. 2003. *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning and Cognition.* Cambridge/New York: Cambridge University Press.

Unbegaun, B. O. 1967. *Russian Grammar*. Oxford: Clarendon.

Van Schooneveld, C. H. 1978. *Semantic Transmutations: Prolegomena to a Calculus of Meaning, Vol. 1: The Cardinal Semantic Structure of Prepositions, Cases, and Paratactic Conjunctions in Contemporary Standard Russian*. Bloomington, Indiana: Physsardt.

Van Schooneveld, C. H. 1958. The So-called 'préverbes vides' and Neutralization. *Dutch Contributions to the Fourth International Congress of Slavists*. The Hague. 159-161.

Vendler, Zeno. 1957. Verbs and times. *The philosophical review.* 66: 143-160.

Vey, M. 1952. Les préverbes 'vides' en tchéque moderne. *Revue des études slaves*. 29: 82-107.

Veyrenc, Jacques. 1980. Pour une Grammaire Syntagmatique des Préverbes. *Etudes sur le Verbe Russe*. Paris: Institut pour d'Etudes Slaves.

Vinogradov, V. V. 1938 (reprinted 2001). *Russkij jazyk: grammatičeskoe učenie o slove.* Moscow.

Winters, M. 1990. Toward a theory of syntactic prototypes. In: S. L. Tsohatzidis (ed.). *Meanings and Prototypes: Studies in Linguistic Categorization.* London: Routledge. 285-306.

Wittgenstein, Ludwig. 1973. *Philosophical Investigations*. trans. G. E. M. Anscombe. New York: Macmillan.

Zaliznjak, Anna A. and Aleksej D. Šmelev. 2000. *Vvedenie v russkuju aspektologiju*. Moscow: Jazyki russkoj kul'tury.

Zemskaja, E. A. 1955. Tipy odnovidovyx pristavočnyx glagolov v sovremennom russkom jazyke. *Issledovanija po grammatike russkogo literaturnogo jazyka*. Moskva: Izdatel'stvo Akademii Nauk SSSR.