

On Resolving Long Distance Dependencies in Russian Verbs

Dirk Saléschus

University of Konstanz, Germany

Abstract. Morphological analyses based on word syntax approaches can encounter difficulties with long distance dependencies. The reason is that in some cases an affix has to have access to the inner structure of the form with which it combines. One solution is the percolation of features from the inner morphemes to the outer morphemes with some process of feature unification. However, the obstacle of percolation constraints or stipulated features has led some linguists to argue in favour of other frameworks such as, e.g., realizational morphology or parallel approaches like optimality theory. This paper proposes a linguistic analysis of two long distance dependencies in the morphology of Russian verbs, namely secondary imperfectivization and deverbal nominalization. We show how these processes can be reanalysed as local dependencies. Although finite-state frameworks are not bound by such linguistically motivated considerations, we present an implementation of our analysis as proposed in [1] that does not complicate the grammar or enlarge the network unproportionally.

1 Secondary Imperfectivization in Russian Verbs

Like tense and person, aspect is one of the grammatical categories of Russian verbs. The aspectual category consists in the opposition between perfective and imperfective aspect. The majority of Russian verbs expresses this grammatical distinction, although its morphological exponence is quite complex. Prefixation, suffixation, suppletion, stem allomorphy, and a combination thereof can be used. We will not consider the intricate semantics of the aspect [2] but only concentrate on the forms, especially on the joint interaction of prefixation and suffixation. An overview of the data can be found in any grammar of Russian, as e.g. in [3].

We start with simple verb stems. The overwhelming majority of them are imperfective, however, there also exist perfective simple stems. The aspect of simple stems is thus an idiosyncratic property and has to be marked for each single stem.

Before turning to the morphological expression of the aspectual opposition it is important to consider first the formation of complex verb stems by means of lexical prefixation. Lexical prefixation of verbs is quite productive in Russian. It is a derivational process and leads to the formation of both new lexemes and complex stems. From a semantic point of view this process of lexeme formation can either be rather opaque or can lead to quite transparent composed meanings

as, e.g., in the case of different semantic classes called *Aktionsarten*. See the following examples for semantically transparent and opaque lexical prefixation:¹

- | | | | | |
|---|-----------------|--------------------|-----------------|-------------------|
| 1 | <i>nosít'</i> | (carry-indet.ipf.) | <i>vnosít'</i> | (carry in-pf.) |
| | <i>vynosít'</i> | (carry out-pf.) | <i>iznosít'</i> | (to wear out-pf.) |
| 2 | <i>begát'</i> | (run-indet.ipf.) | <i>vbegát'</i> | (run inside-ipf.) |
| | <i>vybegát'</i> | (run out-ipf.) | <i>izbegát'</i> | (avoid-ipf.) |

We also give an example for semantically transparent lexical prefixation with the ingressive *Aktionsart* from [4]:388 f.:

- | | | | | |
|---|-----------------|--------------|-------------------|----------------------|
| 3 | <i>govorít'</i> | (speak-ipf.) | <i>zagovorít'</i> | (start speaking-pf.) |
| | <i>igrát'</i> | (play-ipf.) | <i>zaigrát'</i> | (start playing-pf.) |
| | <i>kričát'</i> | (cry-ipf.) | <i>zakričát'</i> | (start crying-pf.) |

From a morphological point of view, there are around 20 prefixes which can be used for lexical prefixation with both perfective and imperfective simple verb stems. Lexical prefixation can also be applied cyclically, leading to complex forms such as:

- | | | | | |
|---|----------------------|-------------------|------------------|--------------|
| 4 | <i>polnít'</i> | (sloppy:fill-pf.) | <i>výpolnít'</i> | (fulfil-pf.) |
| | <i>perevýpolnít'</i> | (overfulfil-pf.) | | |

This phenomenon is also found in other languages like English or German which forms the verbs *füllen*, *erfüllen*, and *übererfüllen*, respectively.

Note, however, that not every complex stem is formed from an actual existing base stem. There are verbs like *dobávit'* (fill-pf.), *pribávit'* (add-pf.), *zabávit'* (amuse-pf.), without an existing verb *bávit'*. Even though they look like complex stems they have to be analyzed as simple stems. This is similar to English morphology with verbs like *perceive*, *receive* with no existing word *ceive* (see [5] for a discussion of such examples).

