

# UC Berkeley

## UC Berkeley Electronic Theses and Dissertations

### Title

The Representation of Morphemes in the Russian Lexicon

### Permalink

<https://escholarship.org/uc/item/8x62v9dq>

### Author

Antic, Eugenia

### Publication Date

2010

Peer reviewed|Thesis/dissertation

**The Representation of Morphemes in the Russian Lexicon**

by

Eugenia Antić

A dissertation submitted in partial satisfaction of the  
requirements for the degree of  
Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eve Sweetser, Co-Chair  
Professor Sharon Inkelas, Co-Chair  
Professor Susanne Gahl  
Professor Johanna Nichols

Fall 2010

**The Representation of Morphemes in the Russian Lexicon**

Copyright 2010  
by  
Eugenia Antić

## Abstract

The Representation of Morphemes in the Russian Lexicon

by

Eugenia Antić

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Eve Sweetser, Co-Chair

Professor Sharon Inkelas, Co-Chair

Different morphological theories assign different status to parts of words, roots and affixes. Models range from accepting both bound roots and affixes to only assigning unit status to standalone words. Some questions that interest researchers are 1) What are the smallest morphological units, words or word parts? 2) How does frequency affect morphological processing? 3) How do experimental results affect existing morphological theories? In this dissertation I attempt to investigate these questions in more detail using results from psycholinguistic experiments and integrating them into the Network theory of morphology.

I first consider what kinds of information a morphological model should take into account. It should reflect the intuition of speakers that words consist of different parts, while avoiding the difficulties of theories that propose that roots and affixes are separate units. In addition, it should incorporate the notion of paradigms and frequency information. I look at how various morphological theories account for Russian data, in particular verbal prefixes and words where they occur.

The status of roots is intensely debated in morphology, and views of the nature of roots range from being complete separate units on their own, to being considered epiphenomena over words, which are the only existing units. Additionally, it is frequently debated whether free and bound roots have the same status, or are different. In the dissertation I describe the results of a prefix separation experiment in Russian, where more than two types of roots can be identified. The results demonstrate that a productive prefix is separated with greatest ease from words with free roots and with greatest difficulty from words with completely bound roots. The remaining two root types are in between those extremes, according to the characteristics of the root types. I interpret these results as meaning that roots form a continuum based on the strength of connections between words. The stronger the form and meaning connections, the more easily identifiable a root. This interpretation fits well within the Network theory of morphology, where only words are units and there are connections of differing strength between them.

Continuing with the investigation of different units of morphological processing, I report the results of a complexity rating experiment in Russian where words with different prefixes were used. The results demonstrate that ratings of word complexity show a direct correlation with the level of transparency of prefixes in those words. Results of both of these experiments fit into the Network theory of morphology; however, two interpretations are possible with respect to prefixes. While I suggest that roots are not units, it is possible that prefixes, and affixes in general, can be represented as generalizations over many words. It is also possible that they are not units, and that strong connections are formed between words containing those prefixes and new words naturally fall into the pattern to form new words containing the prefixes.

In addition to investigating the unit or non-unit status of roots and prefixes, I study the influence of relative frequency in Russian morphological processing. Studying the data from the prefix separation experiment only with free roots, I compare the processing of words that are more frequent than their bases to processing of words that are less frequent than their bases. The difference between processing of those two sets of words demonstrates that relative frequency is an important factor in Russian morphological processing. These results are in accordance with previous studies in English, Italian and Tagalog, and are thus suggestive of a universal principle of morphological organization.

Finally, I incorporate these results into the Network theory of morphology, where units of morphological processing are words and connections between words. The stronger the connections between words, the more apparent are the different constituent parts. However, the roots and affixes within affixed words are not units of processing, but are epiphenomena over the words and connections. The status of prefixes is not completely clear: clearly, generalizations are formed over prefixes in prefixed words, however, it is not certain whether those generalizations are temporary or permanent. Overall, the Network theory of morphology is revised and updated according to the reported experimental results.

*To my family.*

# Contents

<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Overview of the study . . . . .	1
1.2 Brief background on Russian verbal morphology . . . . .	2
1.2.1 Data range . . . . .	3
1.3 Analyses of Russian data in different theories of morphology . . . . .	4
1.3.1 Item-and-Arrangement . . . . .	4
1.3.2 Item-and-Process . . . . .	7
1.3.3 Syntactic theories . . . . .	10
1.3.4 WP and Network models . . . . .	13
1.4 Summary and theoretical considerations . . . . .	21
1.4.1 Evaluation metrics . . . . .	21
1.4.2 Semantics and semantic transparency . . . . .	22
1.4.3 Frequency in morphology . . . . .	24
1.5 Conclusion . . . . .	25
<b>2 An investigation of relative frequency effects and the Network theory of morphology</b>	<b>26</b>
2.1 Theoretical preliminaries . . . . .	27
2.2 Relative frequency effects . . . . .	29
2.3 Status of roots . . . . .	30
2.4 Experimental predictions . . . . .	33
2.4.1 Root status . . . . .	33
2.4.2 Prefix status . . . . .	36
2.5 Summary . . . . .	37
<b>3 Root status and cumulative root frequency effects</b>	<b>38</b>
3.1 Introduction . . . . .	38

3.2	Root status experiment . . . . .	38
3.2.1	Subjects . . . . .	39
3.2.2	Stimuli . . . . .	39
3.2.3	Procedure . . . . .	40
3.2.4	Results . . . . .	40
3.2.5	Discussion . . . . .	43
3.3	Are roots units of representation in the mental lexicon? . . . . .	49
3.4	Conclusions . . . . .	50
<b>4</b>	<b>Relative frequency effects</b>	<b>52</b>
4.1	Introduction . . . . .	52
4.2	Po- and voz- productivity analysis . . . . .	54
4.2.1	Prefix descriptions . . . . .	54
4.2.2	Productivity analysis . . . . .	56
4.3	Po- experiment data analysis . . . . .	59
4.3.1	Analysis . . . . .	59
4.3.2	Discussion . . . . .	61
4.4	Conclusion . . . . .	63
<b>5</b>	<b>Differences between affixes</b>	<b>64</b>
5.1	Introduction . . . . .	64
5.2	Previous work . . . . .	64
5.3	Russian prefix characteristics . . . . .	65
5.4	Complexity ratings experiment . . . . .	69
5.4.1	Subjects . . . . .	69
5.4.2	Stimuli . . . . .	69
5.4.3	Procedure . . . . .	70
5.4.4	Results . . . . .	70
5.4.5	Discussion . . . . .	74
5.4.6	Methodology differences . . . . .	76
5.5	Conclusion . . . . .	76
<b>6</b>	<b>The Network model of morphology and its implications</b>	<b>78</b>
6.1	Hypothesis evaluation . . . . .	78
6.2	Morphological model restrictions . . . . .	79
6.3	Revised network model of morphology . . . . .	80
6.3.1	Relative frequency effects . . . . .	83
6.3.2	Root hierarchy . . . . .	84
6.3.3	Semantic transparency in the network model . . . . .	84
6.3.4	Differences between prefixes . . . . .	85
6.4	Status of roots and affixes . . . . .	85



6.5 Conclusion . . . . .	89
<b>A List of neologisms used for the productivity test</b>	<b>97</b>

## List of Figures

1.1	Representation of <i>bežat'</i> and <i>pobežat'</i> . . . . .	17
1.2	Representation of <i>polit'</i> . . . . .	19
1.3	Representation of <i>prokurit'</i> and a general schema . . . . .	20
3.1	Experiment 1 results . . . . .	41
3.2	Form and semantic connections between words. . . . .	48
4.1	Plot of base versus derived frequency for <i>po-</i> . . . . .	58
4.2	Plot of base versus derived frequency for <i>voz-</i> . . . . .	58
4.3	Average RT for words that are more frequent than their bases (word) and less frequent than their bases (base). . . . .	60
5.1	Histogram of rating untransformed . . . . .	72
5.2	Histogram of rating transformed . . . . .	72
5.3	Model residuals . . . . .	73
6.1	Form and semantic connections between words. . . . .	80
6.2	Paradigm connections. Connections exist between all members of the paradigm. Not all of them are shown for clarity. <i>Stol</i> 'table' nom. sg., <i>stola</i> gen. sg., <i>stole</i> loc. sg., <i>stolom</i> inst. sg., <i>stolu</i> dat. sg. . . . .	82
6.3	Paradigm organized in a satellite-entry fashion. The nominative singular form is central, all others are connected to it. Forms illustrated same as above. . .	82
6.4	Prefix <i>pro-</i> . . . . .	87
6.5	Root <i>-ryb-</i> . . . . .	88
6.6	Representation of <i>prokurit'</i> and a general schema . . . . .	89
6.7	Representation of a pattern match . . . . .	90

# List of Tables

1.1	Sample Russian Verbs . . . . .	3
1.2	Paradigm of <i>stol</i> 'table' . . . . .	14
2.1	Masked priming results from (Pastizzo and Feldman, 2004). . . . .	31
3.1	By-subject RT averages for different root types. p-values are calculated for the comparison of the root type in that row with the root type in the following row. . . . .	40
3.2	By-item RT averages for different root types. p-values are calculated for the comparison of the root type in that row with the root type in the following row. . . . .	41
3.3	Fixed effects of the mixed effects regression model. . . . .	42
3.4	Random effects of the mixed effects regression model. . . . .	43
3.5	The links parameter for the four types of words. . . . .	45
3.6	Number of verb with a particular links parameter, links parameter sorted from 0 to 1. . . . .	46
3.7	Fixed effects of the mixed effects regression model. . . . .	47
3.8	Random effects of the mixed effects regression model. . . . .	47
4.1	Fixed effects of the mixed effects regression model. . . . .	61
4.2	Random effects of the mixed effects regression model. . . . .	61
5.1	Numeric characteristics of 5 Russian prefixes. . . . .	66
5.2	Semantic characteristics of 5 Russian prefixes. . . . .	68
5.3	Transparency scores of five Russian prefixes. . . . .	69
5.4	Number of different root types used in the experiment. . . . .	70
5.5	Fixed effects of the mixed effects regression model. . . . .	73
5.6	Random effects of the mixed effects regression model. . . . .	73
A.1	Words used for the productivity test. . . . .	98

## Acknowledgments

It is a pleasure to thank those who made this thesis possible. I would like to thank my advisors, Eve Sweetser and Sharon Inkelas for their continued guidance, support, and providing a fresh look at my ideas. I am grateful to Eve for introducing me to cognitive linguistics and encouraging my interest in it. I am indebted to Sharon for introducing me to the idea of relative frequency that eventually led to this dissertation. Many thanks to the other members of my committee: to Susanne Gahl for invaluable help with statistics and multiple regression, as well as making me think carefully about my arguments, and to Johanna Nichols for help with Russian linguistic descriptions and reasoning. I would like to thank Keith Johnson for his help with my qualifying paper and with statistical analyses.

Other members of the Berkeley faculty have helped immensely during my years at the Berkeley linguistics department. I would like to thank all professors whose classes made for an interesting and engaging graduate career: Line Mikkelsen, Gary Holland, George Lakoff, Leanne Hinton, Ian Maddieson, Lynn Nichols, as well as the rest of the Linguistics faculty for making the graduate program interesting and stimulating. I would also like to thank members of the Slavic department who made my stay at Berkeley all the more enjoyable and intellectually stimulating: Lisa Little, Ronelle Alexander, David Frick. Thank you to my fellow graduate students who created a very stimulating environment for learning. I would like to thank the Berkeley staff, Belén Flores, Paula Floro and Natalie Babler, who have provided invaluable logistical support and advice. I am grateful to all the participants who have taken part in my experiments, as without their involvement this dissertation would not be possible.

Finally, I would like to thank my family for their continuing support. My parents, Alexander Trusov and Irina Vasilyeva for believing in me and helping me with running my experiments. My mother-in-law, Danica Antić, for occasional help and support at home. My sons, Nikola and Ivan, for their inspiration and the opportunity to watch their everyday linguistics experiments. My husband, Filip Antić, for his unwavering support and humor, as well as technical help with statistical data.

# Chapter 1

## Introduction

### 1.1 Overview of the study

Various morphological theories assign different status to the traditionally described subparts of words, i.e., roots and affixes. Models range from assigning unit status to all morphemes, including bound roots and affixes, to only assigning unit status to standalone words. Both approaches can present problems. If all morphemes are standalone units, and meaning is assigned to them, then cases with one-to-many or many-to-one meaning to morpheme relationships do not fit into the theory. If no morphemes are units, it is unclear where the intuition of words consisting of subparts comes from. A combination of Word-and-Paradigm, Network and Construction/Cognitive Grammar can account for these problems without difficulties. While only words are standalone units with meaning associated with them, there are connections between words that make the word parts apparent.

In this chapter, I consider examples of Russian word structure and the different theoretical treatments, including Item-and-Arrangement (IA), Item-and-Process (IP), Syntactic (S), Word-and-Paradigm (WP) and Network models of morphology. I look at how these theories can account for the structure of several Russian verbs that include productive Russian verbal prefixes, and how these prefixes combine with other word parts. The Network theory accounts for all of the data, while each of the other theories presents at least one major problem in its treatment.

The Network theory of morphology has two types of units: words and connections between words. These connections are stronger or weaker depending on several factors, including word frequency. Word frequency has been found to be important in morphological processing ((Bybee, 2007, pp.5-22)), and thus should be included as a factor in a morphological model. Some unresolved questions in this theory include whether all roots and affixes have separate entries in the lexicon and whether effects of relative frequency of the word and its base have an effect on the structure of the lexicon, and I investigate them in this dissertation.

In her description of Network theory, Bybee (1985) leaves the question of roots unresolved, stating that her definition of ‘word’ is intentionally left vague and may also be language-specific. Other theorists debate this question. Some (Bloomfield, 1933; Lieber, 1980; Kiparsky, 1982; Stump, 2001) assign roots the status of standalone units, while others (Robins, 1959; Bochner, 1993; Blevins, 2006) only recognize words as possible building blocks of morphology. In addition to the question of whether roots are units, which *types* of roots are units is also a debated question. Forster and Azuma (2000); Pastizzo and Feldman (2004) find no differences in processing free and bound roots, while Marslen-Wilson et al. (1994) provide evidence for differences between them. Since Russian has more types of roots than just free and bound, it provides suitable material for investigating this question. I find that the roots form a hierarchy from completely bound to free, and these results can fit well into the Network theory, where the stronger the connections between the words, the easier it is to identify the word parts.

Another question that arises is the status of affixes in the Network theory. Langacker (2002) states that affixes, like the English plural suffix, form generalizations over all the words where they are used. This is somewhat contradictory to the principle of only words being units (form-meaning pairings) in the Network theory, although intuitive. Russian is again very appropriate in investigating this question, since it has about 20 verbal prefixes of different productivity. Although the productivity varies, the overall process of forming differently prefixed words, especially verbs, is very productive. Thus, borrowings freely combine with the prefixes (and combine with some better than others). I use 5 prefixes in a complexity rating experiment to investigate differences in their processing. I find that the more semantically transparent a prefix is, the more complex the word is rated. Although this provides some evidence in favor of unit status of prefixes, it does not resolve the question. Prefixes could form generalizations that are immediately discarded, or stored for further use.

Finally, I use Russian, a morphologically complex language, for testing whether relative frequency effects are relevant for morphological processing. Hay (2002) suggests that relative, not absolute frequency, is the parameter that influences morphological decomposition. Relative frequency is the difference between the frequency of the word itself and that of its unaffixed base. Hay (2002) found these effects in English, Burani and Thornton (2003) in Italian and Zuraw (2009) in Tagalog, and I show that such effects also exist in Russian. This points to a universal principle of morphological organization, and I consider how relative frequency effects can fit into the Network theory of morphology.

In the final chapter of the dissertation, I integrate the reported experimental results into the existing descriptions of the Network theory of morphology. I consider how the relative frequency effects could fit into it, and explore two options for representing affixes.

## 1.2 Brief background on Russian verbal morphology

Russian is a language rich in morphology, both inflectional and derivational. A large part of derivational morphology are Russian verbal prefixes. They attach to a variety of verbs and have a range of functions:

1. Change of meaning. For example, *žit'* 'to live, imp.' versus *vyžit'* 'to survive, pf.'.
2. Changing an imperfective verb to a perfective one. For example, *pisat'* 'to write, imp.' versus *napisat'* 'to write, pf'.

These verbal prefixes are in many cases homophonous with spatial prepositions and have spatial meanings with many verbs. For example, *prygnut'* 'to jump, pf.' versus *vyprygnut'* 'to jump out'.

In addition to prefixes, verbs are formed using imperfectivizing suffixes, usually from perfective prefixed verbs. For example, *čitat'* 'to read, imp.', *pročitat'* 'to read (completely), pf.', *pročityvat'* 'to (iteratively) read (completely), imp.'.

Thus, the structure of a Russian infinitive verb is as follows: optional prefix, root, optional thematic vowel, optional suffix, infinitive suffix. Examples of verbs are given in Table 1.1.

Verb	Gloss	Prefix	Root	Thematic vowel	Suffix	Infinitive suffix
čitat'	to read, imp.	-	čit	a	-	t'
pročitat'	to read through, pf.	pro	čit	a	-	t'
pročityvat'	to iteratively read through, imp.	pro	čit	-	yva	t'

Table 1.1: Sample Russian Verbs

### 1.2.1 Data range

The Russian verbs briefly described above are interesting in that they exhibit all the characteristics that are of interest:

- There are a variety of productive and not very productive affixes that are used with Russian verbs.

- Many Russian verbs also have bound roots in them.
- A given Russian root will participate in various inflectional and derivational paradigms, forming relationships with other words.

I will consider the prefix *po-* and verbs used with that prefix, including verbs with different types of roots and meanings:

- *bežat'* 'to run' impf., *pobežat'* 'to start running' pf. This pair is interesting in that the prefix *po-* in this case exhibits the meaning 'to start doing something'.
- *lit'* 'to pour' impf., *polit'* 'to start pouring' or 'to water (flowers, plants, etc.)' pf. In this case, the prefix *po-* has two meanings, depending on context.
- *polivat'* 'to water (flowers, plants, etc.)' impf. This verb is related to the two above; however, it does not appear without the prefix, i.e. there is no *\*livat'*.
- *pojasnjat'* 'to clarify' impf., *pojasnit'* 'to clarify' pf., *projasnjat'* 'to clarify' impf., *projasnit'* 'to clarify' pf., *vyjasnjat'* 'to find out' impf., *vyjasnit'* 'to find out' pf. This triplet is interesting in that the verbs are relatively frequent and the root appears with three different prefixes. However, there is no unprefixated verb *\*jasnjat'*. This case is different from the one above, because there are no related verbs that appear unprefixated, although there are other unprefixated words that contain the same root, such as *jasnost'* 'clarity'.
- *kurit'* 'to smoke' impf., *pokurit'* 'to smoke for a while' pf., *prokurit'* 'to smoke for a while' pf. This set demonstrates the delimitative meaning of the prefixes *po-* and *pro-*, which I explore in more detail in (Antić, 2006).

## 1.3 Analyses of Russian data in different theories of morphology

### 1.3.1 Item-and-Arrangement

Item-and-Arrangement is perhaps the most well-known type of morphological theory. Of the different IA theories, three are surveyed here: Bloomfield's (1933) description of a traditional IA model, Kiparsky's Lexical Morphology (Kiparsky, 1982) and a model by Lieber (1980).

As Bloomfield describes, in a traditional IA model, words consist of morphemes: phonological form paired with a meaning. Words are assumed to exhaustively consist of morphemes, where morphemes are pairings of form and meaning, and the meaning of combining morphemes is generally assumed to be compositional, a combination of the meanings of the



constituent morphemes. For example, a verb like *baked* is assumed to consist of two parts: the root *bake*, whose meaning is the action of baking (however it is represented) and the affix *-ed*, whose meaning is ‘past tense’. Thus, the verb *baked* combines both the form and the meaning of its constituent morphemes to mean ‘bake at a certain time in the past’. Thus, in a traditional IA model both bound roots and affixes are given the status of morphological units. In fact, bound morphemes are numerous in such an approach, since they include such morphemes as the infamous *cran* of *cranberry*.

The problems with such an approach are numerous and well-known (see Anderson, 1992, pp. 51-56). Most of the problems arise from the restriction in classical IA theory that the semantics-form relationship be one-to-one. Empty, cumulative and portmanteau morphs are three classes of phenomena that violate this restriction. Other problems that are associated with semantics-form mismatches are ‘bracketing paradoxes’. These arise when Level Ordering is assumed, as does Kiparsky (1982), where some affixes are ‘Level 1’ and others are ‘Level 2’. For example, *un-* is assumed to be Level 2, while *-th* is assumed to be Level 1. Thus, following this convention, *untruth* should be bracketed *un[[true]th]*, while its semantics suggests *[un[true]]th*. Such bracketing paradoxes, mismatches in compound structure (*blackboard* is not necessarily black) and other semantic mismatch phenomena force Lieber (1980) to conclude that the final meaning of a word is not necessarily put together in a way isomorphic with the morphological structure, although she does assume that morphemes are paired with a semantic representation.

Thus, Lieber’s theory, although similar to Bloomfield’s original view, differs significantly in that it separates semantics from form in putting together semantic representations. Kiparsky offers different workarounds to solve the bracketing paradoxes (Kiparsky, 1982, pp. 121-124). The main characteristics of his theory, on the other hand, is the separation of morphology into three levels and the assignment of a level to each affix.

The general characteristics of IA models are thus as follows:

- Words consist of morphemes.
- Each morpheme either is associated with meaning. Meaning is compositional and is isomorphic with morphological structure (Bloomfield, 1933; Kiparsky, 1982) or is not (Lieber, 1980).
- Morphemes are put together into words via rules.
- In some theories, morphemes are assumed to belong to different levels and have fundamentally different properties (Kiparsky, 1982; Lieber, 1980).
- Both bound roots and affixes are morphological units in IA theories.

Thus, in IA models both affixes and bound roots have the status of separate units. In some theories, they are paired with meaning, in some they are not. Here I assume the apparently more standard view that morphemes are form-meaning pairings. Looking at the words in Section 1.2.1, the following analyses emerge:

- *bežat'* 'to run' impf., *pobežat'* 'to start running' pf. These words are related to the noun *beg* 'running' and thus in IA would have this root [beg, 'running'] (where instead of the English gloss there would be some semantic representation) as a unit. The *-t'* suffix appears on the majority of Russian infinitives and thus a unit [-t', infinitive] is present in both of these words. A problem arises when we encounter the theme vowel *-a-*, since no meaning can be assigned to it. The verb *bežat'* is imperfective, and in some verbs the theme vowel *-a-* is a reflex of the imperfective suffix *-yva*, so we could possibly assign the meaning 'imperfective' to *-a-*. However, the perfective verb *pobežat'* also contains this theme vowel. Thus, there are two alternatives: either state that *-a-* is an empty morph, or constructing the unit [-a-, impf.], which will lead to problems when analyzing the verb *pobežat'*. If *-a-* is taken to be an empty morph, then it is not clear where the imperfective meaning comes from.

The verb *pobežat'* then would be analyzed as follows: [beg, 'running'], [-t', inf.], [-a-, ??]. The last unit is the prefix *po-*, which is also somewhat problematic. As we can see, the difference in meaning between *bežat'* and *pobežat'* is that the latter is perfective and inchoative. Thus, the following would be the structure assigned to *po-*: [po, perf. & inch.]. However, if *-a-* was previously assigned the meaning 'imperfective', it is not clear how *po-* would overwrite that. If *-a-* was not assigned any meaning, the problem of an empty morph persists.

Another problem with assigning perfective and inchoative meaning to *po-* is that in many cases verbs with *po-* are not perfective, and it has many other meanings aside from inchoative. Thus, we would need to assign diacritics to different meanings of *po-*. As we will see below, some verbs allow more than one reading of *po-*, while *pobežat'* allows only one. The rules that combine the morphemes would then need to list, for each meaning of *po-*, which verb roots allow it, and which do not. Such a listing will probably result in the same amount of information as listing the verbs in a network-style model. As already mentioned, a potential problem might arise in the rules when the meaning 'imperfective' is assigned to *-a-*.

The rule would have put the following together: [po-, perf. & inch.] + [beg, 'running'] + [-a-, impf.] + [-t', inf.]. We will see below that there are instances where almost the same combination of morphemes produces different results; in one case, perfective, in the other, imperfective.

Additionally, affixes have to have a specification of what items they attach to, since they are dependent morphemes. Thus, for example, *po-* would have a specification [\_V], as attaching to verbs. There again would have to be several items for *po-*, since it attaches not only to verbs, but also nouns, adjectives and adverbs.

- *lit'* 'to pour' impf., *polit'* 'to start pouring' or 'to water (flowers, plants, etc.)' pf. In this case, we again have the infinitive morpheme [-t', inf.]. With the first verb, a problem arises as to where the imperfective meaning should be assigned. Apart from

the *-t'* morpheme, both *l* and *i* are segments of the root (cf. the related word *liven'* 'rain shower'), and we must have a morpheme [li, 'pour']. Thus there is no morpheme where the meaning 'imperfective' could be assigned. Even if we suggest that a verb that has no aspect specification receives the meaning 'imperfective', it still has to be assigned to the whole word, not a particular morpheme.

Additionally, as we see, the corresponding perfective verb *polit'* has two meanings: one inchoative and one not. In an IA theory that would mean that *po-* would have two (and more) variants with diacritics. Problems with such an approach are described above.

- *polivat'* 'to water (flowers, plants, etc.)' impf. The morpheme segmentation in this word is as follows: [po-, ??] + [li, pour] + [-iva-, impf.] + [-t', inf.]. It is not clear what meaning would be assigned to *po-* in this case. This word, *polivat'* is related to the word *polit'* 'to water flowers, plants' pf. above. In *polit'*, *po-* seems to only have perfectivizing meaning, while it cannot have that meaning in *polivat'*, since the verb is imperfective. On the other hand, *polivat'* is not synonymous with *lit'*, the imperfective verb that *polit'* is derived from. It seems that these three verbs form a triplet primary imperfective - perfective - secondary imperfective. However, such a network is not easily modeled in IA theories.
- *pojasnjat'* 'to clarify' impf., *pojasnit'* 'to clarify' pf., *projasnjat'* 'to clarify' impf., *projasnit'* 'to clarify' pf., *vyjasnjat'* 'to find out' impf., *vyjasnit'* 'to find out' pf. Not much is to be said about these verbs, apart from the fact that the prefixes in these cannot be associated with aspectual meaning, since both perfective and imperfective verbs in this case bear them. It is possible to assign the meaning 'perfective' to *-i-* and the meaning 'imperfective' to *-a-*, however, the same problems remain: these suffixes do not always appear in perfective and imperfective verbs, respectively. Additionally, it seems that the difference in meaning arises in some sense from the opposition of the two words of a pair.
- *kurit'* 'to smoke' impf., *pokurit'* 'to smoke for a while' pf., *prokurit'* 'to smoke for a while' pf. One peculiarity of these verbs that is different from already discussed above is the difference in meaning between the two perfective verbs, which is discussed in more detail in (Antić, 2006). The verb *prokurit'*, when used in the durative meaning specified, requires an argument (time period), while *pokurit'* does not. While it is not problematic for IA theories per se, it requires modification of the theory to allow for such morphemes to have predicate-level specifications.

### 1.3.2 Item-and-Process

A major characteristic of IP models is that affixes are not 'items'; they are merely markers of certain processes that happen to roots, stems or words. I will discuss three IP models,

(Aronoff, 1976), (Anderson, 1992) and (Stump, 2001)<sup>1</sup>. Aronoff's (1976) and Anderson's (1992) theories are very similar in that the lexicon consists of Word Formation Rules (WFRs) that operate on units to produce other units. The difference is that in Aronoff's theory the WFRs operate only on words, Anderson's WFRs operate on stems to produce other stems (where a stem could be a word). Thus, in a theory like this, the verb *baked* is not seen as being put together from two parts, *bake* and *-ed*, but as *bake* undergoing a rule of the form  $/X/ \rightarrow /Xed/$ , where the result represents a past tense verb. Aronoff puts an additional restriction on the WFRs, suggesting that they can only apply to words and output words. Thus, neither affixes nor bound stems are units in his theory. This introduces a complication: for example, how to derive *nominee* from *nominate*? Aronoff's solution is to introduce truncation rules, where *nominee* is derived from *nominate* by deleting *-ate* and attaching *-ee* at the same time. Anderson, finding such a solution ineffective, allows the existence of stems, possibly bound, to which the WFRs can apply. Both Aronoff and Anderson envision some sort of semantic function that goes along with a WFR and that maps the semantics of the base onto the semantics of the output of the rule, thus, semantics is not divorced from form, as in some IA theories.

Stump's (2001) theory is very similar to the theories of Aronoff and Anderson, especially Anderson's theory. Stump also assumes that affixes are not units, but just markers of processes that happen to words. He distinguishes three kinds of units: roots, stems and words. Thus, bound roots are units in his theory. The root is the ultimate default form, without any inflectional markings. A stem is an intermediate form that arises from application of rules and a word is a syntactically free form. The rules that apply to roots and stems to produce stems and words are paradigm functions that take a lexeme's root and a set of morphosyntactic properties and return a cell in the paradigm. These can also be units in IP theories. The paradigm is an additional characteristic that distinguishes Stump's, and also Anderson's, theories from Aronoff's theory. The paradigm plays a central role in that it structures the output of the paradigm functions and constitutes a separate unit in the morphology. While in Stump's model, the rules, i.e. paradigm functions, are defined through the paradigm, in Anderson's model the paradigm is a unit that is defined over the lexeme's stem set.

The processes that add different affixes to stems, also have a semantic function: they take the semantics associated with a stem and produce semantics associated with the complex item (Anderson, 1992). Thus, semantics is associated with lexical items, but it is not necessarily compositional (it is hard to speak of compositionality in an IP model, since there are no parts that are put together).

To summarize, IP models have the following general characteristics:

- Affixes are not units; they are mere reflections of the processes that happen to other morphological units.

---

<sup>1</sup>While the latter theory is usually classified as Word-and-Paradigm, it has many characteristics of an IP model.

- Roots and stems are either units (Anderson, 1992; Stump, 2001) or are not (Aronoff, 1976).
- In some theories, the paradigm is an additional unit (Anderson, 1992; Stump, 2001).
- Semantics is associated both with the input and the output of the processes that add affixes. However, since the process is a function that takes input and output, the semantics need not be compositional.

In some sense, IP models seem to be even more problematic than IA models for the Russian data I would like to consider, since they derive words from other words or stems. Thus, they assume that one item is basic and others are derived. This is not easy to determine in some cases we consider below.

- *bežat'* 'to run' impf., *pobežat'* 'to start running' pf. The most obvious approach in this case for IP models would be to consider *bežat'* the basic item and *pobežat'* as derived from it through the process of prefixation. Additionally, it seems that in a model like Anderson's (1992), *beža* would be a separate item, where word formation would start. A rule forming *bežat'* would be very general: adding *-t'* to a root forms an infinitive verb. However, only certain roots can form verbs, and those roots have to be specifically marked for the ability to form verbal infinitives.

Formation of the verb *pobežat'* from *bežat'* suffers from the same problem as in IA models, just cast in a different light. The addition of *po-* is the rule that transforms one into the other, however, there are different meanings that can be associated with this attachment. Again, not all verbs can form *po-*prefixed compounds with all its meanings. Thus, the rules will need to have the information about which verbal roots allow undergoing those rules and which do not.

- *lit'* 'to pour' impf., *polit'* 'to start pouring' or 'to water (flowers, plants, etc.)' pf. This pair is analogous to the pair above. Again, one complication is that there are two senses of the verb *polit'* and the rules will need to specify which verb roots allow which senses.
- *polivat'* 'to water (flowers, plants, etc.)' impf. The most logical procedure for deriving this verb would be two rules: one would add the prefix *po-* to *lit'*, as above, and the other would add the suffix *-iva-* to *polit'*. In this case, the IP analysis is superior to the IA analysis, since aspectual complications as in IA do not arise. In IP, *lit'* would change to perfective *polit'* and *polit'* would then change to imperfective *polivat'*.
- *pojasnjat'* 'to clarify' impf., *pojasnit'* 'to clarify' pf., *projasnjat'* 'to clarify' impf., *projasnit'* 'to clarify' pf., *vyjasnjat'* 'to find out' impf., *vyjasnit'* 'to find out' pf. Here the case is more complicated, since the root does not appear on its own, or as part of an unprefixated verb. It is not clear which item would be chosen as basic in this case,

and why. Would, for example, *projasnjat*’ impf. be derived from *projasnjit*’ pf. or the other way around and why? One solution would be to state that in general imperfective verbs are more ‘basic’ than perfective ones, that they are unmarked. This is what Comrie (1976) proposes for Russian. There are complications with this approach as well, since there are prefixed perfective verbs that do not have either an imperfective pair or a corresponding unprefixed verb, for example *pomeret*’ ‘to die’ pf.

- *kurit*’ ‘to smoke’ impf., *pokurit*’ ‘to smoke for a while’ pf., *prokurit*’ ‘to smoke for a while’ pf. Here the only issue that arises is similar to IA models: predicate-level meaning needs to be built into the rule that derives *prokurit*’ from *kurit*’.

### 1.3.3 Syntactic theories

The distinguishing feature of Syntactic theories is that all of words creation, except phonological feature insertion, happens in the syntax. Thus, attachment of affixes to roots is assumed to be governed by syntactic principles. Phonological feature insertion happens after the syntactic tree has been created, in a step called ‘Vocabulary Insertion’. Here I consider the Distributed Morphology model proposed by Halle and Marantz (1993). Additionally, I look at Baker’s (1988) incorporation theory, which is not only a morphology theory, but is relevant to the Russian verb data I will consider later on.

While separating syntax and phonology, Halle and Marantz identify lexical entries as relations between ‘bundles of morphosyntactic features’ and ‘bundles of phonological features’. Thus, in DM, both roots and affixes are units. Distributed Morphology is part of the Minimalist Syntax program, and thus, affixes as units are in a sense necessary, since affixes such as past tense are seen as syntactic heads in Minimalist Syntax.

As is standard in Minimalist Syntax, semantics is independent from both syntax and morphology and, as in Lieber’s theory, there is a separate semantics component that puts together form and meaning and uses lambda-style calculus (e.g., (Chierchia, 2000)).

Baker’s incorporation theory (1988) is very similar in its assumptions to DM. Morphology is assumed to ‘happen’ in the syntax, via syntactic operations. Semantics is assumed to be compositional (via lambda-style analysis, like in Lieber’s (1980) theory and DM), with idiosyncrasies listed in the lexicon. Various phenomena are considered from the point of view of syntactic incorporation, or head movement: causatives, applicatives, benefactives, etc. All of them are hypothesized to arise through different types of incorporation: noun, verb, preposition incorporation. Since the incorporated elements are usually bound and are heads that incorporated into other heads, it follows that bound elements, such as roots and affixes are separate elements in this theory.

Thus, DM and incorporation theory resemble IA theories in the structuring of morphological units: both roots and affixes are standalone items. However, meaning and phonological features are divorced from morphosyntactic features. A major difference between syntactic

models and all others is that words and word pieces are assumed to be syntactic heads in many instances. In the case of Russian, this brings back the problem of assigning aspect to a certain morpheme.

An incorporation analysis relates unprefixes verbs with prepositions to prefixed verbs, and has many of the assumptions of syntactic models such as Distributed Morphology.

Since these syntactic models have many properties that IA models have, they have the same types of problems associated with them. Consider how the Russian data would be analyzed in DM and preposition incorporation.

- *bežat'* 'to run' impf., *pobežat'* 'to start running' pf. The verb *bežat'* in a syntactic model such as DM would be formed from at least two morphemes: *bež(a)* and *-t'*. The latter would also constitute the Tense node in a syntactic tree. Usually syntactic models also assume a Asp (aspect) node as well (for example, (Franks, 1995, p. 183)). The problem that appeared in IA approaches - where to assign the imperfective value - appears here as well. A solution, which could also be applied to IA models, is to introduce null morphemes, and to mark aspect with a null morpheme associated with aspect properties. Null morphemes, however, are a last resort, and model that does not introduce them while describing everything else is superior to a model which does introduce them.

Baker's (1988) preposition incorporation analysis for deriving the verb *pobežat'* does not work here. Many Russian prefixes derive from spatial prepositions and even those that do not in many cases also have primary spatial meaning, thus a preposition incorporation account would be natural to consider. However, in this case it cannot apply. Consider examples in (1.1). While the verb *bežat'* does occur with the preposition *po-*, the corresponding sentence with the verb *pobežat'* is ungrammatical without it. Thus, an incorporation analysis cannot apply in this case.

- (1.1) (a) *Malčik bežal po doroge.*  
 boy.nom.sg run.impf.3.sg.masc.past along road.dat.sg  
 'The boy was running along the road.'
- (b) *Malčik pobežal po doroge.*  
 boy.nom.sg po-run.pf.3.sg.masc.past along road.dat.sg  
 'The boy started running along the road.'
- (c) \**Malčik pobežal doroge/doroga.*  
 boy.nom.sg po-run.pf.3.sg.masc.past road.dat/acc.sg  
 'The boy started running along the road.'

- *lit'* 'to pour' impf., *polit'* 'to start pouring' or 'to water (flowers, plants, etc.)' pf. This pair is similar to the pair of verbs above.

- *polivat* ‘to water (flowers, plants, etc.)’ impf. In this case, problems again arise with assignment of aspectual syntactic heads. If *po-* is considered a perfective aspectual head, how does that work with *-iva-* being an imperfective aspectual head? One solution would again be positing two homophonous items *po-*, where one is an aspectual head and another one is not. Either the verbal roots or the *po-* items would then need to be specified for compatibility with each other.
- *pojasnjat* ‘to clarify’ impf., *pojasnit* ‘to clarify’ pf., *projasnjat* ‘to clarify’ impf., *projasnit* ‘to clarify’ pf., *vyjasnjat* ‘to find out’ impf., *vyjasnit* ‘to find out’ pf. Similar problems arise here with aspectual marking.
- *kurit* ‘to smoke’ impf., *pokurit* ‘to smoke for a while’ pf., *prokurit* ‘to smoke for a while’ pf. A few interesting issues arise in the case of this triplet. All of these verbs can appear with two nouns in the accusative case in the same sentence:

- (1.2) (a) *Ivan*                 *kuril*   *sigaretu.*  
Ivan.nom.sg smoke.impf.3.sg.masc.past cigarette.acc.sg  
‘Ivan was smoking a cigarette.’
- (b) *Ivan*                 *kuril*   *sigaretu*                 *paru*  
Ivan.nom.sg smoke.impf.3.sg.masc.past cigarette.acc.sg couple.acc.sg  
*minut.*  
minute.gen.pl  
‘Ivan was smoking the cigarette for a couple of minutes.’
- (c) *Ivan*                 *pokuril*   *sigaretu*                 *paru*  
Ivan.nom.sg on.smoke.pf.3.sg.masc.past cigarette.acc.sg couple.acc.sg  
*minut.*  
minute.gen.pl  
‘Ivan smoked the cigarette for a couple of minutes.’
- (d) *Ivan*                 *prokuril*   *sigarety*  
Ivan.nom.sg through.smoke.pf.3.sg.masc.past cigarette.acc.pl  
*vsju*                 *noč.*  
whole.acc.sg night.acc.sg  
‘Ivan smoked cigarettes the whole night.’

Various generative accounts (for example, see (Pereltsvaig, 2000)) have trouble explaining this phenomenon, since, in general, the verb in a generative framework can assign only one accusative case. In the examples above, however, both the ‘true direct object’ (‘cigarette’ or ‘cigarettes’) and the time period (‘a couple of minutes’ or ‘the whole night’) are in the accusative case, and there is no ‘case-assigner’ other than the verb.

In this case, Baker’s incorporation account is an interesting alternative to consider. One of the properties that characterizes Baker’s preposition incorporation is that the



noun phrase that used to be the object of the preposition in an unincorporated sentence, behaves like a canonical direct object in the incorporated sentence (Marantz's Generalization, (Baker, 1988, p. 246)). This is exactly what happens in this particular case in Russian. However, there are a number of other consequences that are not compatible with the Russian data presented above.

Baker's analysis works for Russian data only if it is assumed that the verb can assign case to more than one NP, since there are two NPs in the accusative case in the Russian sentences. Baker does make that assumption, and cites Bantu data that has similar properties (pp. 264-266). On the other hand, the preposition incorporation analysis makes the prediction that intransitive verbs cannot have preposition incorporation, since they do not assign structural case to NPs (pp. 252-253). This prediction is not borne out by the Russian data:

- (1.3) *Ljuba*                    *prostojala*    *tam*    *celuju*  
 Ljuba.nom.sg   through.stand.pf.3.sg.fem   there   whole.acc.fem.sg  
*noč.*  
 night.acc.fem.sg  
 'Ljuba stood there the whole night.'

As we see in the above example, the structure similar to Baker's preposition incorporation can be used with intransitive verbs in Russian.

Another reason for not using the preposition incorporation account for the Russian data is that there are no structures with prepositions parallel to the ones cited above:

- (1.4) \**Ljuba*                    *stojala*    *tam*    *pro/čerez*   *celuju*  
 Ljuba.nom.sg   stand.impf.3.sg.fem   there   through   whole.acc.fem.sg  
*noč.*  
 night.acc.fem.sg  
 'Ljuba stood there the whole night.'

Thus, I conclude that the preposition incorporation account does not fit the Russian data.

### 1.3.4 WP and Network models

#### Word-and-Paradigm

Although several of the IP theories claim to be word-based, like Anderson (1992) and Stump (2001) above, they still make use of units smaller than the word - roots and stems. As Blevins (2006) notes, a truly word-based model does not recognize any morphological

units smaller than words. In such a model, ‘roots, stems and exponents [are] abstractions over a set of full forms’ (Blevins, 2006, p. 531).

Main postulates of a WP theory are developed in (Robins, 1959) and (Blevins, 2006). Apart from the word, the paradigm is the other morphological unit in a WP theory. A morphological analysis of a form in a WP theory can only be given in relation to other forms, to its inflectional paradigm (Blevins, 2006, p. 536). For example, consider the Russian noun *stol* ‘table’. The paradigm that *stol* belongs to is given in Table 1.2. Thus, a word like *stola* is considered both as a unit by itself and as a part of the paradigm illustrated, or, more specifically, as filling in either the Acc.Sg. or the Gen.Sg. cell.

	Singular	Plural
<b>Nominative</b>	stol	stoly
<b>Genitive</b>	stola	stolov
<b>Dative</b>	stolu	stolam
<b>Accusative</b>	stol	stoly
<b>Instrumental</b>	stolom	stolami
<b>Prepositional</b>	stole	stolax

Table 1.2: Paradigm of *stol* ‘table’

The paradigm in Table 1.2 thus plays two roles: it describes the paradigm of *stol* ‘table’ and also serves as an exemplary paradigm for Russian inflection 1. In a WP theory exemplary paradigms and principal parts help identify the paradigm of other words. Principle parts are forms that are predictive of the paradigm that an item would have. As Blevins notes (Blevins, 2006, p. 538), in Indo-European languages, the nominative singular is of highly predictive value, but that does not extend to other families.

As Robins (1959, p.128) notes, the morpheme does play a role in WP theories, but it is merely an abstraction over the paradigms and the words that are in those paradigms, and the morpheme cannot be coupled with meaning, as is done in IA and IP theories.

One theory that is developed in detail and is both word-based and uses paradigms is Bochner’s Lexical Relatedness Morphology (Bochner, 1993). As we will see, it also has characteristics of a network theory, such as (Bybee, 1985), to be discussed in the next section. His theory is truly word-based, and morphemes are not specified as independent units. Relationships between words are described using ‘rules’, such as the ones shown in (1.5) and (1.6).

$$\left[ \begin{array}{c} /light/ \\ A \\ \text{not weighing much} \end{array} \right] \leftrightarrow \left[ \begin{array}{c} /lighter/ \\ A \\ \text{weighing less than} \end{array} \right] \quad (1.5)$$

$$\begin{bmatrix} /X/ \\ A \\ Z \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xer/ \\ A \\ \text{more } Z \end{bmatrix} \quad (1.6)$$

Although the rules above are called 'rules' by Bochner, it is clear that they are static relationships and not processes, as the rules in an IP theory might be. The rules above describe the relationship between *light* and *lighter* (rule 1.5) and adjectives and the same adjectives ending in *-er* (rule 1.6). Although morphemes are not referenced as basic building blocks of word formation, both the root and the suffix are indirectly referenced in the rule above: *X* standing for the root and *-er* is the suffix. It is also evident that semantics is an inherent part of morphology in this theory, but that meaning is associated with a whole word, not a morpheme. Note, however, that while (1.5) assigns meaning to the word *light*, such a generic rule as (1.6) assigns meaning to a generalized pattern, *Xer*.

Other data structures in Bochner's theory are cumulative sets and cumulative patterns. Cumulative sets are sets of words that are constructed from a word using the rules such as above. For example, with the word *nice*, we can construct a cumulative set [nice, nicer] according to the rule above. Cumulative patterns are similar to both cumulative sets and rules: they are sets of generalized patterns. For example, one cumulative pattern could be [X, Xer, Xest] to describe the comparative and superlative constructions. This cumulative set could be seen as an inflectional paradigm of English adjectives and paradigms are exactly that: they are cumulative patterns of inflections.

To summarize, WP models have the following general characteristics:

- These models are clearly word-based. Morphemes, although recognized as abstractions over paradigms, are not units in and of themselves and are not paired with meanings. Words (and only words), on the other hand, are paired with meanings. Thus, there is no semantic compositionality involved, since affixes and roots are not associated with meaning.
- Apart from words, the other important unit in a WP theory is the paradigm. Inflectional patterns constitute paradigms, but so could derivational ones as well. Paradigms serve a dual function: they describe relationships between existing words and serve as exemplars for forming new ones.
- Therefore, neither roots nor affixes are units in a WP theory, while words and paradigms are.

## Network models and Construction Morphology

The three Network/Constructional Morphology models I will discuss here are Bybee (1985), Langacker (2002) and Booij (2005). Latest work in this field also include Gurevich

(2006) and Booij (2010). These two theories have some features in common with Bochner's (1993) model, and the latter is also in a sense a network model.

Network models are different from all of the above models in several respects. Most notably, these models take into account the frequency with which words are used by speakers. Thus, network models are usage-based models, as opposed to IA, IP and WP ones. Additionally, the two models described here are based on the idea of Construction/Cognitive Grammar (Goldberg, 1995; Langacker, 2002) that words and other higher-level constructions are pairings of meaning and form.

Bybee's model (Bybee, 1985) is based on two notions: lexical strength and lexical connections. Each word has its lexical strength, which is increased each time the form-meaning pairing is experienced. Thus, more frequent words will have greater lexical strength. Additionally, words are connected via 'lexical connections' to other words with the same semantic or phonological parts. Lexical connections between words form schemas, where new words can be fitted. Paradigms are one class of such schemas, since they are connections between words that are most closely related to each other. Thus, the word *cat* is connected to the word *cats*, since they share the segment *cat*, with the same meaning. In English, the pairing of *cat* and *cats* would constitute its own paradigm. The word *cats* also has a connection to the word *pots*, since they share the segment *s*, with the same meaning. Notably, although affixes have meaning associated with them (meaning that is shared between words they belong to), they are not stand-alone units; they exist only as parts of words.

In a sense, then, the only two units in Bybee's model are words (with the property of lexical strength) and connections between words. Paradigms are not separate units; they arise as an epiphenomenon of connections between related words. In that sense, Bochner's (1993) model is similar: it consists of listings of words and relationships between words, where paradigms arise as relationships between more closely related words. In contrast to Bybee, however, Bochner does not build in frequency as a defining factor (although he does mention that it is important).

Langacker's (2002) model is very similar to Bybee's. One difference is that he explicitly specifies the pattern that emerges from connections between instantiations of affixes. For example, the specification for the English plural suffix *-s* might look something like (1.7) (Langacker, 2002, p. 263), a *schema* in Langacker's terms. It is important to note that in order to be a truly word-based model, Langacker's specifications need to be changed. In his plural representation he specifies that separate meanings is associated with parts of the schema. In a truly word-based model, the meaning 'several Xs' would be associated with the schema's form.

$$\begin{array}{ll} \text{THING} & -s \\ \text{X} & \text{PL} \end{array} \quad (1.7)$$

The structure in (1.7) is very similar to Bochner's (1993) rules and different only in that it consists of one structure instead of two, as in Bochner's model. Bybee implicitly assumes

similar structures, but does not state what they should look like.

Booij (2005) makes such schemas explicit. He proposes schemas of different specificity: very abstract schemas that have almost no meaning specification, intermediate schemas, such as Langacker's plural specification ((1.7) above) and very specific schemas that describe individual lexical items. Booij's theory is not a network model as he specifies it, but it uses very similar ideas to Bybee's and Langacker's theories. The schemas he proposes are in a hierarchical relationship between each other, progressively from more abstract to more specific. The more abstract schemas appear once enough items with the same specification has been seen by the language user; presumably, this is the assumption of Bybee and Langacker as well.

In summary, network models have the following general characteristics:

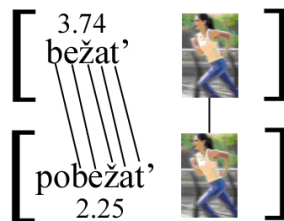
- The Network models are entirely word-based. Affixes and roots are not given a separate representation. Thus, as in WP models, meaning is associated only with words and there is no semantic compositionality to speak of. However, there exist semantic connections between words and parts of words.
- There are connections between words that are formed on the basis of phonological and semantic identity.
- The Network models are based on usage: frequency of words is taken into account explicitly. In Bybee's model, higher frequency words have larger lexical strength and weaker lexical connections to other words.
- Paradigms are sets of connections between closely related words. In a sense, they are an epiphenomenon of lexical connections.
- Thus, the two types of units in Network models are words and connections between words. Langacker also specifies schemas that are generalized over words and connections.

### **Analysis of Russian data in WP and Network**

For these models, we assume that each word is a pairing of meaning and form, there are connections between words and that there are generalizations over words with similar semantics and form.

- *bežat* 'to run' impf., *pobežat* 'to start running' pf. These two words will have a connection between them because of the form and meaning similarity. The representation will look something like Figure 1.1. There are both form and meaning connections, as well as frequency representations (log frequency).

An interesting question to consider is how aspectual meaning is represented in such a model. As was mentioned above, assigning aspectual meaning to parts of words is

Figure 1.1: Representation of *bežat'* and *pobežat'*

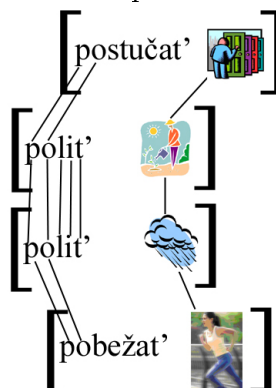
highly problematic for Russian data. What seems to be the case is that assigning it to the whole word might also not be right.

Consider the biaspectual verbs in Russian, of which there are a handful. Although ambiguous with respect to aspect in isolation, they are unambiguously identified as perfective or imperfective in an unambiguously perfective or imperfective context. For example, consider the verb *stabilizirovat'sja* 'to stabilize' pf./impf. (The example is from (Wade, 1992).) Although it is biaspectual, it is unambiguous in the following examples:

- (1.8) (a) *Položenije postepenno stabiliziruetsja.*  
 situation.neut.sg gradually stabilize.impf.pres.3.sg  
 'The situation is gradually stabilizing.'
- (b) *Položenije skoro stabiliziruetsja.*  
 situation.neut.sg soon stabilize.pf.fut.3.sg  
 'The situation will soon stabilize.'

While unambiguously perfective or unambiguously imperfective verbs only appear in one set of sentential contexts (perfective or imperfective), perfectivity or imperfectivity of ambiguous verbs is *defined* by the context where they occur. Thus, it seems that the unambiguous verbs are only unambiguous because they occur in only one set of contexts.

What this seems to imply is that aspectual meaning is not word-level, but predicate or sentence level meaning, at least for the ambiguous verbs. This would need to be tested empirically. If confirmed, it would provide more evidence for the Network theory of morphology, as no other morphological theory deals with sentence or predicate level meaning assignment. This meaning assignment would imply expanding the Network theory to include not only standalone words, but also constructions consisting of several words, essentially what Goldberg describes in her Construction Grammar (Goldberg, 1995).

Figure 1.2: Representation of *polit'*

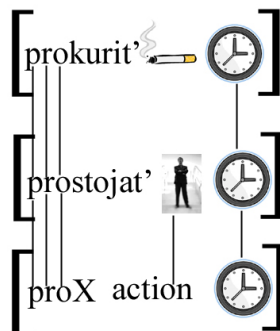
- *lit'* 'to pour' impf., *polit'* 'to start pouring' or 'to water (flowers, plants, etc.)' pf. Representation of these words is similar to the representation of *bežat'* and *pobežat'* above. The interesting part is the two meanings of *polit'*. They both must be represented. Both of the *polit'* items are connected to other verbs with similar meanings<sup>2</sup>.

For example, consider Figure 1.2. The inchoative verb *polit'* 'to start pouring' is connected to the inchoative verb *pobežat'* 'to start running', and the verb *polit'* 'to water the flowers' pf. is connected to the verb *postučat'* 'to knock' pf. A similar issue to the one above, aspectual meaning assignment, arises here. Only with context is it possible to decide which meaning of the verb is being used.

Another important point is that generalizations might arise about the prefix *po-*. It could emerge as a schema, where the meaning 'to start X' will be associated with form *poX*. The other option is that strong connections between similar words provide a pattern where a new word could fit. I discuss the two options in detail in the concluding chapter of the dissertation.

- *polivat'* 'to water (flowers, plants, etc.)' impf. This item will be connected to *polit'*, but not all of its segments are identical. It will also be connected to other items with the imperfectivizing suffix *-iva-*. The problems specified in IA theories are avoided: since we do not assign perfective or imperfective meaning to any word part, there is no conflict arising between the 'perfective' prefix *po-* and the 'imperfective' suffix *-iva-*.