A last fact to note about lexical prefixation is that some stems have only one or a few actual prefixed variants whereas others combine with lots of prefixes [4]. See, e.g., the possible lexical derivations of the stem *xodít'* (go-indet.ipf.):

vxodít', *vsxodít'*, *vyxodít'*, *doxodít'*, *zaxodít'*, *isxodít'*, *naxodít'*, *obxodít'*, *otxodít'*, *perevxodít'*, *poxodít'*, *podxodít'*, *príxodít'*, *proxodít'*, *rasxodít's'a*, *sxodít'*, *uxodít'*; only exception: **nadxodít'*

¹ The following writing conventions are adopted here: the *y* stands for the high back unrounded dorsal [ɨ]. A soft consonant is a consonant that is palatal or has secondary palatalization. The latter feature is signalled by an apostrophe after the consonant (e.g. *t'*). The softness of consonants is predictable when they are followed by the front vowels *i* or *e* and is left out in these contexts. The symbols *č*, *š* and *ž* stand for a soft alveo-palatal affricate [tʃ] and the posterior voiced and unvoiced fricatives [ʃ] and [ʂ], respectively. At the surface, the *č* is always soft whereas *š* and *ž* are always hard. Finally, an accent signals stress.

Our lexicon shall be fully productive and contain all potential complex word forms, whether actually existing in Russian or not.

From a grammatical point of view there is one important biproduct associated with lexical prefixation. All newly formed lexemes are always perfective stems. In other words, lexical prefixation always leads to perfectivization (with just a few exceptions). This shows the intricate connection between lexical and aspectual meaning [6] but its description goes beyond the scope of this work.

Now we can look at the exponence of the aspectual opposition by means of prefixation and suffixation (ignoring other morphological exponences). Imperfective simple verb stems can express the perfective aspect by grammatical prefixation. From a semantic point of view this stem formation does not alter the lexical meaning. From a morphological point of view a new complex stem is formed.

The crucial fact is that the set of prefixes used for grammatical prefixation is identical to the set of prefixes used for lexical prefixation. However, there is one important difference between grammatical and lexical prefixation – for each simple imperfective stem there is exactly one prefix which is used exclusively for grammatical prefixation. All remaining prefixes can be used for lexical prefixation. The choice of the grammatical prefix that can combine with an imperfective simple stem is not predictable and has to be marked for every simple imperfective stem. The following sketch with some prefixes and stems illustrates this (“G” means grammatical prefixation, “L” means lexical prefixation):

<i>stróit'</i> (build-ipf.)	po	G: <i>postróit'</i> (build-pf.)
		L: <i>popísát'</i> (write a bit-pf.)
		L: <i>podélat'</i> (carry on-pf.)
<i>pisát'</i> (write-ipf.)	na	L: <i>nastróit'</i> (adjust-pf.)
		G: <i>napísát'</i> (write-pf.)
		L: <i>nadélat'</i> (cause-pf.)
<i>délat'</i> (do-ipf.)	s	L: <i>sostróit'</i> (look surly-pf.)
		L: <i>spísát'</i> (copy-pf.)
		G: <i>sdélat'</i> (do-pf.)

If a simple imperfective stem has formed a grammatically prefixed partner stem both stems together make up an aspectually complete verbal lexeme. Since the aspectual opposition is not expressed via exponence on the same stem (as is the case for other grammatical categories like number or person) this morphological process is called grammatical derivation [7].

Now let us see how a (simple or complex) perfective stem expresses the imperfective form in order to create an aspectually complete lexeme. In traditional analyses simple perfective verb stems can change the stem vowel to express imperfective partner stems. In our analysis this is considered to be a case of suffixation. There are two allmorphs of the imperfective suffix: an empty V-slot

and the string /yv/. If not filled by an adjacent vowel from some suffix the empty slot is per default filled with the vowel [a]. The right allomorph of the ipf. suffix is determined by morphological class membership. Interestingly the allomorph for simple perfective stems is always the same, namely *a*:

- 5 throw: *brósit'* (pf.) *brosát'* (ipf.)
- deprive: *lišít'* (pf.) *lišát'* (ipf.)

The *i* in *brosít'* could also be analyzed as an aspectual suffix. However, we analyze it as a thematic vowel for one morphological class of simple perfective stems. The root would then be *bros*. It is a fact that morphological verb classes assign thematic vowels to stems of a paradigm in different ways. Sometimes this vowel is kept in only some stems of the paradigm, sometimes in almost all stems. This separation of thematic vowels from roots facilitates the analyses of the imperfective suffix and of the deverbal nominalization.