<sup>2</sup>While I state that there are two verbs *polit'*, I do not give any significance to that statement, theoretical or otherwise. One form could be connected to several meanings, or there could be two items with the same form. At this point, I do not have the means to decide between the two solutions.

Figure 1.3: Representation of *prokurit'* and a general schema

- *pojasnjat'* 'to clarify' impf., *pojasnit'* 'to clarify' pf., *projasnjat'* 'to clarify' impf., *projasnit'* 'to clarify' pf., *vyjasnjat'* 'to find out' impf., *vyjasnit'* 'to find out' pf. As is stated above, these verbs are interesting in that they do not appear unprefixed. Therefore, there is assumed to be no independent item *-jasn-* that has to do with clarifying. However, all these verbs are connected, both through meaning and form. Thus, there might be a generalized item arising through these connections, similar to *po-* above. One difference between *po-* and *-jasn-* is that it seems harder to make meaning generalizations. There is parallelism between verbs prefixed with *po-*: a new verb can be used with *po-* if it fits the template. There is no such parallelism with words with bound roots. For example, the verbs *vyjasnit'* 'to find out' and *pojasnit'* 'to clarify' both have something in common with the adjective *jasnyj* 'clear', but it is hard to extract a generalization about *-jasn-* from all three words.
- *kurit'* 'to smoke' impf., *pokurit'* 'to smoke for a while' pf., *prokurit'* 'to smoke for a while' pf. As I already noted above, these verbs are interesting in that the prefix seems to introduce predicate-level constraints. This is again easily handled in the Network models by associating predicate-level meaning with the form. Thus, in the case of *prokurit'*, various verbs with the prefix *pro-* will be connected to it and a general schema might be extracted. If extracted, it would look like is shown in Figure 1.3.

A verb with *pro-* used in the durative meaning requires the time expression to be there:

- (1.9) (a) *Ivan prokuril sigarety vsju*  
 Ivan.nom.sg on.smoke.pf.3.sg.masc.past cigarette.acc.pl whole.acc.sg  
*noč.*  
 night.acc.sg  
 'Ivan smoked cigarettes the whole night.'





and thus it is not as 'expensive' to list the rule in 1.11 in the grammar when rule (1.10) is already there, as it would be if rule (1.10) were not listed.

$$\begin{bmatrix} /X/ \\ A \\ Z \end{bmatrix} \leftrightarrow \begin{bmatrix} /Xer/ \\ A \\ \text{more } Z \end{bmatrix} \quad (1.10)$$

$$\begin{bmatrix} /σ/ \\ A \\ Z \end{bmatrix} \leftrightarrow \begin{bmatrix} /σer/ \\ A \\ \text{more } Z \end{bmatrix} \quad (1.11)$$

Thus, the pattern-matching evaluation metric is an alternative way of counting independent information, with the assumption that not all information listed in the lexicon is independent, as IA models assume. Using this metric partly eliminates the need to only have the minimum amount of units listed in the lexicon. This eliminates the need to list only parts of regular words and only whole irregular formations: if using the pattern-evaluation metric, redundant information can be listed.

Additionally, as Langacker (Langacker, 2002, p. 262) argues, there is no it is not necessary to assume that the human brain works like a computer (minimal information is stored and the rest is computed) and that the human brain does not list redundant information. Thus, all other things being equal, WP and Network models should not be ruled out just on the basis of the fact that they list redundant information in the lexicon.

## 1.4.2 Semantics and semantic transparency

Semantics can be an important factor when deciding between morphological theories. Several studies (e.g. McCormick et al. (2008); Rastle et al. (2004) show that there is priming between words that are not related semantically, but are 'related' orthographically. Moreover, it seem that there are two components to priming: one orthographic, and one that is semantic. This is demonstrated by bound stem priming (to be discussed in more detail in chapter 3), where there is priming between words with bound stems and related words in visual mode priming (e.g. Forster and Azuma (2000)), while it disappears in cross-modal priming (Marslen-Wilson et al., 1994). Free stem priming, on the other hand, remains robust throughout experimental paradigms. The difference between free and bound stems used in the above experiments is in semantic transparency: bound stems are semantically opaque, while free stems are semantically transparent. Thus, semantic transparency is an important factor in morphologic processing.

What exactly is semantic transparency? This term can be used in several different ways: when describing a part of the word, a word, or a relationship between words. Semantic transparency for a part of a word, such as an affix, means how easy it is to describe that

part's meaning inside a word or words, and if it is consistent across different words. A word is labeled as semantically transparent when its meaning can be predicted from the meaning of its parts, and semantically opaque otherwise. For the priming experiments above semantic transparency is whether the meaning of one word can be partially predicted from the meaning of another word. If it can, the relationship of the two words is semantically transparent, and it is semantically opaque otherwise. All of these definitions have something in common - how easy it is to predict the part's meaning from its form.

While form has something to do with a part or a word's meaning, context is an important contributor. Consider the following examples (the following discussion is partly based on (Langacker, 2002; Fillmore, 1992; Sweetser, 1999)):

1. This beach is safe for the children.
2. The child is safe on this beach.

In the first sentence *safe* describes the beach, while in the second sentence it describes the child. In the first sentence the meaning of *safe* is 'free of potential hazards', and the second sentence it is 'will not find herself in hazardous situations'. Thus, depending on context, the same word can have different meanings.

More specifically, 'context' here means constructions where words occur (Goldberg, 1995). The more constructions two words have in common, the closer in meaning they are, when all other variables are unchanged. Presumably, for words, two words of the same part of speech are closer in meaning than two words that are of different parts of speech, simply because words that are both nouns or both verbs share more constructions than a noun and a verb.

In her discussion of semantic transparency, Hay (2001) names the two characteristics of a 'maximally transparent word'. It is the word that 'has neither shifted nor proliferated in meaning'. Thus, in a historic perspective, words that retain their meaning and do not add new meanings are the ones that are most semantically transparent.

Roots and affixes have different characteristics with respect to word meaning. Roots can occur in a variety of contexts (e.g., in words of different parts of speech), while affixes usually occur in a limited number of contexts, some only with one part of speech, for example. In that sense affix meaning is more restricted than root meaning, and is easier to generalize. For example, the Russian root *ryb* occurs in the following words: *ryba* 'fish', *rybnyj* 'fish' (adj.), *rybnik* 'fish seller', *rybalka* 'fishing trip', *rybak* 'fisherman'. Thus, although all the words have something to do with fish, it is impossible to specify the meaning further. On the other hand, some of the words that have the prefix *vy-* are *vyxodit'* 'to go out' (cf. *xodit'* 'to walk'), *vylezat'* 'to crawl out' (cf. *lezt'* 'to crawl'), *vystavljat'* 'to put out' (cf. *stavit'* 'to put'), *vyduvat'* 'to blow out' (cf. *duvat'* 'to blow'). The meaning of the prefix is much clearer than the meaning of the root in the above example, and there are many words that have the meaning of 'out' because of the prefix *vy-* in them. To be sure, other, non-transparent examples exist, such as *vystupat'* 'to present something at a conference', but there are many

words with the spatial meaning. What this seems to mean is that roots occur in a relatively small number of words and can be assigned relatively vague meaning, while affixes occur in a larger number of words, and have a small number of meanings that are easier to identify.

We can now reinterpret semantic transparency characteristics for word parts and specify criteria for semantic transparency a little more precisely. Since a root's meaning can be identified as 'having to do with X' where X can be almost anything, and since a root can have several meanings, a more semantically transparent root is the one where  $N(X)$  is minimum. Thus, a root where  $N(X)=2$  is more semantically transparent than a root where  $N(X)=5$ . On the other hand, for affixes we can usually discern a few meanings, and the affix with fewer meanings will be more semantically transparent than the affix with more meanings. Also, some meanings are easier to identify than others. Concrete meanings, such as spatial relations, are easier to identify than abstract meanings, such as time relations. Thus a prefix with a clearly identifiable spatial meaning will be more semantically transparent than a prefix without such a meaning. In addition to this, irrespective of the concreteness of the meaning, it is important if the meaning is productive, i.e. appears in a large number of words. The more of the meanings of a word part appear in large numbers of words, the more semantically transparent that word part is.

### 1.4.3 Frequency in morphology

In addition to semantics, frequency can be a distinguishing factor for deciding between morphological theories. One of the major advantages of the Network models is that frequency is a built-in factor, while all other models do not consider frequency to be important. Many studies have now shown that frequency effects are ubiquitous in morphology. (One such study is (Baayen et al., 1997), for a comprehensive review, see Hay (2001)). Thus, models that include frequency as a factor are at a significant advantage over other models that do not.

While the effects of frequency in morphology are well-known, it has been a topic of discussion in recent years as to which frequency is important: absolute frequency of words or relative frequency of words and their constituent parts. As Hay (2001) and Burani and Thornton (2003) show in their experiments, relative frequency is an important factor in morphological processing. Hay additionally argues that previous experiments that confirmed the importance of absolute frequency might have misleading results because of high correlation of absolute and relative frequency.

The Network models considered above only have the absolute frequency of words built in. For example, in Bybee's model, the more frequent a word is, the less strong are the lexical connections of that word with other words. Let us consider if relative frequency of words and their constituent parts can be worked into Network models and how.

As was mentioned in Section 1.3.4, Network models have two types of units: words and connections between words. Additionally, they have generalizations across words (explicitly mentioned in Bochner's and Langacker's theories), which give a unit-like status to affixes (see

rule 1.7 in section 1.3.4). Thus, frequency in these models can be associated with words and affixes, and possibly connections between items. No further apparatus is needed to represent relative frequency in the lexicon. There are a variety of ways of modeling relative frequency effects using these representations, I will mention one.

The gist of the relative frequency effect that has been proposed (Hay, 2001; Burani and Thornton, 2003) is that decomposition of a morphologically complex word happens if the constituent part, usually the root, is more frequent than the word itself, and that decomposition does not happen (or, more specifically, the decomposition route is slower in this case) if the constituent part is less frequent. Thus we might imagine that the lexical connection is weaker if the root's frequency is lower than the word's, and stronger if it is higher. Lexical connection strength is then a function of the frequencies of the two words it connects. Thus, Network models of morphology have a number of advantages as opposed to IA and IP models: they do not assume the symbol-counting metric, which allows for redundant information listing, and they include frequency as a factor, with the possibility of including relative frequency. Therefore, theoretical considerations would appear to favor Network models over all others.

## 1.5 Conclusion

In this chapter I have examined different morphological theories and how they account for Russian verbal data. Most theories are problematic in their treatments. Since the relationship between meaning and form is not one-to-one in Russian verbs, IA and S treatments would need to include null morphemes with meaning that cannot be assigned to any other morpheme and morphemes without meaning assigned to them. IP treatments are problematic because whether an item is basic or derived needs to be specified, and it is an arbitrary choice in many cases. Finally, WP and Network treatments are unproblematic, since they are not based on morphemes but words and their connections. I discuss evaluation metrics, frequency in morphological theories, and the notion of semantic transparency.

In Chapter 2 I examine what particular predictions about processing of Russian words, such as reaction times and complexity ratings, can be made based on the network theory of morphology. I go on to examine the status of roots in Chapter 3, the status of affixes in Chapter 5 and relative frequency effects in Chapter 4. In Chapter 6 I incorporate the results from this dissertation into the network theory of morphology.

## Chapter 2

# An investigation of relative frequency effects and the Network theory of morphology

As we saw in Chapter 1, the network theory of morphology (Bybee, 1988; Langacker, 2002) is attractive on several counts in its account for the semantics of prefixed verbs in Russian. This chapter addresses the question of relative frequency of words and their bases, that is suggested to be important in relations between words in recent studies (Hay, 2001, 2002; Burani and Thornton, 2003). I investigate how relative frequency would fit into a network theory of morphology and what experimental predictions are made on the basis of that fit. The experiments described here, with results reported in the next chapters, are important to morphological theory in the following ways:

- Status of roots, including bound roots, and affixes, is explored in the experiments. This is an important elaboration over Bybee's theory of morphology, which leaves the question of status of words and roots open. I propose a hierarchy of roots that results from differing strength of connections between words, instead of a dichotomy of free and bound roots.
- In addition to results in English (Hay, 2001) and Italian (Burani and Thornton, 2003), results from a morphologically complex language, Russian, are brought in, providing further evidence of the role of relative frequency in morphology.

As discussed in 1, the following morphological facts need to be accounted for in a theory of morphology:

- The intuition shared by speakers that words can be divided into parts, where each part has its own meaning. This intuition led to the well-known morphology theories, Item and Arrangement and Item and Process. However, avoiding the well-known problems

of IA and IP theories, such as empty, cumulative and portmanteau morphs (Anderson, 1992, pp.51-56), basic versus derived forms (see chapter 1 of this dissertation), and others, is also necessary.

- Paradigm effects, such as the basic-derived relation between related inflectional forms and degree of relatedness between those forms, which can vary (Bybee, 1985, pp.49-79), need to be accounted for.
- Frequency information needs to be taken into account. Numerous studies (Hay, 2001, 2002; Hay and Baayen, 2002; Burani and Thornton, 2003; Bybee, 2007 and others) have shown that frequency is a factor in morphological processing. For an overview, see (Bybee, 2007, pp.5-22).

The morphological theory that best fits these facts of different languages is the network theory of morphology that is word-based. In this chapter I discuss the role of relative frequency in this theory and experimental predictions based on it.

## 2.1 Theoretical preliminaries

In this section I describe the network theory of morphology and the necessary modifications that are based on recent experimental evidence.

Network theory of morphology was developed by several researchers, notably Bybee (1988) and Langacker (2002). Bochner's (1993) is in a sense a network model, although he does not explicitly state that. The idea of constructions (Construction Grammar and Morphology Goldberg (1995); Booij (2005); Gurevich (2006)) is also used in the Network theory, in the sense that all words are constructions, or pairings of meaning and form. The following are the most important characteristics of the network models:

- There are only two types of units in these models: words and connections between words. These models are truly word-based: there are no roots or stems, as in other models (Item-and-Arrangement, Item-and-Process, Word-and-Paradigm).
- Each word has the property of lexical strength, which is a function of its overall frequency.
- The other type of units are connections between words based on phonological and semantic identity. Connections can be of different strength.
- Paradigms can fit naturally into the structure of these models, as connections between different words can be ascribed different characteristics. Words in a paradigm can have closer connections than words not in a paradigm.

- These models are usage-based. One of the main characteristics of usage-based models is that word frequency is included as a factor, as opposed to other models.

I will now describe Bybee's and Langacker's network theories in more detail, considering how more recent evidence of relative frequency effects fits in. The following discussion is based on (Bybee, 1988) and (Langacker, 2002, pp.261-288).

While words and connections are the only standalone units in Bybee's theory, it is important to mention that Bybee's conception of 'words' is left open:

In this model, morphologically complex items are stored in the lexicon, and I refer to them as "words" although it is conceivable that some may be larger than traditional words, and some may be smaller, and there even be typological differences among languages regarding the size of the lexical unit. (Bybee, 1988, p.126)

Thus, Bybee does not explicitly state that her model is word-based, and leaves the question open; however, she also states that in her model words are not separated into parts, but the parts are identified (as in Construction Morphology), and this is a clue to a truly word-based model.

The only other units present in this model, apart from words, are connections between words. These connections can be of two types: phonological and semantic. Both types of connections can be formed between either identical and non-identical components (i.e., ones that only have some features in common). Since the character of connections is different for different words, connection strength is a property of all connections. When both semantic and phonological connections run in parallel, the words connected are subject to morphological analysis: parts with identical semantic and phonological connections are identified, but the word is not "split" into separate parts, meaning that those parts can be recognized, but are not stored. However, according to the strength of the connections, words can be more or less related (as is demonstrated in priming and rating experiments), characterized by what Bybee calls "degree of relatedness", or cumulative strength of semantic and phonological connections between two words. Paradigms fit naturally into this model: lexical connections between words in a paradigm are stronger than with other related words.

In addition to words that are connected by semantic and phonological connections, there is a property of lexical strength that is associated with each word, which is an index of word frequency. Bybee sees two effects that result from differences in word frequency. The first one is resistance to change in phonological shape: the more frequent the word, the less likely it is to regularize or be subject to suppletion. The second one is lexical and inflectional splits that occur inside a paradigm: words that are more frequent tend to dissociate from the words that they are related to.

Langacker's theory (Langacker, 2002, ch.10-11) is very similar to that of Bybee's, with the main difference being in his extensive use of higher-level schemas. While Bybee proposes only word-level units that are connected between themselves via semantic and phonological connections, Langacker proposes more abstract units, such as [PROCESS... PAST], a schema



that would fit any past-tense verb. However, he also admits (p. 288) that such higher-level schemas might not be useful for speakers, as their level of abstractness might render them inapplicable to most specific items. I do not choose between the two options and leave this as a question for future research.

I now consider how recent research in relative frequency and status of roots fits into this network theory of morphology.

## 2.2 Relative frequency effects

Several recent studies (Hay, 2001, 2002; Hay and Baayen, 2002; Burani and Thornton, 2003) suggest that, other than absolute frequency of a word, the relative frequency of the word and its base might be a defining factor in morphology and morphological processing. Below I review some of the effects found by these researchers.

The first effect is the route of morphological processing. It has been disputed over the years whether words are processed by morphological decomposition, i.e. necessarily breaking the word up into its component parts (see e.g. Butterworth 1983), or accessing the word as a whole (see e.g. Taft 1985). More recent studies (Caramazza et al. 1988; Baayen 1992; Frauenfelder and Schreuder 1992) argue for a dual-mode processing route, where which route is picked for each word depends on its frequency: the more frequent the word, the more likely it is to be processed as a whole, and not decomposed into its component parts. However, as Hay (2001) argues, experimental results on this correlation between frequency and route of processing are at best contradictory. She argues instead that the processing route depends on relative, not absolute frequency, where relative frequency is the relationship between the word (for example, *conception*) and its base (*conceive*). Her experiment with English speakers, where she solicited complexity judgments, confirms that if the word is more frequent than its base, it is judged less complex than if the word is less frequent than its base.

Another effect is a secondary effect of the route of morphological processing; it is how relative frequency affects affix ordering. Hay (2002) shows using many English affixes that their usual split into level 1 and level 2 (Siegel, 1979) with the stipulation that level 2 affixes cannot appear outside level 1 affixes is better explained by considering relative frequency of words that are associated with those affixes, along with other factors, such as morphotactics. If an affix can be found inside a large number of words that are less frequent than their bases, that affix is more productive, and hence can appear outside other affixes. If, on the other hand, an affix can be found inside a large number of words that are more frequent than their bases, it will be less prone to decomposition, and less likely to appear outside other affixes. The reason is that words with affixes that are less prone to decomposition are more likely to be treated as one unit, i.e. not decomposed into parts.

In addition to the above English experiments, similar experiments were carried out in Italian (Burani and Thornton, 2003). These were lexical decision experiments, where the frequency of the component parts (root and suffix) were controlled for. However, the derived

word frequency was held constant and was always less than both the root and the suffix frequency. Hence, under the theory described above, all those words should have been prone to morphological decomposition. However, the authors also found that words with a more frequent root were processed faster than words with a less frequent root, and no effect of suffix frequency. Hence, even if a word is accessed via the decomposition route, it may be accessed faster or slower, depending on root frequency.

Since relative frequency seems to be a fundamental factor in how words are processed morphologically, it is important to investigate it further. While the arguments for relative frequency effects are based on decomposition of words versus whole word processing, and the Network theory of morphology is based on whole word processing only, the argument is still important, but should be rephrased as follows. Based on previous research, the connections between words, as hypothesized by the Network theory of morphology, might be stronger or weaker, depending on the relative frequency of words and their bases. The stronger connections between parts of words might make them easier to identify. Experimental predictions based on network theory of morphology and effects of relative frequency are discussed below.

## 2.3 Status of roots

In both theoretical and experimental literature the question of status of roots is intensively debated. The main questions that are raised are whether roots by themselves are units, and whether bound roots and free roots<sup>1</sup> have the same status. Theories of morphology take one of the two positions on root status: either they are morphological units along with words (Bloomfield, 1933; Lieber, 1980; Kiparsky, 1982; Stump, 2001), or, as in word-based theories, only words are morphological units (Robins, 1959; Bochner, 1993; Blevins, 2006). Anderson's (1992) theory probably fits more with the root-based theories, as his *stems* are 'word[s] minus (productive) inflectional affixation' (Anderson, 1992, p.71). In word-based theories there is no need to distinguish between free and bound roots, as roots are not units. Theories that make use of roots as morphological units do not in general make a distinction between bound and free roots in the sense of properties of those roots. Thus, although it is recognized that bound roots do not appear by themselves, they have the same exact properties and subject to the same processes as free roots. For example, Stump (2001) assumes roots as morphological units, and makes no distinction between free and bound roots. Similarly, although Anderson (1992) bases his theory on words minus productive inflection, he states that internal structure must be accessible, even for bound roots like *-ceive* (p.297). This kind of non-distinction between free and bound roots makes sense for a theory that is based

---

<sup>1</sup>While the traditional definition of bound roots is roots that never occur by themselves, we can identify at least two types of such roots: semantically transparent and semantically opaque bound roots. In English bound roots are mostly semantically opaque, while in many other languages they are semantically transparent, but do not occur by themselves because of inflectional requirements. In the literature cited above it is usually the semantically opaque bound roots that are discussed.

on roots, since roots are standalone units in such a theory, and all morphologically related words share the same root.

A number of researchers have sought to confirm or disprove the hypothesis that free and bound roots have the same status through priming experiments. Forster and Azuma (2000) report the results of a masked priming experiment where priming between two prefixed words sharing a bound stem (e.g. *submit* and *permit*) and priming between a free stem word and its prefixed relative (e.g. *fold* and *unfold*) are compared. They find that priming is the same for both types of pairs, with no significant difference between them. A second masked priming experiment showed that bound stem priming did not differ from orthographic priming (e.g. *shallow-follow*). A final experiment with closer distractors and a longer prime duration showed priming for bound stems that was significantly different from orthographic controls. Based on these results, the authors conclude that both free and bound stems are units in the mental lexicon.

Pastizzo and Feldman (2004) performed a masked priming experiment with primes and targets sharing either a free or a bound stem. They found that decision latencies to both bound stem and free stem targets were significantly faster after a morphological prime than an unrelated prime. Based on significant priming after a morphologically related prime for both free and bound stem targets, the authors likewise conclude that free and bound stems are equally important in the mental lexicon and constitute units in processing. Interestingly, accuracy rates were significantly higher for free stem words than for bound stem words, and free stem targets were reacted to slower after an orthographically related prime than after an unrelated baseline, while the pattern for bound stems was different: targets after a morphological prime were reacted to faster than after an unrelated baseline (the same after an orthographic prime). This is illustrated in Table 2.1. This difference effectively means that orthographic primes interfere with lexical decision for words with free roots, but not bound roots, which I interpret as meaning that words with free roots have more connections to other words than words with bound roots, and it is harder to distinguish words with free roots from their orthographic neighbors.

	<b>Prime type</b>		
<b>Stem type</b>	Unrelated	Orthographic	Morphological
Free	760 ms	780 ms	736 ms
Bound	760 ms	755 ms	738 ms

Table 2.1: Masked priming results from (Pastizzo and Feldman, 2004).

On the other hand, Marslen-Wilson et al. (1994) performed a cross-modal priming experiment where they failed to find any priming between prefixed words with the same bound stem (such as *submit* and *permit*) while there was robust priming between prefixed words with a free stem (such as *unfasten* and *refasten*). The authors conclude that to represent a word as morphologically complex synchronic semantic links to other morphologically complex words

are needed, thus stating that words with bound roots are represented as morphologically simple in the mental lexicon.

What these results seem to point to is a hierarchy of root types, where words with both free and bound roots elicit priming effects, but those effects are different, and depend on methodology used in the experiment. Thus, in (Pastizzo and Feldman, 2004), although words with both free and bound roots exhibited morphological priming, there were differences between them (accuracy rates were significantly different and reaction times after orthographic primes were faster than after unrelated primes for bound roots, while the pattern was the reverse for free roots). Forster and Azuma's (2000) masked priming experiment showed the same amount of priming between prefixed words with free and bound stems. However, only in an experiment with close distractors and longer prime duration did the difference between bound stems and orthographic primes arise. Thus, again, although there was priming for both free and bound stems, bound stems were different from free stems. Finally, in a cross-modal priming experiment (Marslen-Wilson et al., 1994), there was priming for free stems and not for bound stems. Cross-modal priming used by Marslen-Wilson et al. (1994) seems to exaggerate differences between words with free and bound stems. What all these results show is that priming for bound stems is harder to demonstrate than for free stems, and thus free and bound stems must have different properties.

Overall, while Marslen-Wilson et al. (1994), Pastizzo and Feldman (2004) and Forster and Azuma (2000) do not agree on their conclusions on whether free and bound roots have the same status, they all agree on the fact that roots of some kind do exist as units in the mental lexicon. Other researchers present results that also suggest unit status of roots. The evidence comes from different angles. Burani and Thornton (2003) and Cole et al. (1989) found cumulative root frequency effects for Italian and French, respectively (only for suffixed words in the French case), Caramazza et al. (1988) found differences in processing of differently structured nonwords in Italian, Feldman (1994) found that shifting of letter sequences is significantly faster for morphemic sequences than for non-morphemic in Serbian, Reid and Marslen-Wilson (2003) found priming between words with the same stem in Polish, Fiorentino and Poeppel (2007) found differences between processing of monomorphemic words and compounds matched on whole word frequency in English, Melinger (2003) found more errors in a speech production study in English for morphologically complex words as opposed to morphologically simple words, and Marslen-Wilson et al. (1994) found significant priming in a cross-modal priming task for words sharing the same stem.

On the other hand, Manelis and Tharp (1977) (English), Giraudo and Grainger (2000) (French) failed to find evidence for roots as units. Manelis and Tharp (1977) found no differences in processing between affixed and unaffixed words in a lexical decision task and a decomposition task. Giraudo and Grainger (2000) found no cumulative root frequency effects in a lexical decision task.

Whether roots are considered units or not, differences between the two types of roots must be taken into account. Neither position can accommodate the differences demonstrated by (Pastizzo and Feldman, 2004), (Forster and Azuma, 2000) and (Marslen-Wilson et al., 1994).

If roots are units, both free and bound roots are represented as standalone elements with no differences between them. If no roots are units, there are again no differences between free and bound roots, as they do not exist.

The above experiments were all done assuming a dichotomy between two root types, but in some languages, it is possible to identify more than two types of roots, bound and free. One of those languages is Russian, where at least three types of roots exist. In addition to bound and free roots, there are also modified roots and roots bound within part of speech. Modified roots are roots that are related to other roots, but have a slightly different form, and roots bound within part of speech are roots that do not have unprefixes in their own part of speech, but do have them in others. The following pairs of words illustrate this: *molč-at* ‘to be quiet’ and *po-malk-ivat* ‘to keep quiet’ (modified roots), *pojasnjat* ‘to make clear’ and *jasnost* ‘clarity’ (\**jasnjat*). The experiment in Chapter 3 explores this question in detail.

## 2.4 Experimental predictions

The experiments in this dissertation address some of the shortcomings of the experiments described above. The English experiments were paper-and-pencil, not reaction time based, and a reaction time based experiment is a more appropriate measure of online processing<sup>2</sup>. Additionally, English is not a very rich language morphologically. Although Italian is a language much richer in morphology, and the experiments were RT-based, relative frequency was always the same - skewed in the direction of morphological decomposition. The experiments in this dissertation are based on reaction time in a morphologically rich language, Russian. In addition, it also explores other questions not addressed above, such as the status of bound roots.

The first experiment was a computer reaction time experiment, where participants were asked to answer ‘yes’ or ‘no’ as quickly and as accurately as possible to the question ‘Does this word contain the prefix *po-*?’ The expected results for such an experiment are as follows: if the word is usually decomposed, the reaction time should be shorter than if the word is not usually decomposed, since in a word that is usually decomposed it is easier to separate out the prefix. As is described above, words that are usually decomposed are the ones that are less frequent than their bases.

The second experiment was a complexity rating experiment involving 5 Russian prefixes of similar productivity, but of different semantic transparency, *po-*, *pod-*, *za-*, *vy-* and *ot-*. The question asked in the experiment was ‘How complex in structure is this word?’ and the answer was a rating on a 1 (not complex) to 10 (very complex) scale. The expected result of

---

<sup>2</sup>However, the reaction times resulting in experiments in the following chapters are too long to be indicative of online processing, and replications of these experiments need to be done with other measures, such as lexical decision.

the experiment is that the more semantically transparent a prefix is, the more complex the word containing it will be rated.

Now consider the specific predictions based on these experiments.

### 2.4.1 Root status

As is described above, the only units in the network theory are words and connections between words. Bybee does not give a definitive opinion on whether roots should be considered units or not, and it is interesting to investigate that question further. As we saw in Chapter 1, different morphological theories address this question differently. Item-and-Arrangement and some Item-and-Process theories see the root as the smallest unit of (idiosyncratic) meaning, while others (Word and Paradigm, Network theories) do not view roots as units of morphology. In the theories that do recognize roots as morphological units, as with other morphemes, usually free and bound roots are distinguished (e.g. (Carstairs-McCarthy, 2002, p.18)). Free roots are defined as those that can stand alone, and bound roots never occur by themselves.

Most theoretical treatments of Russian, Slavic and related Slavic languages (e.g. (Townsend, 1975), (Gladney, 2006), (Svenonius, 2004), (Lunt, 2001)) are IA theories, and thus rely heavily on the notion of the root. IP (e.g. (Corbett and Fraser, 1993) - although it claims to be paradigm-based, it is also essentially an IP account, (Elson, 1997)) and Distributed Morphology treatments (e.g. (Weisser, 2006)) also rely on the notion of the root. The only Word and Paradigm (WP) treatment is (Blevins, 2006).

Analyses that do include the notion of the root usually distinguish between free and bound roots. This traditional distinction works well for English, where many bare roots are also standalone words. However, in Russian, a more sophisticated definition is needed, since there is no such clear dichotomy between free and bound roots, and bare roots are rare, with most roots occurring with some inflectional morphology. Some bare roots occur in the imperative (e.g. *moj* ‘wash!’) or in the past tense (e.g. *njos* ‘he carried’), but these are relatively rare. Thus, in identifying ‘free’ versus ‘bound’ roots in Russian verbs I used the heuristic of whether the verb occurs unprefixes or not. If it does occur unprefixes, it falls into the ‘free’ category, if it does not, it falls into the ‘bound’ category, with further subtypes of bound roots. I identify the following types of roots:

- Completely bound roots. Verbs with these roots never occur unprefixes, although they might appear with several prefixes. An example of such a verb is *postigat* ‘to understand’, ‘to befall’. This verb also occurs with other prefixes, e.g. *nastigat* ‘to befall’, but not unprefixes.
- Roots bound within part of speech. Since verbal prefixes are different from prefixes used with other parts of speech (Townsend, 1975, p.20), I hypothesize that the relationship between a prefixed verb and the corresponding unprefixes verb is substantially different from the relationship between the prefixed verb and an unprefixes noun/adjective with

the same root. Thus, these verbs do not occur unprefixated, but other parts of speech with the same root do. An example of such a verb is *ponižat* ‘to lower’. There is no corresponding unprefixated verb, although the root does occur unprefixated in other parts of speech, e.g. *nižnij* ‘lower’ (adj.).

- Modified roots that are otherwise not bound. These roots occur in verbs that are not seen unprefixated, but are connected to other unprefixated verbs with the same root. One example is the verb *pomalkivat* ‘to keep quiet’ that does not occur unprefixated, but is clearly related to the verb *molčat* ‘to be quiet’.
- Free roots. Verbs with these roots occur unprefixated, e.g. *xlopat* ‘to clap’ (impf.) and *poxlopat* ‘to clap’ (pf.).

If we just take into consideration relative frequency, the results should appear fairly discrete: words that are more frequent than their bases should have a certain range of reaction times and words that are less frequent than their bases should have another range of longer reaction times. However, as Frauenfelder and Schreuder (1992) discuss in the description of their dual-route morphological processing model, when there is a race between two routes, processing is faster than when there is only one route of access. The amount of speed-up (statistical facilitation, Raab (1962)) depends on the amount of overlap between the two routes. This model is consistent with the results from Italian experiments (Burani and Thornton, 2003), where all the words are decomposed, according to the relative frequency hypothesis, however, reaction times are different depending on the frequency of roots and affixes. Considering this modification, we might suppose that the results will be more like a continuum in the present experiment as well, depending on the structure of the word. This is discussed below, when I consider the expected results for the different types of roots selected for this experiment and how the above hypothesis affects processing of those roots.

- Completely bound roots. These roots are never seen unprefixated, which means that derived frequency is always larger than base frequency for words with these roots. In this respect we would expect the highest reaction times when asked if the word contains a certain prefix, since these words should not be broken down morphologically. However, since the prefixes are identified in these words, and roots do have connections to roots in other words, however weak, we might expect that more frequent roots will result in shorter reaction times than less frequent roots.
- Roots bound within the class of verbs. While these roots are seen unprefixated, it does not happen within verbs, only other parts of speech. As a result, for these items, the derived word is still always more frequent than the base word, since the base words do not exist. However, the root has connections to other words, which are much stronger than in verbs above. Thus I expect reaction times that are shorter than for roots that are bound completely.

- Modified bound roots, otherwise free. The bases of these words are never seen unprefixes, since these modified roots only happen after prefixes. However, the free roots that these modified roots have connections to, do appear unprefixes, and thus the frequency of these roots (both modified and free combined) might have a facilitating effect on processing these words. Thus, although they are not decomposed, they are hypothesized to be processed faster than the words with two root types above.
- Free roots. Depending on whether the frequency of the base words, these words will either be or will not be decomposed. If the word is more frequent than the base, it will not be decomposed (reaction time will be longer), and if the word is less frequent than the base, it will be decomposed (the shortest reaction time). However, even if the word is not decomposed, its root is free, and thus has strong connections to other words, and I expect reaction times that are longer than for words with free roots that are decomposed, but shorter than the words above.

To summarize, in the model I am proposing only words are units, and thus, for free roots, the relative frequency of word and its base is the most important factor that leads to decomposition when the base word is more frequent than the derived word. However, lexical (semantic and phonological) connections between word parts contribute to morphological processing, and although words with bound roots are not morphologically decomposed, their processing is facilitated by the status of the root: words with completely bound roots are processed slower than words with roots that are modified, but otherwise free, for example.

## 2.4.2 Prefix status

The goal of the other experiments in this dissertation is to compare the processing of different prefixes in Russian, *po-*, *pod-*, *za-*, *vy-* and *ot-*. Hay and Baayen (2002) demonstrate that an important factor in determining an affix's productivity is the ratio of words that are decomposed versus words that are not decomposed associated with a particular affix. Thus, an affix that is associated with more words that are less frequent than their bases, or more prone to decomposition, will be more productive than an affix that is associated with more words that are more frequent than their bases, or less prone to decomposition. While the prefixes in the experiment have similar productivity, based on this metric, they are different how consistently their different meanings are used, or semantic transparency. They differ in semantic transparency, as defined by the degree of consistency with which their different meanings are used. They also vary in number of meanings (polysemy) and in concreteness of meaning. As is discussed in Section 1.4.2, the more meanings a word part has, and the less concrete they are, the less semantically transparent that word part is. Thus, I expect that in the complexity ratings experiment words containing prefixes that have more and less concrete meanings will be rated as less complex.

This result, if confirmed, would also fit into the network theory. The more semantically transparent a prefix is, the stronger are the semantic connections to other words with the



same prefix. Thus, stronger semantic connections of prefixes will manifest themselves as more complex ratings of the words including those prefixes. This view of prefixes is consistent with a completely word-based model of morphology. However, this view must also be consistent with form and semantic connections existing not only between *words*, but also *parts of words*, such as prefixes. This could mean that generalizations over many words containing a certain prefix with strong form and semantic connections could be stored. I discuss this further in Chapter 5.

## 2.5 Summary

I propose a model of morphological processing that combines the following pieces of previous research: network theory of morphology (Bybee, 1988; Langacker, 2002), where only words (with the property of lexical strength) and lexical connections are units, relative frequency effects (Hay, 2001, 2002; Hay and Baayen, 2002; Burani and Thornton, 2003) and the dual-route processing model (Frauenfelder and Schreuder, 1992). In this model, relative frequency plays a major role in whether a particular word is processed via the direct or the decomposition route, while lexical connections between words facilitate that processing, independent of which route is used.

Based on the proposed model and using Russian, a morphologically complex language, I created a set of experiments that will shed light on both the status of roots and affixes and relative frequency effects. The experiments should show a clear difference in processing words that are more frequent than their bases versus words that are less frequent than their bases: the latter ones should be processed faster than the former ones. These results will add to cross-linguistic evidence of importance of relative frequency as a factor in morphological processing.

In addition to that, I expect the roots to form a continuum, where free roots facilitate processing of words, even if processed via the direct route, the most, and completely bound roots, the least. In addition, we should see a difference in processing of words with different prefixes: words with the more semantically transparent prefix should be rated more complex than words with the less frequent prefix.

Thus, we expect two types of effects: one on the level of the word (each word that is less frequent than its base should be decomposed), and at the level across words: roots and affixes with stronger lexical connections should facilitate the processing, even though the word itself might not be decomposed.

## Chapter 3

# Root status and cumulative root frequency effects

### 3.1 Introduction

As described in 2, status of roots, both free and bound, is subject of debate in both theoretic and experimental literature. In this chapter I investigate the status of roots using a prefix separation reaction time experiment. The purpose of the experiment described below is to further investigate the question of root status, both free and bound, adding to the cross-linguistic evidence and using Russian as the study language. In the experiment I used words with different types of roots in Russian, including words with completely bound, bound within part of speech, modified, and free roots. These types of roots include both the semantically opaque bound roots (completely bound roots) and semantically transparent bound roots (modified roots, and roots bound within part of speech). The results of the experiment show clear differences in processing between words with free and different types of bound roots, where the four types of roots form a hierarchy. These findings support a model based on processing of whole words, with links between shared parts that create roots effects.

### 3.2 Root status experiment

In this experiment I look at different root types in Russian, with the hypothesis that different root types are processed differently and that cumulative root frequency does not contribute to said processing. The experiment involved looking at Russian words and answering the question, ‘Does this word contain the prefix *po-*?’ Data items all had the prefix *po-*. The hypothesis is that, depending on root type, the words will have faster or slower reaction times. The types of roots used in the experiment are described in detail in Section 2.4.1, repeated here for convenience.

- Completely bound roots. Verbs with these roots never occur unprefixes, although they might appear with several prefixes. An example of such a verb is *postigat* ‘to understand’, ‘to befall’. This verb also occurs with other prefixes, e.g. *nastigat* ‘to befall’, but not unprefixes.
- Roots bound within part of speech. Since verbal prefixes are different from prefixes used with other parts of speech (Townsend, 1975, p.20), I hypothesize that the relationship between a prefixed verb and the corresponding unprefixes verb is substantially different from the relationship between the prefixed verb and an unprefixes noun/adjective with the same root. Thus, these verbs do not occur unprefixes, but other parts of speech with the same root do. An example of such a verb is *ponižat* ‘to lower’. There is no corresponding unprefixes verb, although the root does occur unprefixes in other parts of speech, e.g. *nižnij* ‘lower’ (adj.).
- Modified roots that are otherwise not bound. These roots occur in verbs that are not seen unprefixes, but are connected to other unprefixes verbs with the same root. One example is the verb *pomalkivat* ‘to keep quiet’ that does not occur unprefixes, but is clearly related to the verb *molčat* ‘to be quiet’.
- Free roots. Verbs with these roots occur unprefixes, e.g. *xlopat* ‘to clap’ (impf.) and *pohlopat* ‘to clap’ (pf.). While I call these roots ‘free’, most of them, as most roots in Russian, do not occur by themselves because of inflectional requirements. ‘Free’ here means ‘occurs unprefixes’.

Words with completely bound roots are predicted to have the slowest reaction times in decisions about presence of a prefix, words with free roots to have the fastest reaction times, and words with modified roots and with roots bound within part of speech are predicted to be in between those two extremes.

### 3.2.1 Subjects

41 native speakers of Russian living in the New York City greater area and the San Francisco Bay area took part in this experiment. They were monetarily compensated for their participation.

### 3.2.2 Stimuli

Stimuli included 48 data items and 50 filler items. Out of 48 data items 8 words had completely bound roots, 10 had roots bound within part of speech, 10 had modified roots and 20 had free roots. Half the 50 filler items were verbs that start with *po*, but did not contain the prefix *po-*. The other half were words that did not begin with *po*.

### 3.2.3 Procedure

The experiment was run on a laptop computer with 2 keys labeled clearly in Russian ‘yes’ and ‘no’. Each participant saw instructions for the experiment with two sample questions with feedback. The instructions explained what constitutes a prefix and how to answer questions. The concept of prefix was explained by showing that the word *izbegat* ‘to avoid’ contains the prefix *iz-* and the root *-beg-*. The subjects were instructed to respond as quickly and as accurately as possible. After that the experiment was run, with each subject receiving a different random rotation of the items. Each item was presented in the middle of the computer screen, with the question ‘Does this word contain the prefix *po-*?’ on top and the instructions ‘Please press YES or NO’ on the bottom. There was no time limit and no feedback for the questions.

### 3.2.4 Results

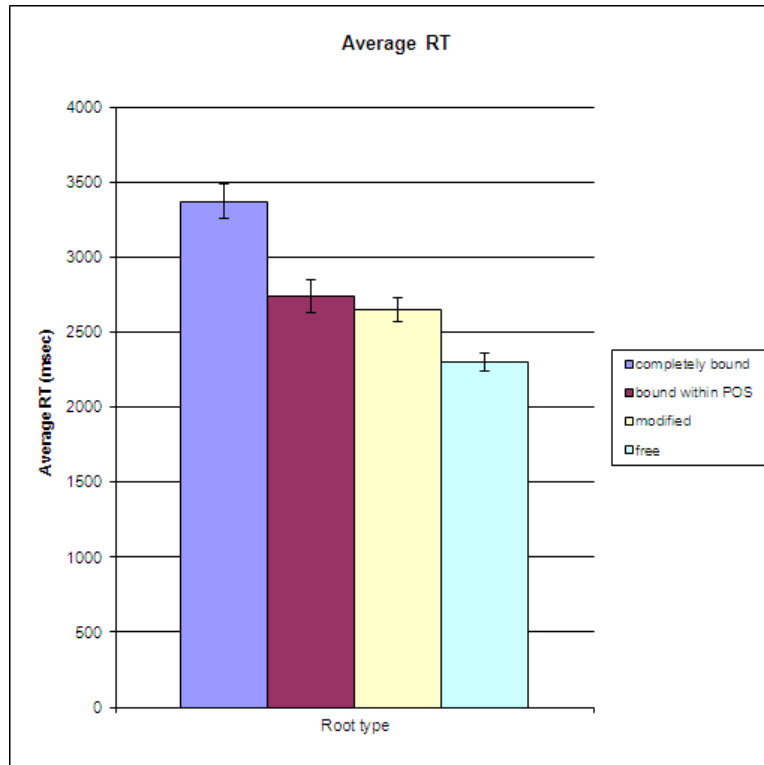
Responses of four subjects were taken out, as their accuracy rates were lower than 75%. In addition, 6 items containing the reflexive suffix *-sja* were taken out, as the presence of this suffix complicated the structure of the word by a degree and complicated data analysis. The average reaction time results are presented in Figure 3.1. Planned by-subject and by-item analyses of the data are presented in Tables 3.1 and 3.2. Completely bound roots were reacted to slower than roots bound within part of speech ( $p < 0.001$  for both by-subject and by-item comparison). Roots bound within part of speech were reacted to slower than modified roots ( $p = 0.03$  in the by-subject comparison, but  $p > 0.1$  in the by-item comparison). Modified roots were reacted to slower than free roots ( $p < 0.01$  in the by-subject comparison and  $p = 0.01$  in the by-item comparison). The only comparison that was not significant is the by-item comparison between modified roots and roots bound within part of speech, which I attribute to the small number of items compared (10 and 8, respectively). Thus, what we see is indeed a continuum of different root types: completely bound roots, roots bound within part of speech, modified roots and free roots (from slowest to fastest reaction time).

Root type	Average RT	Standard deviation	t-test
Completely bound	3440 ms	1092 ms	$p = 0.0004$
Bound within verbs	2720 ms	1040 ms	$p = 0.03$
Modified	2495 ms	820 ms	$p = 0.007$
Free	2203 ms	773 ms	

Table 3.1: By-subject RT averages for different root types. p-values are calculated for the comparison of the root type in that row with the root type in the following row.

Comparing the mean reaction times of the words with the four root types, we see that their processing is different. In order to more precisely investigate the factors contributing to the processing of these words, a mixed effects regression model was built. In the model, I

Figure 3.1: Experiment 1 results



Root type	Average RT	Standard deviation	t-test
Completely bound	3365 ms	260 ms	$p = 0.0003$
Bound within verbs	2748 ms	303 ms	$p = 0.3$
Modified	2656 ms	364 ms	$p = 0.01$
Free	2298 ms	351 ms	

Table 3.2: By-item RT averages for different root types. p-values are calculated for the comparison of the root type in that row with the root type in the following row.

used the logarithm of the reaction time as the dependent variable, since the residuals of the model better fit the assumptions when the dependent variable was log-transformed. Subject number and item number were random effects in the model. The following variables were tried as fixed effects: logarithm of word frequency, logarithm of cumulative root frequency, accuracy, length in letters and syllables, the phonotactic transition (V-CV or V-CC) between the prefix and the root, root type (4 types: completely bound, bound within part of speech, modified and free), semantic transparency (entered the word as semantically transparent if the definition of the prefixed word on <http://www.gramota.ru> contained the unprefixing base and semantically opaque if that was not the case), trial number, family size. Insignificant effects were eliminated one by one, using the ANOVA test for comparing models. The model was then refitted.

I found the following fixed effects to significantly contribute to the model: the type of root (RT for type of root: completely bound < modified < bound within part of speech < free, compared using level contrasts as described in (Crawley, 2007)), semantic transparency (facilitatory effect), a quadratic effect of verbal unprefixing family size, or the count of all unprefixing verbs with the same root (initial facilitatory, later inhibitory effect) and trial number (facilitatory effect).

The model has an  $R^2$  of 0.34, a small reduction from the  $R^2$  of the model with just random effects (0.35). However, the reduction in variance is 88% for the random effect of item number (stays the same for the random effect of subject number). The estimates and  $p$ -values for the different factors based on Markov Chain Monte Carlo sampling are shown in Table 3.3 and the variance and standard deviation for random effects are shown in Table 3.4.

<b>Factor</b>	<b>Estimate (Log RT)</b>	<b><math>p</math>-value</b>
Intercept	7.9003	$p = 0.0000$
Completely bound versus bound within POS	0.1035	$p = 0.0007$
Free versus modified	-0.0638	$p = 0.0023$
All bound versus free and modified	0.0607	$p = 0.0071$
Semantic transparency	-0.1655	$p = 0.0000$
Trial number	-0.0015	$p = 0.0006$
Verbal unprefixing family size	-0.0384	$p = 0.0600$
Verbal unprefixing FS (quad term)	0.0035	$p = 0.0200$

Table 3.3: Fixed effects of the mixed effects regression model.

The trial number effect shows that subjects got faster as the experiment progressed, probably getting better at reacting to the task. Another interpretation is that the participants were faster as the experiment progressed because they made policy decisions about their responses. The verbal unprefixing family size is facilitatory at first, but then becomes inhibitory. This is similar to the U-shaped family size effect that was found by Wurm et al.

Random effect	Variance	Standard deviation
Item number	0.004345	0.065916
Subject number	0.073566	0.271230

Table 3.4: Random effects of the mixed effects regression model.

(2006), the difference being that the present family size measure only includes unprefixes verbs with the same root. In addition, the model confirms the results of the analysis of the means: words with completely bound roots are the slowest, words with free roots are the fastest, and words with modified roots and roots bound within part of speech are in between, with former slower than the latter. The means analysis showed that the difference between modified roots and roots bound within POS was significant only by subjects and not by items, and in this respect the mixed effects model provides somewhat different results. A more detailed investigation with new data items would be needed to confirm or disprove that particular difference. This points to a hierarchy of roots, where free and completely bound roots are at the extremes and modified roots and roots bound within part of speech are in between. Finally, there is a significant factor of semantic transparency, where semantically transparent words elicit a faster reaction time than semantically opaque words. I turn to a detailed analysis of these effects next.

### 3.2.5 Discussion

The average reaction time elicited in this experiment is high: it ranges between 2200 and 3400 ms, while usually average reaction times in psycholinguistic studies are between 500 and 700 ms. Thus, the prefix separation experiment was an off-line task, as compared to the other studies. The latter studies usually involve lexical decision (where priming experiments also include lexical decision). The question ‘Is this a word of language X?’ is less complex than the question ‘Does this word contain the prefix Y?’. The second question necessarily involves some kind of analysis of the word, while the first not necessarily so. I believe this complexity of the second task is the reason why the reaction times are higher than in lexical decision tasks.

The question that the experiment participants were answering was ‘Does this word contain the prefix *po-*?’. I presume a slow reaction time to mean that it is relatively harder to separate the prefix from the rest of the word and a faster reaction time to mean that it is relatively easier to separate the prefix. Thus, it is easiest to separate the prefix from a free root word, and hardest from a bound root word, with roots that are bound within part of speech and modified roots in between. This means that the rest of the word stands out the most in a word with a free root, and the least in word with a bound root.

Let us now look at the reasons why the rest of the word could stand out more in a word with a free root than in word with a bound root, or root bound within part of speech, or modified root. As Hay (2002) argues in her investigation of affix separability, several different reasons could be at play: phonotactics, relative frequency and proportion of more decomposable words associated with that affix. Thus, a word that contains a less probable within-word phonotactic transition will be more decomposable than a word than contains a more probable within-word phonotactic transition. A word that is less frequent than its base will be more decomposable than a word that is more frequent than its base. And finally, a word with an affix that is associated with more words that are more decomposable will be more decomposable itself than a word with an affix that is associated with less decomposable words. While the last factor is not important in this investigation, since all the words contain the same prefix *po-*, both phonotactics and relative frequency are important.

To see if phonotactics was important in this particular experiment, I coded whether the word had a V-CC or a V-C transition between prefix and root. The prefix is V-final; some roots had one consonant in the beginning, while some had two. Model comparison strongly suggests that including this factor did not improve the model ( $p = 1$ ). While phonotactics is important in how separable a prefix is, in this particular experiment it plays no role. In addition, more sophisticated analyses of phonotactics might be more informative.

Relative frequency is a factor in a sense that free root words have corresponding unprefixes bases, while words of all other root types do not. Thus, the prefix is more easily discernible in free root words than in all other words because the same words without the prefix exist. If unprefixes bases are important, then so are other unprefixes words with the same root. In terms of the Network theory of morphology, these words, unprefixes bases and other unprefixes relatives, are important because of the lexical connections formed between related words. And while words with completely bound roots do not have connections to unprefixes words with the same root, words with other types of bound roots do. Thus, words with roots bound within part of speech are connected to unprefixes words with the same root, just not in the same part of speech. Words with modified roots have connections to unprefixes words with the same root, but the root is of a slightly different shape. These connections to other unprefixes words seem to be the defining difference between different types of roots, and it parallels the reaction time results: the more and the closer matched the connections, the more separable is the prefix from the root. I carry out a more detailed investigation of relative frequency effects on the data from this experiment that only includes free roots in the next chapter.

The importance of these connections to unprefixes words with the same root is confirmed by the presence of another factor - unprefixes verbal family size, or the number of unprefixes verbs with the same root. This factor follows the same U-shaped pattern, as in Wurm et al. (2006): it facilitates reaction time for moderately large families, and then inhibits it when the family grows too large. The authors explain this effect by arguing that a larger family size makes the word likely, but only to a certain point, when the family is so large that it inhibits word identification. The question that arises in the present experiment is why



the contributing factor is unprefixing verbal family size. Both restrictions on family size probably stem from the nature of the experiment. All items presented to subjects were verbs, restricting the domain of operation to verbs. The question asked was ‘Does this word contain the prefix *po-*?’ and retrieving a verb with the same root, but without a prefix, is one easy strategy to decide whether the word with *po-* contains a prefix. An interesting question is whether the grouping by part of speech and only of unprefixing words happens online or stems from storage. In either case, this grouping must be readily available to the language user. This is an argument for language schemas that are more general than just individual lexical items.

Finally, semantic transparency provides a facilitatory effect. If the prefixed word is semantically transparent, it is easier to decide that it contains *po-*, since it is easier to connect that word to other words with the same root and separate the word into its constituent parts.

If the encoding of root type is just a reflection of the underlying connection strength (completely bound roots having the weakest connections, free roots having the strongest connections, and the other two types in the middle), then we can attempt to model the connections in a more gradual way. To accomplish this, I built another multiple regression model, where instead of root type I entered number of form links with the closest unprefixing neighbor divided by word length, my attempt at a crude estimate of connection strength. For words with completely bound roots, I used 1 for the number of links, so that no singularities appear in the statistical matrix. I call this variable the ‘links parameter’. This is similar to the ‘physical overlap’ between words computed by Napps (1989) for her experiment. For the example words above, the links parameter is calculated in Table 3.5. It can vary from 0 (where a word has no unprefixing relatives) to 1, although in theory the links parameter can never reach 1, since that would represent connections to an identical word.

Root type	Word	The links parameter
Completely bound	<i>posjagnut</i> ‘to infringe’	1/9 = 0.11
Bound within POS	<i>pojasnjat</i> ‘to make clear’	4/8 = 0.5
Modified	<i>pomalkivat</i> ‘to keep quiet’	5/11 = 0.45
Free	<i>pohlopat</i> ‘to clap’	7/9 = 0.78

Table 3.5: The links parameter for the four types of words.

Summarizing the number of words with a particular links parameter, where the links parameter is sorted from 0 to 1 gives us Table 3.6. We see almost complete non-overlap between root types, where the links parameter becomes progressively higher with root type. The only overlap that occurs is between modified and bound within part of speech roots, the same types of roots that had one non-significant p-value in the comparison of the means. Thus, the links parameter is the same as the root type parameter, just ordered.

The multiple regression model using the links parameter instead of the root type parameter is summarized in Table 3.7 and the variance and standard deviation for the random

Links parameter	Root type			
	Completely bound	Bound within POS	Modified	Free
0.1	<b>1</b>	0	0	0
0.111	<b>5</b>	0	0	0
0.125	<b>2</b>	0	0	0
0.3	0	<b>1</b>	0	0
0.333	0	<b>1</b>	0	0
0.375	0	<b>5</b>	0	0
0.4	0	0	<b>1</b>	0
0.444	0	<b>3</b>	0	0
0.455	0	0	<b>2</b>	0
0.462	0	0	<b>1</b>	0
0.556	0	0	<b>1</b>	0
0.583	0	0	<b>3</b>	0
0.6	0	0	<b>2</b>	0
0.75	0	0	0	<b>4</b>
0.778	0	0	0	<b>7</b>
0.8	0	0	0	<b>2</b>
0.818	0	0	0	<b>3</b>
0.833	0	0	0	<b>3</b>
0.846	0	0	0	<b>1</b>

Table 3.6: Number of verb with a particular links parameter, links parameter sorted from 0 to 1.

effects are summarized in Table 3.8. The resulting model has an  $R^2$  of 0.34, the same as the model using the root type parameter, and the variance for the item number random effect is reduced by 87% (the variance for the subject random effect stays the same). The links parameter has a facilitatory effect on reaction time.

<b>Factor</b>	<b>Estimate (Log RT)</b>	<b><i>p</i>-value</b>
Intercept	8.0926	$p = 0$
Links parameter	-0.4097	$p = 0$
Semantic transparency	-0.1655	$p = 0$
Trial number	-0.0015	$p < 0.001$
Verbal unprefix family size	-0.0353	$p = 0.08$
Verbal unprefix FS (quad term)	0.0028	$p = 0.05$

Table 3.7: Fixed effects of the mixed effects regression model.

<b>Random effect</b>	<b>Variance</b>	<b>Standard deviation</b>
Item number	0.0044989	0.067074
Subject number	0.0736326	0.271353

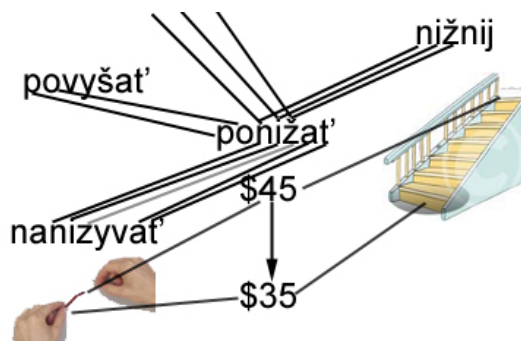
Table 3.8: Random effects of the mixed effects regression model.

As we see, the model with the links parameter is very similar to the model with root type.

The results of this experiment fit well with the Network theory of morphology, where lexical connections between words are emphasized. I illustrate these interword connections with *ponižat'* ‘to lower’ using the same model as presented in (Bybee, 1988). This word is connected to many other words with the same root, including *nižnij* ‘lower’ (adj.) and *nanizyvat'* ‘to string’. These connections are shown in Figure 3.2. There are form connections between the words, where the connection between *z* of *nanizyvat'* ‘to string’ and *ž* of *ponižat'* ‘to lower’ is weaker than the form links between identical segments. There are more form links between *ponižat'* ‘to lower’ and *nanizyvat'* ‘to string’, since they are both verbs. There are semantic connections as well: all the words share an up and down structure with the focus on the element that is below. There are also connections to words that have the same prefix as *ponižat'*, illustrated with connections to the word *povyšat'* ‘to raise’. There are additional connections to other words that share the same root, as well as connections to other words that share other word parts that are not reflected in the figure.

The word used in the example, *ponižat'* ‘to lower’, is a word with a root bound within its part of speech: there are no unprefix verbs with the same root, but there are nouns and adjectives with the same root that do occur unprefix. As we see from the illustration above, there are strong semantic and form connections between *ponižat'* and other words.

Figure 3.2: Form and semantic connections between words.



Strong semantic and form connections also occur between words with modified roots and their morphological relatives, as well as between words with free roots and their morphological relatives. However, in most cases, only form connections are present between words with completely bound roots and their morphological relatives, since the semantic connections are nearly always absent or are very weak.

The three most important factors in the links between words are number of form connections, semantic transparency and family size. What separates words with free roots from words with other types of roots are unprefixed bases with almost identical form connections and strong semantic transparency. On the other hand, what separates words with completely bound roots is semantic opacity. Words with modified roots and roots bound within part of speech are in between: there are clear semantic connections, but not as many form connections. Thus, this model parallels the experiment results: completely bound roots are the least separable, free roots are the most separable, and modified roots and roots bound within part of speech are in between. We see that even very strong form links (as in completely bound roots, where everything except the prefix is linked to another word) are not nearly enough in making roots separable, while semantics plays a very important role, and even when there are not very strong form links present (as in bound within part of speech and modified roots), the presence of semantic transparency makes those roots separable.

This model not only parallels the results of the present experiment, but also corroborates the results of the experiments by Pastizzo and Feldman (2004), Forster and Azuma (2000) and Marslen-Wilson et al. (1994), where a hierarchy between bound and free roots emerged. A hierarchy of roots is exactly what follows in the present experiment: completely bound roots, root bound within part of speech, modified roots and free roots, where form links and semantic transparency are the hierarchy placement factors.

Summarizing the results of this experiment, we see that the best predictors of prefix separability are root type, unprefixed verbal family size and semantic transparency together,

where root type can be modeled equally well with a non-numeric four-way parameter and a numerical links parameter that represents the number of links between a word and its closest unprefixes relatives.

### 3.3 Are roots units of representation in the mental lexicon?

The results of the present experiment together with previous research show that roots indeed form a hierarchy. The hierarchy is based on the number of links with closest unprefixes relatives and semantic transparency, as demonstrated in the analysis above. Do these results support a view of roots as units of morphological processing?

Based on the results of the experiment in this chapter, we can conclude that a hierarchy of roots exists. Several possible underlying phenomena are compatible with these results. The first, most obvious interpretation, is that there are four different types of roots in Russian and that, depending on the type of root, they are more or less separable in the word. This kind of explanation fits morpheme-based theories, since roots can then be separate units, and there can be several types. Different languages would then have different numbers and types of roots. The other interpretation fits the word-based theories. It supposes that the different root types are just a convenient way of coding the strength of connections between different words. This explanation is superior in that it does not postulate different numbers and types of roots for different languages; the underlying mechanism is the same no matter what the language. In addition, it avoids the difficulties associated with the linearity of morpheme-based models that is problematic for languages like Hebrew and Arabic. Since connections can be both linear and non-linear, they work equally well for linear and non-linear languages. The two models above show that both root type and connection strength (simplistically coded) model the morphological phenomena equally well. Thus, I propose that the results reported in this chapter are most compatible with the view that roots are not units on their own, but are epiphenomena that arise over connected words, just as suggested by (Blevins, 2003).

An argument might be raised that roots might be units, but only in form, not involving any meaning connections. This is possible, and does not contradict the model proposed here. In fact, several studies found priming between words that are orthographically similar, but not morphologically related (e.g., Rastle et al. 2004; McCormick et al. 2008). If there is only storage of roots as form units, it would be no different than storage of any orthographically similar elements.

Frequency effects arise from storage (Stemberger and MacWhinney, 1988), thus stored items give rise to frequency effects, and the absence of a cumulative root frequency effect argues against storage of roots. Moreover, cumulative root frequency effects, in the studies where they are demonstrated (e.g., (Burani and Thornton, 2003) and (Cole et al., 1989)),

are not necessarily arguments for root storage. Clearly, family size effects in lexical decision experiments (e.g., Baayen et al. 1997) show that both the word in question and words connected to that word morphologically are accessed during a lexical decision task. Cumulative root frequency effects might arise for the same reason. In addition, cumulative root frequency and family size are usually correlated. For example, for the words used in the present experiment, there is a highly significant positive correlation of family size and logarithm of cumulative root frequency (0.77,  $p < 0.0001$ ). It is possible that in the studies that do find a cumulative root frequency effect, a better predictor would be family size. In addition to this, cumulative root frequency effects, when they are reported (e.g., Wurm 1997), are *inhibitory*, not *facilitatory*, which is an argument against root storage.

We can now reinterpret previous research results that provide evidence for roots as independent units. In Section 2.3 I list numerous studies from several languages that conclude that roots are stored. The number of studies ‘for’ roots as units is overwhelmingly greater than the number of studies ‘against’ roots as units. I also present evidence for root effects in this study. However, these effects need not arise from storage of roots, but from form and semantic connections between words, as hypothesized in the Network model of morphology.

The root continuum in Russian varies from completely bound roots to roots bound within part of speech to modified roots and finally to free roots. Semantic relationships between words with completely bound roots are almost completely opaque, and the question arises, are words with completely bound roots more tightly linked than words with very similar form and no semantic connections? This question has been asked by several researchers. Forster and Azuma (2000) find the same amount of priming for free and bound stems, as well as orthographic controls. The difference appears only after increasing the prime duration and choosing closer distractors. Both Rastle et al. (2004) and McCormick et al. (2008) find priming between semantically unrelated words such as *committee* and *commit*. On the other hand, Beauvillain and Segui (1992) argue against priming based solely on orthography. Based on the present results we can make the prediction that words that have a purely orthographic relationship to other unprefixated words would be lowest on the root hierarchy, lower than words completely bound roots, which do have some, albeit vague, semantic links.

### 3.4 Conclusions

In this chapter I have added to the evidence on the controversy present in previous research, where it appears that both free and bound roots are units with somewhat different properties. These different properties of roots do not fit either the root-based morphological model, where both free and bound roots exist and are not different from each other, nor the word-based morphological models, where existence of the root is denied altogether. I report the results of a prefix separation experiment that uses Russian words with different types of roots.

I show that words with different types of roots form a continuum, and are not all the same. The prefix is separated the fastest from words with free roots, and slowest from words with completely bound roots, with words with modified roots and roots bound within part of speech in between. If roots were units, a hierarchy like this would not be possible, since all roots would have the same status, and would be equally separable. The differences are illustrated both with a mean analysis and using a multiple regression statistical model. In the model, the following factors are important: root type, semantic transparency, unprefixing verbal family size (a quadratic effect) and trial number. I discuss how these results may be explained in two different ways: one with different root types, in a morpheme-based model, and another with varying strength of connections, in a word-based model. The latter explanation is superior since it fits any kind of language, including languages with non-linear morphology, and there is no need to postulate different numbers and types of roots for different languages.

Overall, the controversy that did not fit either the root-based nor the strictly word-based morphological models fits well with the Network theory of morphology: depending on the strength of connections, the roots appear either more unit-like (free roots), or less unit-like (bound roots). While the experimental results do not provide conclusive evidence for or against roots being standalone units in the lexicon, the Network theory of morphology explains the experimental results without requiring roots to be units. The root hierarchy found in Russian is a further argument against root storage, as it is difficult to model such storage with different roots taking different places in the hierarchy.

All together, these results fit well with the explanation that form links together with semantic connections, both between whole words, are the factors that are important in morphological processing, presenting strong evidence for the Network morphological model that is based on whole words and emphasizes lexical connections between them.

## Chapter 4

# Relative frequency effects

### 4.1 Introduction

In the previous chapter we established that an important factor in morphological processing is form and meaning links between words, whether or not roots are units in the said processing. While in the last chapter I considered roots and how their storage might be affected, in this chapter I look at the status of prefixes, and what factors might influence their processing. One of the factors that might be important in processing of prefixes, and affixes in general, is the relative frequency of words and their bases. I address the question of whether the relative frequency of words and their bases affects the processing of Russian prefixes in this chapter.

As we saw in Chapter 3, the question of whether or not word parts are stored is a central one in theories and models of morphology. A variety of experiments has been carried out in the last four decades aimed to answer the question of whether words are decomposed into their constituent parts, or processed as a whole. Proponents of decomposition (e.g., Taft (1985)) argue that every word is decomposed into its constituent morphemes and a lexical search is carried out on the root. On the other hand, proponents of whole-word processing argue that all words are accessed as one whole entity (e.g., Butterworth (1983)). The latest view is that both routes of processing, whole-word and decomposition, exist. In different models each encountered word is processed using both routes, and the faster one prevails (Frauenfelder and Schreuder, 1992), one is employed for known words, the other for novel (Caramazza et al., 1988), or both operate at the same time (Wurm, 1997).

In race models, determining which route prevails in a particular item is usually done by manipulating cumulative root frequency and surface frequency of that item. In this experimental paradigm, cumulative root frequency is defined as the combined frequency of the root of the word and surface frequency is the frequency of the word as a whole. For example, Taft (1985) cites the following frequencies for the following words: *approach* 123, *reproach* 3, *persuade* 17, *dissuade* 3. These are the surface frequencies for those words.



Cumulative root frequencies of *proach* and *suade* are 126 (123+3) and 20 (17+3), respectively. In an experiment, if the surface frequency is held constant and the cumulative root frequency is manipulated, faster reaction times for higher base frequency items is taken as evidence that that set of words is accessed via the morphological decomposition route. If, on the other hand, cumulative root frequency is held constant and surface frequency is manipulated, and more frequent items elicit a faster response, it is taken as evidence that the direct route is favored for words whose frequency is large enough. Usually, the set of words tested includes words with a certain affix and the findings are assumed to apply to all words with that affix. Several studies used this methodology in prefix stripping experiments. Cole et al. (1989) argue against prefix stripping and for suffix stripping. In a set of French lexical decision experiments, the authors find differences between processing of prefixes and suffixes. They find a significant difference in reaction time between suffixed words with high versus low cumulative root frequency, but no significant difference in reaction time between prefixed words with high versus low cumulative root frequency. They explain this effect by proposing that, since words are processed with the prefix first, then root and then suffix, root effects only appear in suffixed words, where the root is processed first. Also in French, Giraudo and Grainger (2003) find opposite results. In masked priming experiments, they find prefix priming, but not suffix priming. Based on three English experiments, Taft and Forster (1975) propose a model of word recognition based on the root where the prefix is stripped first. In a theoretical investigation, Schreuder and Baayen (1994) show that such a model would be highly inefficient and thus improbable.

Other studies show that an important factor in morphological processing not taken into account in the above experiments is relative frequency (Cole et al., 1997; Hay, 2001, 2002; Burani and Thornton, 2003; Zuraw, 2009). Relative frequency is the difference between the frequency of the derived word and the frequency of its base. Using the English words *approach* and *inaccurate* I illustrate these terms:

1. Derived frequency: frequency of the word itself. For both *approach* and *inaccurate* that would be the word frequency.
2. Base frequency: the frequency of the unprefix word. Since *\*proach*, the base of *approach*, is a bound root, the base frequency of *approach* is zero. On the other hand, *accurate*, the base of *inaccurate*, exists as a separate word, and thus the base frequency of *inaccurate* is the frequency of *accurate*.

Hay (2002) predicts that words that are more frequent than the bases they contain are accessed via the direct route, and that words that are less frequent than the bases they contain are accessed via decomposition. For example, a word like *inaccurate* (frequencies are from (Hay, 2002)) should be accessed via decomposition, since the derived frequency of *inaccurate* (53) is less than its base frequency (377). A word like *unleash*, on the other hand, should be processed as a whole word, since its derived frequency (65) is larger than its base frequency

(16). These predictions were borne out in (Hay, 2001), where she asked subjects to provide judgments on relative complexity of pairs of words. She asked subjects to rate which word in a pair was more complex, one that is more frequent than its base or the one where the base is more frequent. Consistently, subjects rated words that are less frequent than their bases as more complex. Results of Hay's experiments in English are corroborated by Burani and Thornton's (2003) results in Italian and Zuraw's (2009) results in Tagalog. According to Hay, relative frequency and derived frequency are highly correlated, and previous experimental results might be inconsistent because of this. Additionally, relative frequency plays a role in determining affix productivity. Hay and Baayen (2002) show that relative frequency is one of the most important factors in determining affix productivity, where affixes that are associated with more words whose derived frequency is less than their base frequency (and thus these words are presumed to be decomposed) are more productive. What this means for the dual route models is that access to the morphological route might depend on relative frequency of base and derived words. It is plausible that the prefix stripping results described above are contradictory because relative frequency was not taken into account in the design of those experiments.

In this chapter I present an analysis of productivity of two Russian prefixes, *po-*, a very productive prefix, and *voz/vos/vz/vs-*, an unproductive prefix. This analysis shows that the correlation of base and derived frequencies and the slope and intercept of the regression line on these two variables are all important predictors of productivity of these two prefixes. Once the analysis is complete, I use relative frequency information to analyze the reaction time of free root data from the experiment in Chapter 3. The analysis shows that relative frequency is an important factor in morphological processing of Russian verbs, suggesting that the decomposition of the prefix out of the words depends on the relative frequency of the derived and base words. This evidence, together with previous results in English, Italian and Tagalog, suggest a universal principal of organization and should be taken into account by models of morphological processing.

As in describing status of roots (Chapter 3, the Network theory of morphology once again provides an adequate model for these results. The strength of connections could be dependent not only on the form or semantic connections, as described in Chapter 3, but on relative frequency of the word and its base as well. Thus, for example, a word like *inaccurate* would have strong connections with its base, *accurate*, since *inaccurate* is less frequent than *accurate*. On the other hand, *unleash* would have weaker connections to its base, *leash*, since *unleash* is more frequent than *leash*.

## 4.2 Po- and voz- productivity analysis

### 4.2.1 Prefix descriptions

The two prefixes I chose for the productivity analysis are *po-* and *voz/vos/vz/vs-*. *Po-* only has one form, while *voz/vos/vz/vs-* has four allomorphs: *voz-*, *vos-*, *vz-*, and *vs-*. *Voz-* and *vz-* occur before vowels and voiced consonants, while *vos-* and *vs-* occur before voiceless consonants. In the rest of the chapter, I refer to the latter prefix as just *voz-* for simplicity. Intuitively, the two prefixes are different in their numeric characteristics and also in their meanings. Townsend (1975) lists the following meanings for the two prefixes:

*Voz-*:

1. Up: physical or abstract.  
*vs-prygnut* 'to jump up'  
*vos-pitat* 'to bring up'
2. Intensity or suddenness.  
*vs-kriknut* 'to utter a sudden shriek'  
*vz-boltat* 'to shake up'
3. Back.  
*voz-vratit* 'to return'  
*voz-obnovit* 'to renew'

*Po-*:

1. Begin to.  
*po-nesti* 'to start carrying'  
*po-ljubit* 'to become fond of'
2. Do for a short time.  
*po-sidet* 'to sit for a while'  
*po-govorit* 'to have a talk'
3. Do somewhat, to some extent.  
*po-lečit* 'to cure a little bit'  
*po-veselit* 'to amuse somewhat'

4. Do from time to time and/or with diminished intensity.

*po-kurivat* ‘to smoke from time to time’

*po-čityvat* ‘to read a little bit from time to time’

In addition to the meanings listed above, the prefixes also have a ‘pure’ aspectual meaning, where it only adds perfective aspect to a verb (e.g., *slat* ‘to send’ (impf.) and *po-slat* ‘to send’ (pf.), *pomnit* ‘to remember’ (impf.) and *vs-pomnit* ‘to remember’ (pf.)).

We see that both prefixes have several well-defined meanings. However, my intuition is that there are more words where the meaning of the prefix is not clear for *voz-* than for *po-*. This intuition is confirmed in the next section where I analyze the numeric characteristics of these two prefixes.

### 4.2.2 Productivity analysis

In this section I perform a numeric analysis of the productivity of the prefixes *po-* and *voz-*. I find that relative frequency, along with other factors, is a good predictor of prefix productivity.

In his discussion of quantifying productivity of morphological units, Baayen (1992) lists the criteria of a good productivity measure: it should provide productivity rankings that correspond to linguistic intuitions (*intuitiveness*), it should reflect how well the morphological particle combines with new words (*hapaxability*), words with idiosyncratic properties should lower the productivity value (*idisyncraticness*) and it should reflect the fact that productivity does not simply equal the number of types associated with that morphological unit (*going beyond types*). In addition, as Hay and Baayen (2002) argue, the number of decomposed forms, or forms whose base frequency is larger than derived frequency, associated with an affix affects its productivity as well: the higher the number of those forms, the more productive the affix (*decomposed forms*).

I compared the two prefixes on a few of these criteria. First, intuitively, *po-* is much more productive than *voz-*. There are many more words with *po-* (4278) than with *voz-* (1236). Next, there are many more new words used with *po-* than with *voz-*. This is a notable characteristic, since one of the most important indicators of productivity of an affix is how readily it enters into new formations. In order to show that *po-* is used with new words more, I selected 47 verbs that entered the Russian language in the past two decades, mostly computer terms from the English language, such as *fludit* ‘to flood’ and *frendit* ‘to friend’ (on Facebook, Livejournal, etc.). Then I entered those verbs plus the prefix *po-* and *voz-* into a Russian search engine, Yandex<sup>1</sup>, to see if any results appear. Since I was only interested in whether the prefixed neologisms exist in usage, I only needed to make sure that the words were not misspellings when there was a small number of returned results. The actual count of the occurrences was not important as long as it was above zero. The complete list of

---

<sup>1</sup><http://www.yandex.ru>

words used for this test is in Table A.1. Out of the 47 verbs used for the test, 46 words, or 98%, are also used with *po-*, as evidenced by results of a Yandex search. In contrast, *voz-* is only used with 6 verbs out of 47 (or 13%). This is further evidence that *po-* is productive, while *voz-* is not.

Next, I studied the *po-* and *voz-* prefixed words based on relative frequency of the derived and base words. Relative frequency of base and derived words is important not only for morphological processing, as Hay (2001) argues, but also for affix productivity, as is discussed in Hay and Baayen (2002). Hay and Baayen analyzed 80 English affixes and plotted log derived versus log base frequency for them. Several factors are important for those affixes: correlation between the two variables, the slope of the resultant line, and its intercept. They argue that a positive and significant correlation between two variables, a higher intercept and steeper slope of the resultant line are all characteristics of a more productive affix. A positive and significant correlation is important, since the more transparent the relationship between bases and corresponding derived words, the more predictable the relationship between frequencies should be. A higher intercept and steeper slope of the resultant line effect in more points being above the  $x=y$  line, meaning more words where the derived form is less frequent than the base, i.e. words that are more prone to morphological decomposition.

To test whether relative frequency of the prefixed words is reflective of the prefixes' productivity, I plotted the words with existing unprefixes bases using their derived and base frequency. The calculations were done as follows: all words starting with the relevant letter sequence (*po*, *voz*, *vos*, *vz* or *vs*) were selected from the Russian orthographic dictionary<sup>2</sup>, only prefixed words with those sequences were selected, and their frequencies were calculated using the main subcorpus of the Russian National Corpus<sup>3</sup>. Then the base frequency was calculated by stripping of the prefix and querying the RNC with the result. Overall, 70% (1944 out of 2755) of words used with *po-* are less frequent than their bases, and 54% (431 out of 787) of words used with *voz-* are less frequent than their bases.

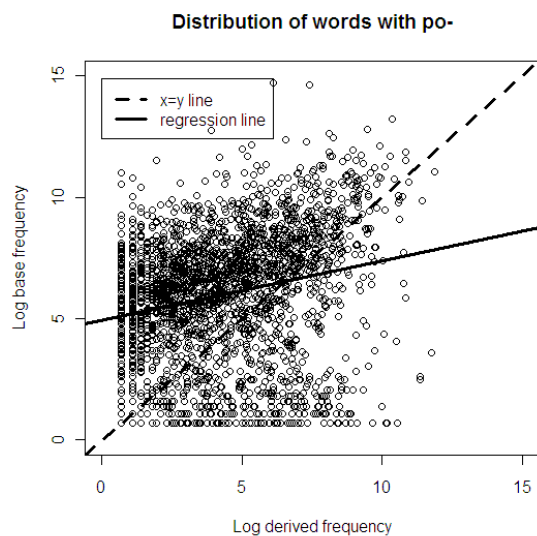
To determine the correlation, intercept and slope for *po-* and *voz-*, I plotted derived versus base frequency for all the *po-* and *voz-* prefixed words, excluding the words with zero base or zero derived frequency. The resulting plots are shown in Figure 4.1 and Figure 4.2. For *po-*, the correlation between log base and log derived frequency is 0.22 ( $p=0$ ). The intercept of the regression line is 4.95 and the slope is 0.24. Thus, there is a positive and significant correlation between log base and log derived frequency of words with *po-*, the resulting regression line has a high intercept and a positive slope.

On the other hand, for *voz-* there is a positive and significant, correlation for log base and log derived frequency, 0.07 ( $p=0.03$ ). Although the correlation is positive and significant, it is much lower than for *po-*. The intercept of the regression line is 4.66 and the slope is 0.07. Thus, the correlation, the intercept and the slope are all much lower than for *po-*.

These data show that the proportion of words less frequent than their bases used with a

<sup>2</sup>[http://ru.wikisource.org/wiki/Орфографический\\_словарь\\_русского\\_языка](http://ru.wikisource.org/wiki/Орфографический_словарь_русского_языка)

<sup>3</sup><http://www.ruscorpora.ru>

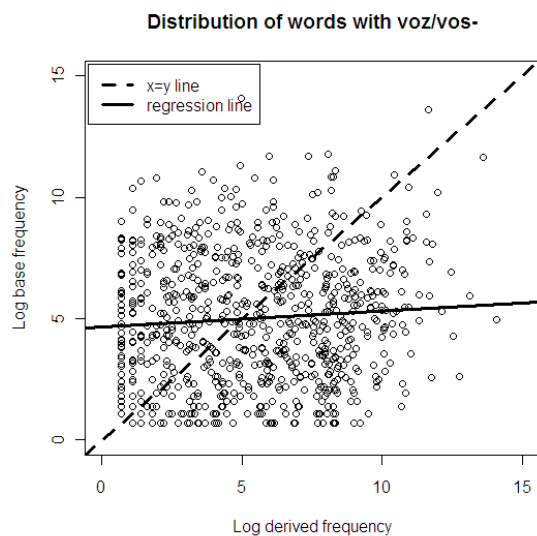
Figure 4.1: Plot of base versus derived frequency for *po-*

particular prefix is a good predictor of prefix productivity.

To summarize, *po-* and *voz-* differ by all relevant parameters. There are many more words used with *po-* than with *voz-*, the correlation for base and derived frequency for *po-* is positive and significant, while it is positive but insignificant for *voz-*, borrowings combine freely with *po-* and almost not at all with *voz-*. Overall, this confirms the intuition that *po-* is more productive than *voz-*. In addition, we see that all measures we selected for the productivity analysis are well-suited: there is a difference between the prefixes in the expected direction in the overall number of words used, in the ratio of words more frequent than their bases to words less frequent than their bases and in the number of neologisms used with that particular prefix. Thus, we can conclude that these measures, and in particular relative frequency of words and their bases, are reliable in informing us of prefix productivity. Next I report the results of a prefix separation experiment with verbs with the prefix *po-* that show a difference in processing words that are more frequent than their bases and words that are less frequent than their bases.

### 4.3 Po- experiment data analysis

Now that we have established that relative frequency is an important factor in Russian verbal morphology, I go on to analyze part of the reaction time data from the experiment in Chapter 3 that includes the free roots. For the details of the experiment, see Section 3.2.

Figure 4.2: Plot of base versus derived frequency for *voz-*

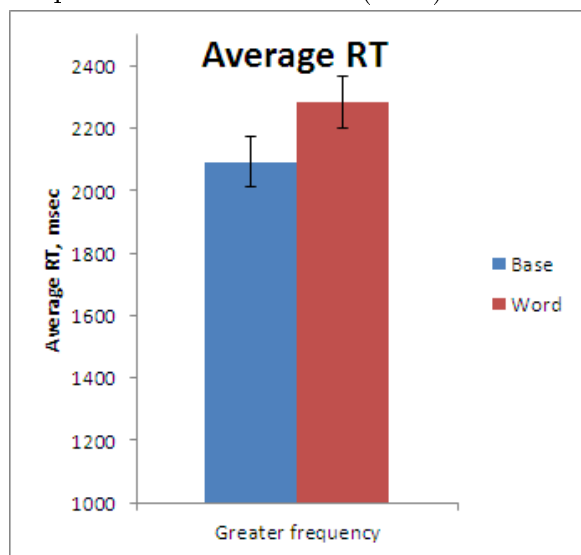
There are several predictions about this analysis.

Since this task required separating the prefix out of the word, the prediction is that the reaction time will be longer for those words that are generally *not* decomposed into constituent parts. Words that have a greater than derived base frequency are hypothesized to be decomposed, while words with smaller than derived base frequency are hypothesized to be processed as whole words. That means that the words whose base frequency is smaller than their derived frequency, are predicted to have longer reaction times than the words, whose base frequency is larger than their derived frequency. However, if, as Cole et al. (1989) argue, prefixes are never decomposed out of words containing them, there should be no difference in reaction times between these two groups of words. Thus, a difference in reaction times would demonstrate the validity of two hypotheses: that prefixes are separated out of some morphologically complex words and that relative frequency is an important factor in morphological processing.

### 4.3.1 Analysis

Four items were excluded. These items contained the reflexive suffix *-sja*, which made the morphological structure of those words more complex and thus harder to analyze. One item (*poumnet* 'to become smarter') was excluded because it was the only item whose base started with a vowel, and contained a V-V transition between the prefix and the root, an extremely unlikely within-morpheme transition. This item's average reaction time was 1546 ms, almost

Figure 4.3: Average RT for words that are more frequent than their bases (word) and less frequent than their bases (base).



500 ms less than the average of all the other items. This is according to expectations; an item containing an extremely unlikely within-morpheme phonotactic transition is expected to be decomposed easier than other items (Hay, 2002). Thus, this item was excluded. After this exclusion, there were 7 items less frequent than their bases and 8 items more frequent than their bases. Results of three subjects were excluded because of high error rates (more than 25%).

Two analyses were performed on *po-* prefixed data, a means analysis and a mixed regression analysis, in order to evaluate which other factors might have influenced the RT. In the mean analysis results were analyzed by item and by subject. The results are shown graphically in Figure 4.3. The difference was significant both by subject ( $p < 0.001$ ) and by item ( $p = 0.048$ ).

What we see from the mean analysis is that there is a difference between the two sets of words, significant both by subject and by item (the borderline  $p$ -value of the by-item analysis might be attributed to the small number of items). To investigate further, I carried out a multiple regression analysis with the logarithm of reaction time as the dependent variable and subject and item as random effects. I performed the analysis according to (Crawley, 2007), and retained all the factors whose  $p$ -values were under 0.1. While designing the model, I input factors that could have affected the response time, including information about frequency, family size, semantic transparency, trial number and phonological and orthographic information. Hay (2002) showed that phonological transitions can affect morphological decompositionality, thus word length (in letters and syllables) and the prefix-root



transition (VCC or VCV) were included as possible influencing factors in the model. Word frequency was included in the model, as it has been shown to be an important factor in morphological processing (Bybee, 2007). Family size has been shown to affect morphological processing of English words (e.g., (Baayen et al., 1997)), even monomorphemic ones, and thus was included as a possible influencing factor. Semantic transparency has also been shown to affect morphological processing (Wurm, 1997), and thus it was included in the model. Semantic transparency was calculated as follows: after inputting each word into the dictionary on <http://www.gramota.ru>, I counted the number of unprefixated words with the same root appear in the definition. This procedure is similar to the calculation of semantic transparency in (Hay, 2001), and the reasoning is that a word that is more semantically transparent should include more words with the same root in its definition than a word that is less semantically transparent. Finally, to test whether relative frequency of derived and base words is important in morphological processing, I included the difference between the logarithms of derived and base frequencies as a possible influencing factor.

The factors included in the final model were accuracy (accurate answers were faster), semantic transparency (semantically transparent items were faster), trial number (the later in the experiment the item was, the faster was the reaction time), unprefixated family size (a small inhibitory effect), and relative frequency (words that are more frequent than their bases were reacted to slower than words less frequent than their bases). Relative frequency and unprefixated family size are marginally significant ( $p = 0.8$ ), and that might be again attributed to a small number of items. The resulting model is shown in Table 4.1 (fixed effects) and Table 4.2 (random effects).

<b>Factor</b>	<b>Estimate (Log RT)</b>	<b><i>p</i>-value</b>
Intercept	7.6743	$p = 0$
Inaccurate answer	0.3441	$p = 0.0014$
Semantic transparency	-0.1114	$p = 0.0050$
Base-derived frequency difference	-0.0376	$p = 0.0788$
Trial number	-0.0020	$p = 0.0018$
Unprefixated family size	0.0024	$p = 0.0757$

Table 4.1: Fixed effects of the mixed effects regression model.

<b>Random effect</b>	<b>Variance</b>	<b>Standard deviation</b>
Subject number	0.0805	0.2838

Table 4.2: Random effects of the mixed effects regression model.

The  $R^2$  of this model is 0.39, compared to the of the model without fixed effects, which is 0.38. Although the increase in is relatively small, the variance for the adjustment by item

is reduced by 100% to 0 (and thus is taken out of the model), while the variance for the adjustment by subject is reduced by 6%. This means that the fixed effects model explains a little more variance than the model without the fixed effects, but now the same amount of variance is explained with fixed effects instead of the random effects of subject and item number.

### 4.3.2 Discussion

Overall, both the mean analysis and the mixed-effects regression models confirm that there are differences in processing between words which are more frequent than their bases and words which are less frequent than their bases. It is easier to separate the prefix from words that are less frequent than their bases. A better predictor is the difference between the frequency of the base and the derived word, where the larger the difference, the easier it is to separate the prefix. I will take this as evidence that the relative frequency effect is present. Thus, the prediction that there are relative frequency effects in Russian is borne out.

Another factor that turned out to be significant in all of the models was unprefixing family size, with a small inhibitory effect. We might hypothesize that that stems from a strategy by subjects to make a lexical decision on the unprefixing base: the word starts with *po-*, and if the unprefixing base is a word, there is a prefix in that word. Usually, a facilitatory family size effect is observed (e.g.), and Wurm (1997) cites U-shaped family size effects in lexical decision tasks, where large family size is to the advantage, and inhibits reaction time when it is larger than a certain threshold. The reasoning underlying this effect is as follows. In the early stages of lexical decision, a large family size is facilitatory, as it raises the probability that the string is a word, while in later stages, where the exact identification of the word is necessary, a large family size makes the probability of that particular word low, and thus is inhibitory. Here we see an inhibitory family size effect, and we might hypothesize that it is due to the fact that the unprefixing base is already very word-like, since it is a part of another word, and only the exact identification of the string is necessary, where a large family size is inhibitory. This is an interesting question for a future more thorough investigation.

The last important factor in this prefix separation experiment is semantic transparency. If the word was semantically transparent (or its unprefixing relative appeared in the dictionary on <http://www.gramota.ru>), it was easier to decompose the prefix out of it. This is in accordance with the results of the experiment from Chapter 3. A clear semantic connection makes lexical connections between words stronger, and the word parts easier to discern.

Overall, while it is not clear whether all the words predicted to be decomposed are indeed decomposed, the experimental results clearly show that *some* words are decomposed, and that is evidence for morphological decomposition of prefixes out of words, at least in some cases. Since in the experiment the participants were asked to answer ‘yes’ or ‘no’ to the question ‘Does this word contain the prefix *po*?’ , difference in reaction times suggests that in some words the prefix is separated out more easily than in others. This finding goes against

previous findings by Cole et al. (1989), where they found that suffixes are decomposed out of words, while prefixes are not. It is possible that previous prefix separation experiment results would be reinterpreted if relative frequency were to be taken into account.

These results agree with Cole et al.'s (1997) findings for French, Hay's (2001) and Losiewicz's (1992) findings for English, Zuraw's (2009) findings for Tagalog and Burani and Thornton's (2003) findings for Italian: relative frequency of base and derived word is an important factor in morphological processing. This experiment, using Russian and a different experimental paradigm, adds cross-linguistic evidence to these findings.

As Stemberger and MacWhinney (1988) argue, frequency effects are usually taken as evidence of storage. The result we see in this experiment is the larger the difference between the base and the derived frequency, the easier it is to separate the prefix. However, this frequency difference does not necessarily mean that there is a difference in storage. One possibility is that there are two representations of a word, a whole-word one and a decomposed one, and the one accessed is the more frequent one. Another possibility is that there is only one representation of a word, a whole-word one, and that the stronger the links to the unprefix base the easier the decomposition. The former option presupposes a separate representation for affixes, since, if a word is stored decomposed, its affixes must be detached from the rest of the word. In the latter option, on the other hand, there is no separate storage of the affix.

Apart from the question of how the words are stored, we can now consider how they are processed. As described in the beginning of the chapter, there are several dual route processing models. Caramazza et al. (1988) suggest a model where the decompositional route is accessed only by novel words, while the whole-word route by known words. However, even if the word is accessed via the decompositional route, its morphological representation is activated. For example, *walked* is accessed via the whole-word route, and it activated the representations of its morphemes, *walk*<sub>V</sub>- and *-ed*. In the light of the current results, the model would need to be modified to take into account relative frequency effects for known words. The race model of Frauenfelder and Schreuder (1992), where the two routes race, would need to be modified, where the likelihood of activation of the decompositional route for morphologically complex words would depend on the relative frequency of the word and its base. Finally, Wurm's (1997) model, where there is an obligatory whole-word route and a decompositional route that is selective about which words it considers, would also need to be modified, and the decompositional route might be accessed only when the relative frequency difference is large enough.

Any model that takes into account relative frequency results would also need to separate the issues of storage and processing, which are independent. A dual-route access model does not need to presuppose decomposed storage. If a prefixed word is decomposed during processing, it could mean that both the decomposed representation and the whole-word representation are accessed (two representations per word), or that the word and its base are accessed (one representation per word). These issues are explored in more detail in the next chapter.

## 4.4 Conclusion

To summarize, in this chapter I presented the results of an experiment that confirm the existence of relative frequency effects in Russian, where the prefix is separated more easily from words that are more frequent than their bases. The larger the difference between the base and the derived frequency, the easier it is to separate the prefix. While these results are consistent both with two representations for a word, one decomposed and one as a whole word, and with one whole word representation with links to the base, they call for a morphological processing model where relative frequency is taken into account, and a decompositional route is more or less likely depending on the relative frequency of the word and its base. The question of storage of stems and affixes is taken up in the next chapter.

## Chapter 5

# Differences between affixes

### 5.1 Introduction

In the last two chapters we saw that roots form a hierarchy in Russian and that relative frequency of the word and its base is an important factor that contributes to processing. These findings describe whole word and roots; in this chapter I compare how different Russian prefixes are processed. In the next chapter I investigate the status of affixes (prefixes in particular) in the network theory of morphology.

Two basic questions are addressed in this chapter. First, are all affixes processed similarly or are there differences in processing between affixes - and if so, what do these differences reflect? Based on my investigation of these issues, I further pose the question of whether affixes are units and how they might be represented in the network theory. A complexity rating experiment where words with five different Russian prefixes were used indeed demonstrates differences between processing of differently prefixed words, where semantics of those prefixes is the important factor influencing the results.

### 5.2 Previous work

Hay and Baayen (2002) performed a very extensive analysis of 80 English affixes. Their analysis included calculating the affixes' type and token parsing ratios (i.e., the ratio of type or token frequency of words whose base frequency is larger than derived frequency and the type or token frequency of all the words with that affix), the intercept and slope of the best fit lines when plotting base versus derived frequency for all the words containing the affix. They found that affixes differ greatly on these parameters, and that these parameters are important for affix productivity. Thus, they found that type and token parsing ratios positively and significantly correlate with the productivity estimate  $P$ , which is the ratio of the total number of hapaxes (words appearing only once in a corpus) with that affix and the total token frequency of that affix. Therefore, an affix's productivity is directly related

to the number of words with that affix that are more likely to be parsed. The larger the number of these words, the more productive it will be.

Hay and Baayen thus demonstrate that affixes are not the same, and that type and token parsing ratios of these affixes directly affect their productivity. In this chapter I intend to investigate other, semantic properties of five Russian prefixes, and show that they, too, affect processing of prefixed words.

### 5.3 Russian prefix characteristics

In this section I consider five Russian prefixes, *po-*, *pod-*, *za-*, *vy-* and *ot-* and their characteristics. I estimated the type and token parsing ratios of these prefixes, and compared those numbers to the type and token parsing ratios of all other Russian prefixes. The estimation procedure was as follows. For a certain prefix, I selected all the words that start with that letter combination from the Russian orthographic dictionary, containing 161,734 words. Thus, for example, when estimating the type and token parsing ratios for *po-*, I selected from the dictionary all the words that start with *po*. After that I selected all the words that contained the selected prefix, as opposed to the words with a pseudo-prefix, and calculated the difference between word frequency and base frequency (frequency of the word with the prefix removed) for each word. I calculated the type parsing ratio by dividing the natural logarithm of the count of words that are less frequent than their bases over the natural logarithm of the count of all the words with that prefix. I calculated the token parsing ratio in a similar fashion, where I used the natural logarithm of the cumulative frequency of the words instead of their counts.

The productivity of the five prefixes used in this chapter is demonstrated by their type and token parsing ratios in Table 5.1. As we see, the type parsing ratio is very similar for all five prefixes, with *po-* and *ot-* having somewhat higher ratios. Compared to other Russian prefixes, whose type parsing ratios range from 0.81 to 0.94, the five prefixes selected are relatively more productive and close in productivity level, based on the type parsing ratio assessment. The token parsing ratios are a bit more spread out, with *za-* and *po-* having the highest ratios. In addition, the token parsing ratios of these five prefixes are spread out as compared to all Russian prefixes, whose token parsing ratios range from 0.84 to 0.95. The type parsing ratio is probably more accurate, as the token frequency count is not as precise as the type frequency count. Thus, I assume that these five prefixes have similar productivity levels.

Let us now look at the meaning of these five prefixes. One of the ways to classify meaning of the prefixes is described in (Townsend, 1975). According to his description, which is a standard view of Russian prefixes, it is possible to divide prefix meaning contributions into three types: aspectual, lexical and sublexical (Townsend, 1975, p.118). Aspectual meaning is added when the prefix is purely perfectivizing, e.g., *pisat'* 'to write' (impf.) and *napisat'* 'to write' (pf.). Lexical meaning is added when the prefix introduces a 'new lexical element'

Prefix	Type parsing ratio	Token parsing ratio	Number of types	Number of tokens
<i>po-</i>	0.93	0.92	3951	3971232
<i>pod-</i>	0.92	0.87	1639	1169171
<i>za-</i>	0.94	0.92	861	695110
<i>vy-</i>	0.91	0.88	2255	811634
<i>ot-</i>	0.92	0.89	2311	1456870

Table 5.1: Numeric characteristics of 5 Russian prefixes.

to the meaning of the verb, in addition to making that verb perfective, e.g., *nesti* ‘to carry’ (impf.) and *prinesti* ‘to bring, to carry to’ (pf.). Finally, sublexical meaning is added when the prefix ‘modifies the verbal action in some way, usually with respect to time or intensity’, e.g., *pisat’* ‘to write’ (impf.) and *popisat’* ‘to write for a while’ (pf.). There are 21 prefixes in Russian, and each prefix has a different combination of the aspectual, lexical and sublexical meanings. For example, the prefix *v/vo-* only has a lexical meanings (‘in’, ‘into), the prefix *po-* only has sublexical meanings (‘do for a while’, ‘begin to’ and others), and the prefix *pro-* has both lexical and sublexical meanings (‘through’, ‘do for a specific length of time’, and others).

While the meaning classification above is standard in Russian linguistics, it could be more useful to see how many words with a certain prefix contain one of the prefix’s systematic meanings. This is a criterion in addition to the three aspects of semantic transparency I identified in Section 1.4.2: the number of meanings, how frequently they occur, and their concreteness. The smaller the number of possible meanings, the more words occur with those meanings, and the more concrete those meanings, the more semantically transparent a unit is. Thus, prefixes that have less meanings are more semantically transparent than prefixes that have more meanings. In addition to that, prefixes that have more spatial meanings are more semantically transparent than prefixes that have fewer spatial meanings, because spatial meanings are the most semantically transparent of all. Spatial meaning arises when the prefix is purely spatial in the word, for example, *vybežat’* ‘to run out’, cf. *bežat’* ‘to run’. Finally, we need to distinguish metaphoric meaning. Metaphoric meaning is added when the use of the prefix is a metaphor on the spatial meaning, for example *vybaltyvat’* ‘to blab out’, cf. *boltat’* ‘to chatter’. While more abstract concepts can be understood in terms of more concrete ones in metaphor, such as time understood through space (Boroditsky, 2000), metaphoric meaning can be processed through its base domain or acquire a meaning on its own (Bowdle and Gentner, 2005; Gentner and Bowdle, 2008), and thus I do not use it in measuring semantic transparency.

In order to evaluate the systematic meaning contribution of the five prefixes I am con-

sidering, I chose a random sample of size 100 of prefixed verbs with each prefix, using the `sample()` function in R. After that, I assigned each verb a meaning that is listed in Townsend (1975, pp.124-133), slightly modified. The list of meanings was as follows:

- **po-**
  - Begin to; *pojti* ‘to start going’.
  - Do for a short time; *posidet’* ‘to sit for a while’.
  - Do somewhat, to some extent; *polečit’* ‘to cure a little bit’.
  - Do from time to time; *pokurivat’* ‘to smoke from time to time’.
  - Distributive, do to a multitude of objects or obliques; *porasprodavat’* ‘to sell multiple objects to multiple recipients’.
- **za-**
  - Alter course, with verbs of motion; *zajti* ‘to drop in (on the way)’.
  - Fix or make permanent by some action; *zakrepiť* ‘to fasten’.
  - Acquisition; *zarabotat’* ‘to earn’.
  - Close, block, fill; *zadelat’* ‘to stop up, close off’.
  - Subject to extreme or excessive action; *zadarit’* ‘to load or overload with gifts’.
  - Begin to; *zaplakat’* ‘to start crying’.
- **ot-**
  - Movement away, off, from, separation; *otojti* ‘to step away’.
  - Movement/figurative movement back; *otdat’* ‘to give back’.
  - Finish; *otslužit’* ‘to serve out time’.
- **pod-**
  - Movement closer to something; *podojti* ‘to approach’.
  - Under, from under; *podnjat’* ‘to lift’.
  - Add, supplement; *podrabotat’* ‘to earn extra’.
  - A little, not completely; *podsoxnut’* ‘to dry a little, not completely’.
- **vy-**
  - Movement out; *vyjti* ‘to go out, exit’.
  - Do or finish successfully; *vyučit’* ‘to learn by heart’.



– Finish; *vykurit* ‘to finish smoking’.

As we see from the above lists, the prefixes that have the most meanings are *po-* and *za-*. *Vy-* and *ot-* have the least number of meanings. *Pod-* is somewhere in the middle.

In addition to the meanings listed, each prefix also had an aspectual meaning (changing an imperfective verb to a perfective, such as the prefix *na-* in the pair *pisat* ‘to write (impf.)’ and *napisat* ‘to write (pf.)’), and an ‘other’ meaning that could not fit in any other category. I then calculated which proportion of the verbs in the sample had one of the systematic meanings above, and called this proportion the transparency ratio. Thus, the ratio was the number of verbs with the meanings above divided by 100, the total number of words in the sample. Verbs where the prefix had an aspectual meaning, or a meaning falling into the “other” class, were not included in the calculation. Aspectual meaning was not included since it is highly abstract, and in many cases it is hard to say whether it is systematic or not. The transparency ratios are presented in Table 5.2. These numbers are a crude estimate of semantic transparency of these ratings, and need to be made more consistent by having more people rate the meanings in future studies.

Prefix	Transparency ratio	Number of spatial meanings
<i>za-</i>	0.65	3
<i>po-</i>	0.66	0
<i>pod-</i>	0.81	14
<i>ot-</i>	0.85	37
<i>vy-</i>	0.85	38

Table 5.2: Semantic characteristics of 5 Russian prefixes.

The table lists the transparency ratios, as well as the number of verbs with a spatial meaning in the sample. We see that these two criteria are closely aligned: prefixes that have a lower transparency ratio also have a smaller number of verbs with spatial meaning. The pattern that emerges is that *za-* and *po-* are the least semantically transparent prefixes, *vy-* and *ot-* are the most semantically transparent prefixes, and *pod-* is somewhere in between.

We can calculate an additional score that takes into account the number of meanings and the number of words per each meaning. This is useful, since, as I mention above, prefixes with fewer meanings are more semantically transparent, and prefixes with a large number of words associated with their meanings are also more semantically transparent. I calculate this transparency score for a sample as follows: for each meaning, divide the number of words with that meaning by the number of meanings, and sum the ratios that result. For example, *vy-* has three meanings, and there are 40, 35 and 10 words associated with those meanings. Thus, the transparency score for *vy-* is  $\frac{40}{3} + \frac{35}{3} + \frac{10}{3} = 28.3$ . It is easy to see that the more meanings the prefix has and the fewer words are associated with each particular meaning,

the lower the score. The transparency scores for the five prefixes discussed in this chapter are presented in Table 5.3.

<b>Prefix</b>	<b>Transparency score</b>
<i>za-</i>	10.83
<i>po-</i>	13.2
<i>pod-</i>	20.25
<i>ot-</i>	28.3
<i>vy-</i>	28.3

Table 5.3: Transparency scores of five Russian prefixes.

The transparency score, like the transparency ratio and the number of spatial meanings above, places *po-* and *za-* as the least semantically transparent prefixes, *ot-* and *vy-* as the most semantically transparent ones, and *pod-* as being in between. What kind of implications does this semantic continuum have for the words where these prefixes appear? In the prefix separation experiment in Chapters 3 and 4 semantic transparency had a facilitative effect on reaction time, so it was easier to separate a prefix in a word that was more semantically transparent than in a word that was less semantically transparent. I hypothesize that the same effect will appear across prefixes, that more semantically transparent prefixes will be easier to discern. To test this prediction, I performed an experiment that asked subjects to rate the structural complexity of words with the above prefixes.

## 5.4 Complexity ratings experiment

In this online experiment I asked subjects to rate the structural complexity of different words on a 1 (simple) to 10 (complex) scale. The results show differences in processing words with different prefixes, and confirm the results of the prefix separation experiment with respect to different types of roots.

### 5.4.1 Subjects

34 native Russian speakers between the ages of 26 and 52 from the USA recruited on several social networks participated in the experiment. The subjects did not receive compensation for their participation.

### 5.4.2 Stimuli

The stimuli included 230 words with the five prefixes described above; 45 items with *po-*, 45 items with *pod-*, 50 items with *za-*, 41 items with *vy-* and 49 items with *ot-*. In addition,

112 unprefixed words were included in the experiment. Words with each prefix included words with bound, modified and free roots. The number of different root types used with each prefix are shown in Table 5.4. The stimuli were chosen from lists of prefixed verbs compiled from the Russian orthographic dictionary referenced above to include items with all root types, with only one prefix and no reflexive suffix *-sja*.

Type of root	<i>po-</i>	<i>pod-</i>	<i>za-</i>	<i>vy-</i>	<i>ot-</i>
bound	10	9	10	6	9
modified	14	11	12	13	13
free	21	25	28	22	27
total	45	45	50	41	49

Table 5.4: Number of different root types used in the experiment.

### 5.4.3 Procedure

This study was presented to subjects online. Both the instructions and the study were in Russian. The instructions described what is meant by ‘structural complexity’ on two examples, a prefixed noun (*prigorod* ‘suburb’) and a monomorphemic noun (*sneg* ‘snow’). The parts of the two nouns were shown (*pri-* and *gorod* for *prigorod* and *sneg* for *sneg*) and it was stated that the prefixed noun was more structurally complex than the monomorphemic noun. Then the subjects were told that they were to answer questions about how structurally complex different Russian words are, using a scale of 1 (simple) to 10 (complex). They were urged to use the whole scale rather than one or two numbers. Then words were presented to subjects one by one with the question ‘How complex in structure is this word?’ and the rating scale.

### 5.4.4 Results

The results were analyzed using multiple regression, where the inclusion of each fixed effect was tested by an ANOVA comparison between a model with that fixed effect and without it. In what follows I discuss which factors were considered in the model, and later on report which ones turned out to be significant. This was a visual test, and the subjects were asked to estimate the structural complexity of the word by visual inspection, and as quickly as possible. Thus, it is reasonable to include word length as a possible influencing factor, since the longer the word looks, the more complex it seems. As word frequency is a standard influencing factor in psycholinguistic studies (Bybee, 2007, pp.5-22), it was also considered. Since phonotactic factors are important in word segmentation (Hay, 2002), whether the stress shifted to the prefix or stayed on the base and whether the prefix was consonant-final or vowel-final were considered. Estimates of how connected a word is to

other words with the same morphological constituents were included: the number of words with the same root in the definition of that word (on <http://www.gramota.ru>), the number of unprefixated words with the same root in the definition of that word, and the number of words with the same prefix in the definition of that word. These were included as measured of semantic transparency of the word, similarly to (Hay, 2001). Since the type of root that appears in the word was a significant factor in the previous experiment (Chapter 3), root type was also included as a possible factor. Since the difference between modified roots and roots bound within part of speech was not significant in the mean analysis in Chapter 3, I used three root types in this experiment, collapsing modified roots and roots bound within POS into one category. The following prefix characteristics were added as possible factors: the type parsing ratio, the token parsing ratio and the transparency ratio. Since some of the verbs contained the imperfectivizing suffix *-iva-*, and some did not, I included the presence of this suffix as a possible influencing factor in the model. The relative frequency factor that turned out to be important in the prefix separation experiment, described in Chapter 4, was difficult to calculate for all the words in this case, since not all words had free roots. However, I did include a prefix relative frequency measure, which was the difference between the natural logarithm of the type frequency of words more frequent than their bases and the natural logarithm of the type frequency of words less frequent than their bases.

The rating given by study participants was transformed to the power of 0.2 in order to make it conform to the normal distribution, and centered around zero. The histogram of the untransformed rating is shown in Figure 5.1 and the histogram of the transformed rating is shown in Figure 5.2. The skewness and kurtosis of the transformed vector was 0.02 and -0.16, respectively, within the accepted ranges for normality, -0.8 to 0.8 for skewness and -3 to 3 for kurtosis. The skewness and kurtosis of the untransformed vector were 0.94 and 0.67, respectively. The model included the transformed and centered rating as the dependent variable and subject number and item number as random effect predictors. 187 outlying data points were removed from the data. Level contrasts as described in (Crawley, 2007) were used for the root type predictor in order to compare the different root types against each other. Significant fixed effects were: word length, presence of the imperfectivizing suffix *-iva-*, number of words (types) with the same root in the definition of the word on <http://www.gramota.ru>, modified roots, free roots, transparency ratio (these predictors were associated with a higher numeric rating); bound roots, word frequency (log transformed) (these predictors were associated with a lower numeric rating). The prefix relative frequency measure was marginally significant (ANOVA test of the model with and without it returned a p-value of 0.65), and thus it is not included in the present model. Fixed effects are summarized in Table 5.5, and random effects are summarized in Table 5.6. The plot of the resulting residuals of the model is in Figure 5.3. Since the transparency ratio only has 5 values, one for each prefix, the model might be overfitting. However, currently cross-validation to check for this on mixed-effects regression models is not possible, and would need to be done once the function necessary to do this becomes available.

The  $R^2$  of the resulting model is 0.783, as compared to the  $R^2$  of the null model, 0.786.

Figure 5.1: Histogram of rating untransformed

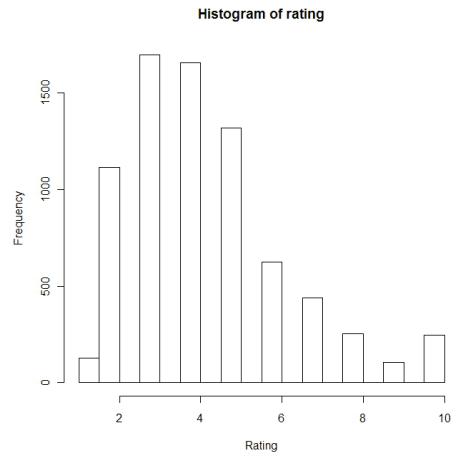
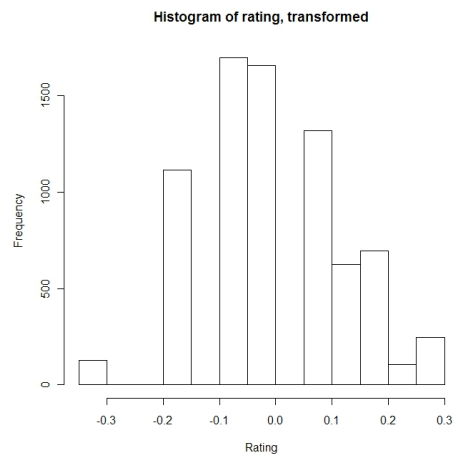


Figure 5.2: Histogram of rating transformed



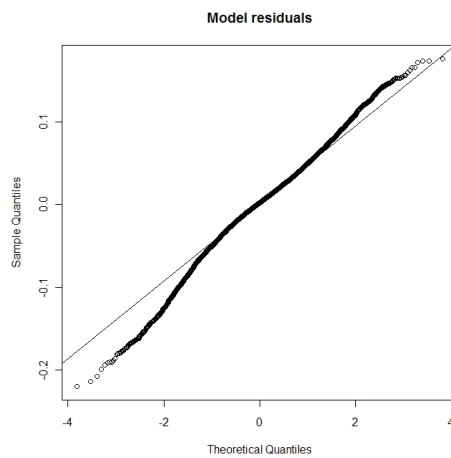
<b>Factor</b>	<b>Estimate (transformed centered rating)</b>	<b><i>p</i>-value</b>
Intercept	-0.2527	$p = 0$
Word length	0.0246	$p = 0$
Free v. modified root	0.0032	$p = 0.0284$
Presence of <i>-iva-</i> suffix	0.0200	$p = 0$
Same root word in def	0.0019	$p = 0.0193$
Transparency ratio	0.0555	$p = 0$
Bound v. modified root	-0.0049	$p = 0.0105$
Log word frequency	-0.0010	$p = 0.0004$

Table 5.5: Fixed effects of the mixed effects regression model.

<b>Random effect</b>	<b>Variance</b>	<b>Standard deviation</b>
Item number	0.00013266	0.011518
Subject number	0.00874602	0.093520
Residual	0.00301229	0.054884

Table 5.6: Random effects of the mixed effects regression model.

Figure 5.3: Model residuals



Although there is a slight decrease in the  $R^2$ , there is a decrease in variance of the item number random effect (the subject number random effect variance stays the same). The variance for item number goes down by 94%, from 0.0022 in the null model to 0.0001 in the model with fixed effects. That means that, while the fixed effects model explains approximately the same variance as the null model, the variance is explained partly by the fixed effects, and not only the random effects.

### 5.4.5 Discussion

In general, the model confirms the results of previous chapters, and factors that emerge as significant are form and semantic factors. Below I consider the contributing factors in detail.

The characteristics of the word are important: word length, presence or absence of the *-iva-* suffix, word frequency. The longer the word, the more complex it is judged to be, an expected effect; even a monomorphemic word that is longer than average can be perceived as containing more than one part. If the *-iva-* suffix was present, the word was judged as more complex than if it was not. This is also an expected result: stimuli had either 3 or 4 morphemes, and those that had the *-iva-* suffix always had 4<sup>1</sup>. Thus, as expected, the more morphemes were in a word, the more structurally complex it was rated. Word frequency also played a role: the more frequent a word, the less complex it was rated. This is also an expected result: more frequent words have weaker connections to related words and their morphological structure is not as evident (Bybee, 2007, p.13).

Root type was also a significant predictor, as in the experiment in Chapter 3, which is additional evidence for differences between root types. Words with bound roots were rated less complex, and words with free roots were rated more complex than words with modified roots. Thus, morphological structure was the least evident when the word had a bound root, more evident when the word had a modified root, and the most evident when the word had a free root. This result parallels the results of the reaction time experiment, where words with bound roots had the longest reaction times, words with modified roots had shorter reaction times, and words with free roots had the shortest reaction times. Therefore, the root hierarchy has been confirmed in two experimental paradigms, a prefix separation experiment and a complexity rating experiment.

Characteristics of the word's place in a larger system, as well as characteristics of the word itself, were relevant. The more words with the same root appeared in the word's definition, the more complex the word was rated, no matter whether these words were prefixed or not. This is in accordance with Hay's (2001) findings of more decomposable words having their bases mentioned in the definition. We can view this characteristic as a rough estimate of how well the meaning of the root is preserved across different words that contain it. If there

---

<sup>1</sup>I tried to input the number of morphemes as a factor, but the presence of the *-iva-* suffix was a better predictor, most likely because it is usually 3 letters long, as opposed to other morphemes that might be shorter.

are many different words with the same root that are used to explain the meaning of another word with the same root, then the meaning of the root is fairly well preserved across words, and the structure of such a word is viewed as more complex.

A final factor that was important in the model was the prefix transparency ratio defined in Section 5.3. The higher the transparency ratio of the prefix, the more complex the word was rated. Thus, as predicted, there was a prefix transparency continuum: the more semantically transparent the prefix, the more complex the word containing it was rated. Both the semantic transparency of the root (as roughly estimated by the number of words with same root used in the definition of the word), and the semantic transparency of the prefix are important factors in determining the structure of a Russian word.

To summarize, factors that contributed to the differences in complexity ratings were characteristics of the word: length, presence of the suffix *-iva-*, frequency, root type; semantic transparency of the root and semantic transparency of the prefix. What does the significance of these factors tell us about the status of prefixes? The only prefix characteristic that emerged as significant in the model is the semantic transparency of the prefix, since words with the prefixes that had a higher transparency ratio were rated as more complex than words with the prefixes that had a lower transparency ratio. The prefix semantic transparency factor cannot give us a definitive answer on the status of prefixes, but can shed some light on the issue.

In order to see how the semantic transparency factor might contribute to the status of prefixes, it is instructive to look at words with bound roots. In the words with bound roots that have the more semantically transparent prefixes the spatial meaning might be more easily discernible than if the prefixes are more semantically opaque. For example, in the words *otnimat'* 'to take away', *podnimat'* 'to pick up' and *vynimat'* 'to take out' the meaning of the prefix is clear, while in the words *ponimat'* 'to understand' and *zanimat'* 'to borrow' it is not. Thus, while the word *\*nimat'* does not exist, the root is completely bound and its synchronic meaning is not available in any of the above words, the meaning of the prefix is more available in some words, and less available in others, depending on the prefix. Therefore, we can hypothesize that prefixes with spatial meanings retain them even in words with bound roots, while prefixes without spatial meanings lose their separate meaning in such words.

While spatial meanings of some prefixes are clearly evident in words with bound roots, other prefixes demonstrate another kind of meaning 'availability'. As was demonstrated in Section 4.2.2, prefix *po-* can attach to many neologisms. The meaning used in new words is 'to do for a while', for example, *frilansit'* 'to freelance' and *pofrilansit'* 'to freelance for a while', or *sjorfit'* 'to surf' and *posjorfit'* 'to surf for a while'. This meaning can be applied to virtually any verb, hence the large number of neologisms that can be used with *po-*. Thus, as in words with bound roots above, the meaning contribution of the prefix is clear.

I have thus identified two criteria of prefix meaning contribution: availability of meaning in a word where the meaning of the root is not clear, and ability to form new word with neologisms. While prefixes with clear spatial meanings might have their meanings readily



available in words with bound roots, prefixes with many sublexical meanings are more likely to attach to neologisms, as spatial meanings easily attach only to a handful of motion verbs, and sublexical meanings modify the Aktionsart of the verb, and this modification can apply to a larger subset of verbs. While the same sort of argument could be made about aspectual meaning, all Russian prefixes can change the meaning of an imperfective verb to a perfective one, and thus this same reasoning does not apply.

This meaning availability, both for existing and new verbs, suggests that Russian verbal prefixes are form-meaning pairings, or units on their own. If they are units on their own, they also must include some reference to other units, as they are conceptually dependent (Langacker, 2002). Thus, we might imagine that the prefix *po-* and its meaning ‘do for a while’ might be represented as a general template over all verbs, [po-V](pf.):[V for a while](impf.), where the form ‘perfective verb with a prefix *po-*’ is paired with the meaning ‘do for a while’. However, it is not clear what kind of generalization is formed with prefix form and meaning. Langacker (2002) suggests that generalizations are necessarily made and stored, while Bochner (1993) proposes that these generalizations are temporary pattern matches, created and discarded after use<sup>2</sup>. If we view affixes as kinds of categories, then both are plausible options. As Barsalou (1983) shows, categories can be both established and ad-hoc, and both have similar structure. Future studies with more prefixes and variables, such as type and token frequency, should be performed, to investigate this question.

### 5.4.6 Methodology differences

The results of this experiment were interesting as to how they compare to the results of the prefix separation experiment described in Chapters 3 and 4. Some results are consistent across the two experiments: root type was a significant factor in both models. On the other hand, relative frequency of base and derived words was significant in the prefix separation experiment, and was not in the present structural complexity rating experiment, although the prefix measure was marginally significant. There are two possible explanations for this difference: one is that the data is too noisy for the factor to emerge, and the other is that there is a difference as to which properties the two methodologies tap into. Future experiments with these and other methodologies would shed light on this issue.

## 5.5 Conclusion

In this chapter I looked at Russian verbal prefixes in order to establish whether there are differences in processing between them. A complexity rating study was performed, where participants rated how complex the morphological structure of Russian verbs is, where the verbs had one of the five prefixes, *po-*, *pod-*, *za-*, *vy-* and *ot-*. The transparency ratio was a

---

<sup>2</sup>Bochner models prefixation as rules, here I do not make a specific choice on whether prefixes would be rules or units.

significant predictor of the complexity rating; the higher the transparency ratio, the more complex was the rating. I suggest that this difference could be reflective of the unit status of prefixes. I also identified two criteria of meaning availability of prefixes: more concrete prefixes have their meaning readily available in established words, while prefixes with more abstract meanings enter into formations with new words more easily. This availability of meaning of different prefixes in different words suggests that generalizations are formed using prefix form and meaning; however, the type of generalization, permanent or temporary, is not obvious.

## Chapter 6

# The Network model of morphology and its implications

In previous chapters I presented experimental evidence for a root hierarchy in Russian, relative frequency effects and processing differences between Russian prefixes. In this chapter I evaluate the hypotheses set forth in Chapter 2, present requirements of a morphological model, evaluate the Network theory of morphology relative to those requirements and different language families, and point to directions in future research.

### 6.1 Hypothesis evaluation

The following experimental results became evident in the previous chapters:

- A root hierarchy was indeed found, with reaction time differences as predicted. Reaction times in a prefix separation experiment were slowest for words with completely bound roots, faster for words with roots bound within POS, and for modified roots, and fastest for words with free roots. In addition, root type was a significant predictor in mixed-effects regression modeling, both for the prefix separation and complexity rating experiments.
- Relative frequency effects were discovered in the same prefix separation experiment, where there was a significant difference in reaction time between words that are more frequent than their bases and words that are less frequent than their bases, with faster reaction times for the latter. In addition, in mixed-effects regression modeling, the difference between base and derived frequencies was a significant predictor.
- In a complexity rating experiment, there emerged a difference between more and less semantically transparent prefixes. The transparency ratio of the Russian prefixes, which measured the availability of prefix meanings in a verb sample, was a significant predictor of the word complexity ratings. Words with prefixes with the higher transparency

ratio were rated more complex than words with prefixes with the lower transparency ratio.

- Additional findings. Other significant predictors that emerged in regression models of the experiments' results were semantic transparency (both the prefix separation and the complexity rating experiments) and family size (prefix separation experiment only).

These findings now need to be fitted into the network theory of morphology. I first consider restrictions that need to be set on any morphology theory.

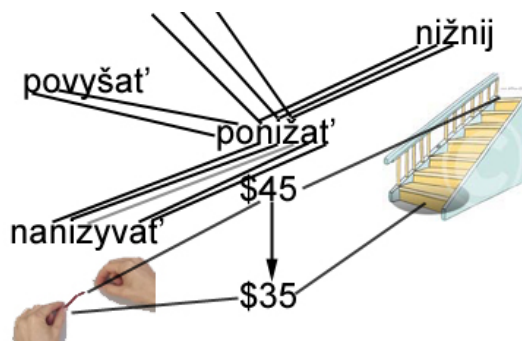
## 6.2 Morphological model restrictions

Any morphological theory is a model of morphological storage and processing in some sense. Each morphological theory has a main goal of representing the morphological facts of a language in some way. However, the best morphological theory will do that in a psychologically plausible way. In the end, we want to describe how people use language, and not just the abstract facts of language. Thus, although how words are stored should be explained by any morphology theory, the explanatory power of such a theory should not be restricted to just that. Processing and usage of words, while not usually included in morphology theories, has a direct bearing on word storage. Therefore, any psychologically plausible theory of morphology should take processing and usage of words into consideration.

It is a complex undertaking to evaluate the appropriate bearing of word processing and word usage on morphological theory. Words can be processed and used in two different modalities: auditory and visual. Studies have shown (Wurm, 1997; Marslen-Wilson et al., 1994) that auditory and visual processing might have some important differences. For example, in visual masked priming tasks, free and bound roots produce results that are the same (Pastizzo and Feldman, 2004; Forster and Azuma, 2000), i.e. both words with free and bound roots produce significant priming, while in cross-modal priming tasks (Marslen-Wilson et al., 1994), where the prime is presented audially and the target visually, words with bound roots do not prime other words with the same root, while words with free roots do. A controlled study with differences in modality taken into account would need to confirm that these different results are indeed caused by the contrast between presentation in the same modality and presentation in the other modality. It would also be interesting to see if the experimental results presented in this dissertation would be the same or different dependent on modality. Other studies (e.g. Rastle et al. (2004)) have also shown priming between pairs of words that are not related morphologically, but have the same letter sequences (such as *fete* and *fetish*). Thus, the visual priming results, where words with bound roots display priming just as words with free roots, might be just effects of orthography.

In addition, the following restrictions need to be taken into account (repeated from Chapter 2).

Figure 6.1: Form and semantic connections between words.



1. The intuition, shared by speakers, that words can be divided into parts, where each part has its own meaning. This intuition has led to the well-known morphology theories, Item and Arrangement and Item and Process. We also need to avoid the recognized pitfalls of IA and IP theories, such as empty, cumulative and portmanteau morphs (Anderson, 1992, pp.51-56), basic versus derived forms (see chapter 1 of this dissertation), and others.
2. Paradigms, often described as units in traditional grammars, should be included as units in a theory of morphology, as the items in paradigms are confirmed by various evidence to be more closely connected than other items (Bybee, 1985, pp.49-79).

In the following section I consider how these restrictions and experimental results from previous chapters can be fit into the network theory of morphology.

### 6.3 Revised network model of morphology

Since the revisions I am going to propose here are all based on experimental evidence, they are all psychologically plausible to a certain degree. Of course, this is contingent on future replications of these experiments, especially in other paradigms, such as lexical decision and priming experiments. Additionally, all the experiments in this dissertation were performed in the visual modality, and future research should include cross-modal or purely auditory experiments.

The basic architecture of the Network model, as is described in the previous chapters, includes form and meaning connections between words. The strength of those connections depends on different factors, including word frequency, and as is discussed in Chapter 4 and described below, relative frequency of words and their bases. This is illustrated in Figure 3.2, repeated here as Figure 6.1.

Addressing the restrictions above, we first have to specify how frequency information is collected in such a model. The logarithmic transformation of frequency has been shown to be an important factor in morphological processing (for an overview, see (Bybee, 2007, pp.5-22)), and thus it has to be collected in some form. Each word has to have its logarithmic frequency information available. Or, what is more likely, the word's availability for access is dependent on the logarithm of frequency. Since context plays a role in word activation (e.g., Zwisterlood (1989)), there will be some interaction between word frequency and context.

Moving on to morphology model restrictions, the first point above is addressed by the network morphology theory in its unmodified form: since it is word-based, it avoids all the problems inherent in IA and IP theories. Rather, morpheme effects are to be attributed to differing strengths of connections. Thus, when there are many form and meaning connections between words that span the same letter/phoneme sequence, that letter/phoneme sequence becomes prominent to the language user.

Paradigms, mentioned in the second point, should be included in the theory. Paradigms fall more into the realm of inflectional morphology, and are not addressed here in much detail, although derivational paradigms might also exist and would need to be considered in future studies. However, I offer some insights into how they might fit into the network theory. Baayen et al. (1997) show that for monomorphemic words the frequency of the plural influences the reaction time in lexical decision experiments, so that words matched on singular frequency differ in reaction time, where the word with the higher plural frequency is reacted to faster. On the other hand, the cumulative root frequency of the root in such a monomorphemic word did not have an effect. The variable that was relevant was family size, a type count, not cumulative root frequency, a token count. In this example we notice that where a paradigmatic relationship is relevant (singular-plural), token frequency had an effect, while where a derivational relationship was relevant (cumulative root frequency of all words and compounds with that root), type frequency had an effect. While this is a topic for future research, it is possible that token frequency is relevant in paradigmatic relationships, and type frequency in derivational ones. If this is the case, we can imagine two ways of representing a paradigm, illustrated in Figures 6.2 and 6.3. One is where the same connections appear between members of a paradigm as between all other words, but paradigmatic connections are stronger (Figure 6.2). The other is a satellite-entries organization: a basic form connected to all other forms of the paradigm (Lukatela et al., 1980) (Figure 6.2). Support for this representation comes from an experiment performed by Lukatela and colleagues. They presented Serbo-Croatian nouns for lexical decision, and nominative singular forms were reacted to faster, even if they were not the most frequent in the paradigm. In addition, the decision times for two other cases - instrumental singular and dative singular - did not differ, although they did differ in frequency.

Depending on which way paradigms are represented, their basic members are represented differently as well. If paradigms have all the same connections as the other words, just stronger, there are no basic members in the paradigm. If, on the other hand, the paradigm is organized as a satellite, then the entry that shows the fastest responses in lexical decision

Figure 6.2: Paradigm connections. Connections exist between all members of the paradigm. Not all of them are shown for clarity. *Stol* ‘table’ nom. sg., *stola* gen. sg., *stole* loc. sg., *stolom* inst. sg., *stolu* dat. sg.

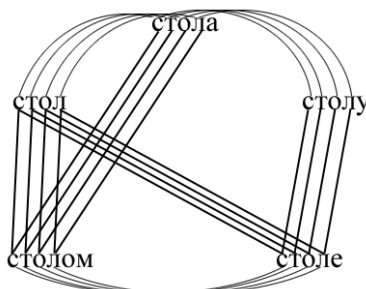
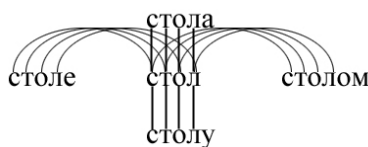


Figure 6.3: Paradigm organized in a satellite-entry fashion. The nominative singular form is central, all others are connected to it. Forms illustrated same as above.



will be the basic member. The details of such an organization are best left to future research.

It is important to note, however, that care should be taken in defining paradigms. While in the nominal singular paradigm there is no difference in meaning between the entries aside from syntactic context, even changing the number to plural changes meaning slightly. Similarly, in a verbal paradigm, there are always meaning differences between forms, and that needs to be taken into account when working with paradigms.

Based on the above discussion of paradigms, we can imagine that there are different possibilities as to how frequency information is collected in the mental lexicon. One possibility is that the entry for each form contains frequency information of some kind. On the other hand, if the entries are organized in a satellite fashion, frequency information might be collected only for the basic member of the paradigm. Since we are dealing mostly with verbs in this investigation, and there are always meaning differences between verb forms, I will assume that every verbal type has some frequency information associated with it.

To summarize, I assume that every lexical entry has meaning and form associated with it, as well as frequency information. Entries have links of varying strength between each other, also based both on form and meaning. Paradigms are units, and there are two possibilities as to how they are organized - either in a satellite-entry fashion where there is a basic member

and all other members linked to it, or when there are the same links between paradigm members, just stronger.

### 6.3.1 Relative frequency effects

Now consider how relative frequency effects could appear in such a system. According to previous studies (Cole et al., 1997; Hay, 2001, 2002; Burani and Thornton, 2003; Zuraw, 2009), and results of the experiment in Chapter 4, words that are more frequent than their bases are less complex morphologically than words that are less frequent than their bases. I offer some ideas as to why such an effect might appear.

Morphological priming is a well-established effect that appears in many studies (e.g., Forster and Azuma (2000) and Pastizzo and Feldman (2004)). There are several ways of interpreting this phenomenon; two most prominent ones are that there are common parts between words that result in priming (e.g., Napps (1989)), and that when a word is activated, related words are also accessed (e.g., Gonnerman and Andersen (2001)). Since in this dissertation, and in chapter 3 specifically, I argue for whole-word processing, I am going to take the latter position. Assuming that when a word is accessed, related words are also accessed through the connections between the words, we then conclude that when a base is activated, derived words related to it are also activated. The question that now arises is what happens when a derived word is accessed? If all the words that share connections with that derived word are accessed, then the base must be accessed as well. However, as the derived word becomes more frequent, the link between the base and the derived word becomes weaker, since the derived word develops a representation of its own. Similarly, if the base word becomes less and less frequent, while the derived word preserves its frequency, the link between the two words also becomes weaker. Thus, the strength of the connection between a base and derived word is modulated by the frequency difference between the two words.

Assuming that the activation of related words proceeds with activating the words that begin with the same sequence of letters as the main word, then there must be differences between connection strength between a base word and its prefixed relative, and a base word and its suffixed relative. In case with prefixed words the derived word always starts with another letter/phoneme sequence than the base word, and we might imagine that the related words that are activated first, start with the same letter/phoneme sequence as the main word being activated. Thus, words that are related to the word being activated, but start with a different letter/phoneme sequence, are activated, but not to the same degree as words that do start with the same sequence. This would presuppose a difference between prefixes and suffixes, where the relationship between a base and a derived suffixed word is more obvious than a relationship between a base and its derived prefixed word.

In fact, some evidence exists in favor of such an asymmetry between prefixes and suffixes. Some experiments have been performed in order to investigate prefix/suffix asymmetries. Gonnerman and Andersen (2001) performed a cross-modal lexical decision task, where



the primes were presented audially, and targets visually. Stimuli varied on semantic and phonetic similarity, and included both prefixed and suffixed words. Mostly, there were no differences between processing of prefixed and suffixed words. The only difference was in the orthographic condition, where primes and targets had the same beginning (in case of suffixes, e.g., *spinach* and *spin*) or ending (in case of prefixes, e.g., *coffee* and *fee*) sequence, but were not related morphologically. In case of suffixes, there was no priming or inhibition between such primes and targets, while in case of prefixes there was significant inhibition. On the other hand, Randall and Marslen-Wilson (1999) performed a self-paced reading experiment and measured reaction times in sentences that were read word-by-word. Overall, sentences that contained prefixed words were read more slowly than suffixed words, and prefixed novel words were read more slowly than suffixed novel words. In addition to that, Hay and Baayen (2002) in their study of 80 affixes found that ‘50% of suffixes (27/54) and 31% of prefixes (8/26) show a significant correlation between base and derived frequency. This provides evidence that suffixes tend to be more decomposable and lead to more semantically transparent forms than prefixed forms.’ These results supports our tentative conclusion that there is an asymmetry between prefixed and suffixed words relative to their base words in that there is stronger activation of related suffixed words than related prefixed words. This could be further experimentally tested. We can predict that in such an experiment, the same difference between base and derived frequencies would amount to a larger processing effect in suffixed-base pairs than in prefixed-base pairs.

### 6.3.2 Root hierarchy

In Chapter 3 I showed that roots form a hierarchy in Russian, where words with free roots are more easily decomposable than words with modified roots, and words with bound roots are the least decomposable. Root type was also a significant predictor in the same direction in the complexity rating experiment. This effect fits into the network model straightforwardly: since words with free roots have more and stronger connections to related unprefixated words, the root is more discernible. The connections are weakest for bound roots, stronger for modified roots and strongest for free roots.

### 6.3.3 Semantic transparency in the network model

As semantic transparency was an important factor in all the models of experimental results, the question of semantics, and how it fits into the network model is an important issue that needs to be addressed. Gonnerman and Andersen (2001) performed cross-modal priming experiments with prefixed and suffixed words, where prime-target pairs were morphologically related and were rated on semantic similarity. They found that words that were more semantically related produced larger priming effects than words that were less semantically related. For example, pairs like *lately-late* (similarity rating 3.9) had a priming effect of 19 ms, while pairs like *boldly-bold* (similarity rating 6.1) had a priming effect of 40 ms, and

a similar difference appeared in processing of prefixed words. Based on the simple model of word activation described above, we can assume that more semantic similarity leads to higher levels of activation of related words.

Another result that is important for considering semantics in the network model is the result of a self-paced reading experiment performed by Randall and Marslen-Wilson (1999). Sentence reading times varied depending on the item included in the sentence. Items used in sentences were either novel or established words, they varied on whether they were prefixed or suffixed, suffixes and prefixes were either productive or unproductive, and sentence context was either constrained or not constrained. All these manipulations produced differing results. In addition, there were differences in processing of sentences with regular and irregular verbs. The conclusion the authors make is that morphological and syntactic processing are not independent of each other.

These two studies demonstrate that semantics is an important component of morphological processing, and that the effect is gradient, and that semantic processing is not independent of syntactic processing. What this means is that more semantic similarity between a pair of words leads to stronger connections between these words and, as a result, stronger activation.

### 6.3.4 Differences between prefixes

In Chapter 5 I report on the results of a complexity rating experiment that suggests that there are differences between processing of different prefixes in Russian verbs. Verbs with prefixes that have a higher transparency ratio were rated as more complex structurally than verbs with prefixes that have a lower transparency ratio. This finding also fits well into the network theory of morphology. Words with the more concrete prefixes have stronger connections to other words with the same prefixes than words with the less concrete prefixes, and hence the meaning of the prefix is more ‘available’. It is not clear whether the prefix generalization over prefixed words is temporary (not stored, and only used for new words), or permanent (stored), and I discuss this issue more below.

## 6.4 Status of roots and affixes

The network theory of morphology, which is word-based, predicts that no morphemes will be units on their own. Thus, neither roots nor affixes are predicted to be standalone form-meaning pairings. In Chapter 3, I present the results of a prefix separation experiment that demonstrate that roots in Russian form a hierarchy, where bound roots are least easily separated, modified roots more so, and free roots are easiest to separate. This result is confirmed in the complexity ratings experiment where root type is a significant predictor. This difference between types of roots does not fit well into either the all-root theories, nor into the no-root theories of morphology, and I conclude that the treatment of the network

theory of morphology provides the best explanation: roots are generalizations over whole words, and only words are units. The stronger the connections between words, both form and semantic, the more evident are the generalizations.

We can treat suffixes in the same way. The stronger the connections between words with the same prefix, the more evident the generalization. However, there are important differences between roots and affixes. There are many more roots than affixes, and the meaning of a root is not definable without a word where it occurs. Most affixes, on the other hand, have templatic, easily definable meanings, independent of the word where they occur. In the following discussion I mean ‘Russian verbal prefixes’ when I discuss prefixes, and ‘Russian roots’ when referring to roots.

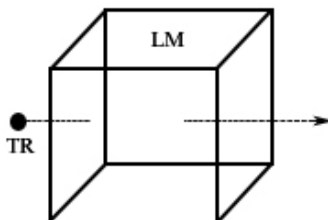
Consider a few examples. The root *-ryb-* in Russian has something to do with fish: *rybnyj* ‘fish’ (adj.), *ryba* ‘fish’ (n.), *rybak* ‘fisherman’. The suffix *-nik-*, according to (Townsend, 1975), has several meanings:

- Denoting persons. For example, *pomosch* ‘help’, *pomoschnik* ‘helper, assistant’.
- Denoting objects. For example, *gradus* ‘degree’, *gradusnik* ‘thermometer’.
- Denoting places. For example, *korova* ‘cow’, *korovnik* ‘cowshed’.

If we add the suffix to the root *-ryb-*, we get the word *rybnik*. According to the meanings of the suffix above, we get the following meanings:

- A person having to do with fish. For example, ‘fisherman’, or ‘fish seller’. Wikipedia gives the first two meanings, and the following additional ones: ‘supervisor of fishing industry’, ‘bird that eats fish’.
- Some object having to do with fish. Could be an instrument of fishing, or a dish with fish. The second meaning is given by Wikipedia.
- A container for fish. This meaning is given by Wikipedia.
- An additional meaning that is given in Wikipedia, but does not fit into the meanings of the suffix above is ‘fishing’.

Using the same suffix, we could easily think up new meanings of this word, for example, ‘someone who eats a lot of fish’, ‘fish cook’, ‘someone who feeds the fish’, ‘someone who raises fish’, etc., all these meanings in the realm of the first meaning of the suffix *-nik-*. Thus, many words can be denoted by the combination of this root and a nominal suffix. It might be argued that the nominal suffix itself has many meanings, and thus the number of resulting meanings is high. Take, however, a more semantically constrained suffix, *-ovschik* which forms names of professions from nouns, for example, *chasovschik* ‘watchmaker’ from *chasy* ‘watch’. Adding *-ovschik* to the root *-ryb-*, we get *rybovschik*, a word that does not

Figure 6.4: Prefix *pro-*

exist, but would mean ‘someone whose profession has to do with fish’. And again, it is not possible to tell which profession it is - a fisherman, a fish seller, a fish cook, etc.

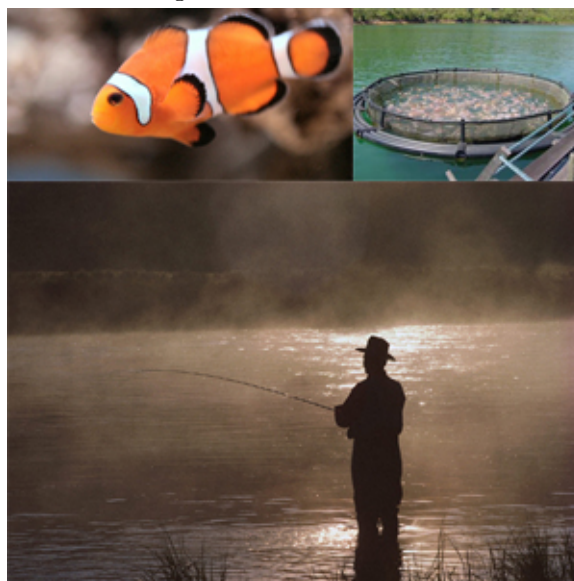
On the other hand, prefixes have very constrained meanings that are easily applicable to new words. For example, one of the meanings of the prefix *po-* is ‘to do for a while’. If the prefix is applied to a verb with a known meaning, the resulting meaning will be quite predictable. For example, a new verb *gejmit* ‘to play computer games’, when combined with the prefix *po-*, will mean ‘to play computer games for a while’ (*pogejmit*). Similarly, when combined with the prefix *za-* ‘to start doing’ (one of the meanings), *zagejmit* will mean ‘to start playing computer games’.

The most clear difference between roots and prefixes is that prefixes are conceptually dependent (Langacker, 2002), because prefixes require a verb as part of their meaning. Since I argue that roots are not meanings units on their own, they cannot be either conceptually dependent or conceptually autonomous. However, words formed from roots are conceptually autonomous.

The conceptual dependency of prefixes, independent of language, is similar to the way that constructions (Goldberg, 1995) are structured in that they have meaning, form, and slots for their arguments. In fact, Goldberg cites Saussure stating that ‘morphemes are clear instances of constructions in that they are pairings of meaning and form that are not predictable from anything else (Goldberg, 1995, p.4). While I agree that in many instances affixes are, in fact, constructions, I argue that roots are just generalizations over related words. Prefix meaning involves arguments (or ‘roles’ as Goldberg (1995) calls them), and root meaning does not. In that sense prefixes are like constructions, and roots are not.

As an example, consider Figure 6.4 and Figure 6.5. They depict the meaning of the prefix *pro-* that roughly means ‘through’, used both temporally (e.g. *prostojat* ‘to stand for a long time’) and spatially (e.g. *projti* ‘to go through’, along with other meanings), and the meaning of the root *-ryb-* that has to do with fish, respectively.

Figure 6.4 illustrates the prefix *pro-* using the concepts proposed by Johnson (1987), where the landmark (LM) is the background and the trajectory (TR) is the moving object

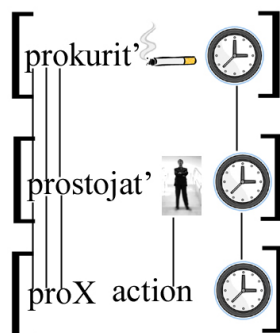
Figure 6.5: Root *-ryb-*

or figure. Thus, in case of *pro-*, there is a trajectory that moves *through* a landmark. In spatial uses the landmark is spatial (a container, etc.), and in temporal uses it is time. As I show in Antić (2006), in case of Russian verbal prefixes, the trajectory is in fact a process that manifests itself as a verb. The meaning of the prefix is thus quite concrete, although it is conceptually dependent on its argument.

On the other hand, if we try to represent the meaning of the root *-ryb-* that has to do with fish, we must represent it with several pictures, shown in Figure 6.5. These pictures represent the meaning of the words where the root appears: *ryba* ‘fish’, *rybak* ‘fisherman’, *rybozavod* ‘fishfactory’. Of course, the list is not exhaustive, since there are other words that involve this root.

Thus, it is possible that, while roots are not standalone units, prefixes are. It is not clear whether this prefixal ‘template’ that can be generalized to new words is a storage unit or created on the fly. Langacker (2002, p.288) discusses this question in more detail comparing low-level and high-level templates (or schemas, as he calls them).

At least two options exist for representing affixes in the Network model, I discuss them in more detail. It is possible that, as Langacker (2002) suggests, generalizations are formed over many words to produce generalized templates. Considering the Russian prefix *pro-* discussed above. As I discuss in Section 1.3.4, the Russian data shows that this prefix is responsible for introducing predicate or sentence level constraints. This is shown in example (6.1), repeated from Chapter 1.

Figure 6.6: Representation of *prokurit'* and a general schema

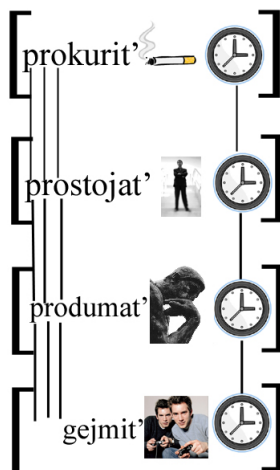
- (6.1) (a) *Ivan*                      *prokuril*    *sigarety*                      *vsju*  
 Ivan.nom.sg   on.smoke.pf.3.sg.masc.past   cigarette.acc.pl   whole.acc.sg  
*noč.*  
 night.acc.sg  
 'Ivan smoked cigarettes the whole night.'
- (b) \**Ivan*                      *prokuril*    *sigarety.*  
 Ivan.nom.sg   through.smoke.pf.3.sg.masc.past  
 'Ivan smoked cigarettes (for a while).'

The fact that the prefix calls for an argument could be seen as an argument for the prefix to be represented as a generalization, shown in Figure 6.6, repeated from Chapter 1. The generalized template would reflect that there are two arguments required: a verb to which the prefix would attach, and a time expression. This generalized template could then be used with new words to create new prefixed verbs with time expressions.

However, it is also possible that the creation of new words with existing prefixes happens as a pattern match to other words with the same parts. This is shown in Figure 6.7. There are many connections between verbs with the prefix *pro-* and a time expressions that are used with those verbs. A new verb, such as *gejmit'* 'to play computer games' falls naturally into such a pattern, and a new verb is created without the use of a generalization.

Future investigations that compare prefixes of different frequency would shed more light on which exact representation is used for affixes. Whichever representation is used, extensions of the Network theory of morphology can be easily used to represent predicate and sentence level meaning, as the same principles of form and meaning representation apply. Meaning can be associated with any type of form, including sentences and sentence pieces, as in Construction Grammar (Goldberg, 1995).

Figure 6.7: Representation of a pattern match



## 6.5 Conclusion

In this dissertation I present the Network theory of morphology that is based on words and connections between them. The connections between words are more or less strong depending on word frequency. Based on existing research, and experiments described in this dissertation, I proposed a modification where relative frequency is also taken into account in the model. In addition, evidence for a hierarchy of roots in Russian and differences in processing between different Russian prefixes was presented.

Overall, the Network theory of morphology proved to be a flexible framework able to accommodate different experimental results. In addition, since it is based on connections between words, it is in principle compatible with both strictly word-based and also morpheme-based theories.

As is discussed in Chapter 1, an extension of the Network theory of morphology, with future experimental investigations to show its validity, can handle contexts larger than the word, i.e., sentences and phrases. This extension, where there are form and meaning connections between phrases and parts of sentences, is very similar to Construction Grammar approaches. Such a general approach has the advantage of applying the same general principles to both morphology and syntax. It would be interesting to see if the same probabilistic connections could be applied to organization of syntactic knowledge.

In addition to potentially being able to cover both morphology and syntax, the Network theory of morphology is easily applicable to different types of languages. For example, word connections are easily described in Semitic languages, where theories of morphology based on consecutive word parts have great trouble. In the Network theory of morphology the form

connections need not be linear, and semantic connections can correspond to these non-linear form connections in Semitic languages.

Thus, the Network theory of morphology presents at least two advantages: general applicability, and being informed by experimental results, and I hope that the model will be used and further extended in future investigations.



## Bibliography

- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge University Press.
- Antić, Eugenia. 2006. Russian bare accusatives of time: a systematic treatment. MS.
- Aronoff, Mark. 1976. *Word formation in generative grammar*. MIT Press.
- Baayen, Harald. 1992. *Yearbook of morphology*, chapter Quantitative aspects of morphological productivity, 109–149. Kluwer Academic Publishers.
- Baayen, R. Harald, Rochelle Lieber, and Robert Schreuder. 1997. The morphological complexity of simplex nouns. *Linguistics* 861–877.
- Baker, Mark. 1988. *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.
- Barsalou, L. 1983. Ad hoc categories. *Memo* 11:211–227.
- Beauvillain, Cecile, and Juan Segui. 1992. Representation and processing of morphological information. In *Orthography, phonology, morphology, and meaning*, ed. R. Frost and L. Katz, chapter 19, 377–388. Elsevier Science Publishers B.V.
- Blevins, James P. 2003. Stems and paradigms. *Language* 79:737–767.
- Blevins, James P. 2006. Word-based morphology. *Journal of Linguistics* 531–573.
- Bloomfield, Leonard. 1933. *Language*. New York: Holt.
- Bochner, Harry. 1993. *Simplicity in generative morphology*. Berlin: Mouton de Gruyter.
- Booij, Geert. 2005. Compounding and derivation: evidence for construction morphology. In *Morphology and its demarcations*, ed. Wolfgang U. Dressler, Dieter Kastovsky, Oskar E. Pfeiffer, and Franz Rainer, 109–132. John Benjamins Publishing Company.
- Booij, Geert. 2010. *Construction morphology*. Oxford University Press.

- Boroditsky, Lera. 2000. Metaphoric structuring: understanding time through spatial metaphors. *Cognition* 75:1–28.
- Bowdle, Brian F., and Dedre Gentner. 2005. The career of metaphor. *Psychological Review* 112:193–216.
- Burani, Cristina, and Anna M. Thornton. 2003. The interplay of root, suffix and whole-word frequency in processing derived words. In *Morphological structure in language processing*, ed. Harald Baayen and Robert Schreuder, 157–207. Mouton de Gruyter.
- Butterworth, Brian. 1983. Lexical representation. In *Language production*, ed. Brian Butterworth, 257–294. Academic Press.
- Bybee, Joan. 1988. Morphology as lexical organization. In *Theoretical morphology*, 119–141. Academic Press.
- Bybee, Joan. 2007. *Frequency of use and the organization of language*. Oxford University Press.
- Bybee, Joan L. 1985. *Morphology: a study of the relation between meaning and form*. John Benjamins Publishing Company.
- Caramazza, Alfonso, Alessandro Laudanna, and Cristina Romani. 1988. Lexical access and inflectional morphology. *Cognition* 28:297–332.
- Carstairs-McCarthy, Andrew. 2002. *Introduction to english morphology*. Edinburgh University Press.
- Chierchia, Sally, Gennaro McConnell-Ginet. 2000. *Meaning and grammar*. The MIT Press.
- Cole, P., C. Beauvillain, and J. Segui. 1989. On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language* 28:1:1–13.
- Cole, P., J. Segui, and M. Taft. 1997. Words and morphemes as units for lexical access. *Journal of Memory and Language* 37:312–330.
- Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge University Press.
- Corbett, Greville G., and Norman M. Fraser. 1993. Network morphology: a datr account of russian nominal inflection. *Journal of Linguistics* 29:113–142.
- Crawley, Michael J. 2007. *The r book*. John Wiley & Sons, Ltd.

- Elson, Mark J. 1997. The evolution of macedonian verbal paradigms in the context of lexeme-based morphology. *The Slavic and East European Journal* 41:645–659.
- Feldman, Laurie. 1994. Beyond orthography and phonology: Differences between inflections and derivations. *Journal of Memory and Language* 33:442–470.
- Fillmore, Beryl T., Charles J. Atkins. 1992. Toward a frame-based lexicon: the semantics of risk and its neighbors. In *Frames, fields and contrasts: New essays in semantic and lexical organization*, ed. Adrienne Lehrer Eva Feder Kittay. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fiorentino, Robert, and David Poeppel. 2007. Compound words and structure in the lexicon. *Language and Cognitive Processes* 22:953–1000.
- Forster, Kenneth I., and Tamiko Azuma. 2000. Masked priming for prefixed words with bound stems: Does *submit* prime *permit*? *Language and Cognitive Processes* 15:539–561.
- Franks, Steven. 1995. *Parameters of slavic morphosyntax*. Oxford University Press.
- Frauenfelder, Uli H., and Robert Schreuder. 1992. Constraining psycholinguistic models of morphological processing and representation: the role of productivity. In *Yearbook of morphology*, ed. Geert Booij and Jaap van Marle, 165–183. Kluwer Academic Publishers.
- Gentner, D., and B. F. Bowdle. 2008. Metaphor as structure-mapping. In *The cambridge handbook of metaphor and thought*, 109–128. New York, NY: Cambridge University Press.
- Giraud, Hélène, and Jonathan Grainger. 2000. Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes* 15:421–444.
- Giraud, Hélène, and Jonathan Grainger. 2003. On the role of derivational affixes in recognizing complex words: Evidence from masked priming. In *Morphological structure in language processing*, ed. Harald R. Baayen and Robert Schreuder, 209–232. Mouton de Gruyter.
- Gladney, Frank Y. 2006. Slavic morphology. *Glossos* 8:1–48.
- Goldberg, Adele. 1995. *Construction grammar: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Gonnerman, Laura, and Elaine Andersen. 2001. Graded semantic and phonological similarity effects in morphologically complex words. In *Morphology 2000*, ed. W. Dressler. John Benjamins.

- Gurevich, Olga I. 2006. Constructional morphology: The georgian version. Doctoral Dissertation, University of California-Berkeley.
- Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20*, ed. Kenneth Hale and S. Jay Keyser. Cambridge, MA: MIT Press.
- Hay, Jennifer. 2001. Lexical frequency in morphology: is everything relative? *Linguistics* 1041–70.
- Hay, Jennifer. 2002. From speech perception to morphology: affix ordering revisited. *Language* 527–555.
- Hay, Jennifer, and Harald Baayen. 2002. Parsing and productivity. In *Yearbook of morphology*, ed. Geert Booij and Jaap van Marle. Kluwer Academic Publishers.
- Johnson, Mark. 1987. *The body in the mind: The bodily basis of meaning, imagination, and reason*. Chicago: University of Chicago Press.
- Kiparsky, Paul. 1982. Lexical morphology and phonology. In *Linguistics in the morning calm*, ed. The Linguistic Society of Korea, 3–91. Seoul: Hanshin Publishing.
- Langacker, Ronald. 2002. *Concept, image and symbol: the cognitive basis of grammar*. Berlin: Mouton de Gruyter.
- Lieber, Rochelle. 1980. *On the organization of the lexicon*. Indiana University Linguistics Club.
- Losiewicz, Beth. 1992. The effect of frequency on linguistic morphology. Doctoral Dissertation, University of Texas-Austin.
- Lukatela, G., B. Gligorišević, A. Kostić, and M.T. Turvey. 1980. Representation of inflected nouns in the internal lexicon. *Memory and Cognition* 8:415–423.
- Lunt, Horace G. 2001. *Old church slavonic grammar*. Mouton de Gruyter.
- Manelis, Leon, and David A. Tharp. 1977. The processing of affixed words. *Memory & Cognition* 5:690–695.
- Marslen-Wilson, William, Lorraine Komisarjevsky Tyler, Rachel Waksler, and Lianne Older. 1994. Morphology and meaning in the english mental lexicon. *Psychological Review* 101:3–33.
- McCormick, Samantha F., Kathleen Rastle, and Matthew H. Davis. 2008. Is there a ‘fete’ in ‘fetish’? effects of orthographic opacity on morpho-orthographic segmentation in visual word recognition. *Journal of Memory and Language* 58:307–326.

- Melinger, Alissa. 2003. Morphological structure in the lexical representation of prefixed words: Evidence from speech errors. *Language and Cognitive Processes* 18:335–362.
- Napps, Shirley. 1989. Morphemic relationships in the lexicon: Are they distinct from semantic and formal relationships? *Memory & Cognition* 17:729–739.
- Pastizzo, Matthew, and Laurie Feldman. 2004. Morphological processing: A comparison between free and bound stem facilitation. *Brain and Language* 90:31–39.
- Pereltsvaig, Asya. 2000. On accusative adverbials in russian and finnish. In *Adverbs and adjunction. linguistics in potsdam*, ed. A. Alexiadou and P. Svenonius, 155–76. University of Postdam.
- Raab, D.H. 1962. Statistical facilitation of simple reaction time. *Transactions of the New York Academy of Sciences* 574–590.
- Randall, Belinda, and William Marslen-Wilson. 1999. The relationship between lexical and syntactic processing. In *Proceedings of the Twentieth Annual Conference of Cognitive Science Society*, ed. M.A. Gernsbacher and S.J. Derry. Lawrence Erlbaum Associates.
- Rastle, Kathleen, Matthew H. Davis, and Boris New. 2004. The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin & Review* 11:1090–1098.
- Reid, Agnieszka Anna, and William David Marslen-Wilson. 2003. Lexical representation of morphologically complex words: evidence from polish. In *Morphological structure in language processing*, ed. Harald Baayen and Robert Schreuder, 287–336. Mouton de Gruyter.
- Robins, Robert H. 1959. In defence of wp. *Transactions of the Philological Society* 116–144.
- Schreuder, Robert, and Harald Baayen. 1994. Prefix stripping re-visited. *Journal of Memory and Language* 33:357–375.
- Siegel, Dorothy. 1979. *Topics in english morphology*. New York: Garland.
- Stemberger, Joseph Paul, and Brian MacWhinney. 1988. *Theoretical morphology*, chapter Are Inflected Forms Stored in the Lexicon?, 101–116. Academic Press.
- Stump, Gregory T. 2001. *Inflectional morphology: a theory of paradigm structure*. Cambridge University Press.
- Svenonius, Peter. 2004. Slavic prefixes and morphology. *Nordlyd* 32:177–204.
- Sweetser, Eve. 1999. Compositionality and blending: semantic composition in a cognitively realistic framework. In *Cognitive linguistics: Foundations, scope, and methodology*, ed. Gisela Janssen, Theo Redeker. Mouton de Gruyter.

- Taft, Marcus. 1985. Recognition of affixed words and the word frequency effect. *Memory and Cognition* 7:263–272.
- Taft, Marcus, and Kenneth I Forster. 1975. Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior* 14:638–647.
- Townsend, Charles E. 1975. *Russian word-formation*. Slavica Publishers, Inc.
- Wade, Terrence. 1992. *A comprehensive russian grammar*. Blackwell Publishers.
- Weisser, Philipp. 2006. A distributed morphology analysis of croatian noun inflection. In *Linguistische arbeits berichte 84*, ed. Gereon Müller and Jochen Trommer, chapter A Distributed Morphology Analysis of Croatian Noun Inflection, 131–142. Universität Leipzig.
- Wurm, Lee H. 1997. Auditory processing of prefixed english words is both continuous and decompositional. *Journal of Memory and Language* 37:438–491.
- Wurm, Lee H., Mirjam Ernestus, Robert Schreuder, and R. Harald Baayen. 2006. Dynamics of the auditory comprehension of prefixed words: Cohort entropies and conditional root uniqueness points. *The Mental Lexicon* 1:125–146.
- Zuraw, Kie. 2009. Frequency influences on rule application within and across words. In *Proceedings of the 43rd Annual Meeting of the Chicago Linguistic Society*.
- Zwisterlood, P. 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32:25–64.

## Appendix A

### List of neologisms used for the productivity test

Word	Gloss	Number of Google results
<i>avangardit'</i>	'to do something vanguard'	62
<i>balansirovat'</i>	'to balance'	358000
<i>brendit'</i>	'to brand'	3910
<i>burokratizirovat'</i>	'to bureaucratize'	29000
<i>gangsterit'</i>	'to be a gangster'	29
<i>gejmit'</i>	'to game'	5220
<i>guglit'</i>	'to google'	50700
<i>dajvit'</i>	'to SCUBA dive'	4280
<i>developit'</i>	'to develop'	6820
<i>diversificirovat''</i>	'to to diversify'	294000
<i>dizajnit'</i>	'to design'	31000
<i>imejlit'</i>	'to e-mail'	107
<i>investirovat'</i>	'to invest'	1380000
<i>indeksirovat'</i>	'to index'	113000
<i>insajdit'</i>	'to earn through inside information'	168
<i>integrirovat'</i>	'to integrate'	835000
<i>kastingovat'</i>	'to cast'	2160
<i>kvotirovat'</i>	'to impose a quota'	72300
<i>klonirovat'</i>	'to clone'	216000
<i>kolbasit'sja</i>	'to fool around'	24300
<i>kommercializirovat'</i>	'to commercialize'	15900
<i>kompilit'</i>	'to compile'	129000
<i>kompromitirovat'</i>	'to compromise'	6250
<i>konvertirovat'</i>	'to convert'	708000

<i>konsaltit'</i>	'to consult'	317
<i>kreativit'</i>	'to do something creative'	143000
<i>kritikovat'</i>	'to criticize'	835000
<i>kserit'</i>	'to copy' (on a copy machine)	73000
<i>lizingovat'</i>	'to lease'	1030
<i>liftingovat'</i>	'to do face lifting'	74
<i>pilingovat'</i>	'to do face peeling'	1600
<i>piratit'</i>	'to pirate'	28400
<i>pressingovat'</i>	'to pressure'	162000
<i>provajdit'</i>	'to provide'	1260
<i>programmirovat'</i>	'to program'	608000
<i>rejtit'</i>	'to rate'	440
<i>rekrutit'</i>	'to recruit'	1780
<i>roumit'</i>	'to roam'	251
<i>seksit'sja</i>	'to have sex'	36000
<i>servisit'</i>	'to service'	667
<i>skanit'</i>	'to scan'	21900
<i>skrabit'</i>	'to do body scrubbing'	12600
<i>spamit'</i>	'to spam'	124000
<i>tjuningovat'</i>	'to tune up'	26700
<i>figačit'</i>	'to make'	14600
<i>fludit'</i>	'to flood'	498000
<i>frilansit'</i>	'to freelance'	20300
<i>šejpingovat'</i>	'to do fitness'	56
<i>šopit'sja</i>	'to shop'	22800
<i>esemesit'</i>	'to sms'	4380
<i>kommentit'</i>	'to comment'	59900
<i>frendit'</i>	'to friend'	21700
<i>daunlodit'</i>	'to download'	2410
<i>apgrejdit'</i>	'to upgrade'	111000
<i>juzat'</i>	'to use'	407000

Table A.1: Words used for the productivity test.