Complex perfective verb stems also use suffixation. For these stems this process is called secondary imperfectivization. This is a grammatical process only and is never an option for simple imperfective verb stems. Complex pf. stems normally take the allomorph *yv* but some have the allomorph *a* and some even have both:

- 6 manufacture: *izgotóvit'* (pf.)
- izgotovl'át'* (ipf.) or *izgotávlivat'* (ipf.)

Finally, some (simple or complex) perfective stems show consonant alternations when imperfectivized while others do not. This also has to be marked lexically:

- 7 stem allmorph: render: *javít'* (pf.) *javl'át'* (ipf.)
- manufacture: *izgotóvit'* (pf.) *izgotávlivat'* (ipf.)
- 8 no stem allmorph: throw: *brósit'* (pf.) *brosát'* (ipf.)
- copy: *perepisát'* (pf.) *perepísyvat'* (ipf.)

The following table gives an overview of prefixation and suffixation in Russian verb stems with two example verbs. Empty cells signal that the morphological process is not possible.

		simple verb stem					
		perfective (<i>brós-it'</i>)			imperfective (<i>pis-át'</i>)		
prefixation	gramm	lex	none	gramm	lex	none	
	↓						
input to		<i>výbrosit'</i>	<i>brósit'</i>	<i>napisát'</i>	<i>spisát'</i>	<i>pisát'</i>	
	↓						
ipf. suff		<i>vybrásyvat'</i>	<i>brosát'</i>		<i>spísyvat'</i>		

Before turning to the analysis we want to point out an interesting typological fact. In morphological approaches based on word syntax there is some discussion about the notion of a head (e.g. [8] or [9]). On the assumption that Russian complex stems are indeed derived and not stored in the lexicon, both prefixes and suffixes can determine the aspect of the complex stem and thus constitute the head. The last applied affix always has the last word.

2 Blocking Affixation with Derived Flag Diacritics

The crucial question is how secondary imperfectivization can be blocked for grammatically prefixed perfective complex stems. We will first sketch the solution informally and then consider possible implementations.

In fact, the imperfective suffix needs two kinds of information. First, is the stem perfective or imperfective? Suffixation is only possible for perfective verb stems.

The second question is whether the verb is lexically prefixed. Secondary imperfectivization applies only if the complex stem is created by lexical prefixation, not by grammatical prefixation. Thus, imperfectivization is accomplished in order not to just create imperfective verb stems but to create imperfective partner verbs, i.e. aspectually complete lexemes.

How can this long distance dependency be captured in a morphological framework? How can a morphological type of affixation be made sensible to the complex morphological as well as lexical structure of a verb stem?

One possibility would be the following. Assume that the imperfective suffix requires a stem which is not yet aspectually complete. Assume also that every stem can signal whether it is aspectually complete by some kind of feature.

Next let us assume that every simple imperfective stem is marked for its matching grammatical prefix (if there is one at all). Let us call this marking the stem-prefix-feature. Once this prefix is encountered the lexeme is saturated and signals that it is aspectually completed by setting the mentioned feature. The imperfective suffix is then blocked from application by reference to that feature. Thus, if the prefix and the stem-prefix-feature match, then a new feature is set and suffixation is blocked by that new feature.

There are several ways to implement these ideas in computational linguistics. [10] discusses different strategies, among them using concurrent rule transducers as in two-level morphology or composing in constraints at compile time. The first solution has the disadvantage of slower performance at runtime whereas the latter solution leads to an enormous increase in network size.

Beesley favours a solution with flag diacritics (also described in [1]). In *xfst*, flag diacritics are part of the normal alphabet insofar as they are interpreted like epsilons. However, the enhanced *xfst* lookup routines process them specially and enforce the dependencies between morphemes. The lookup routines do this by introducing a small amount of memory which suffices to capture the long distance dependencies [10]. The overall flag system in *xfst* is non-monotonic and thus exceed the expressive power of regular transducers [1]. With flags it is in

principle possible to create a blatantly overgenerating lexicon and let the lookup routines rule out impossible or undesired combinations. The only disadvantage of flags is a possible slower performance due to backtracking.

We try to solve the problem with the Russian data by reference to flag diacritics, a feature-like symbol in *xfst*. The solution seems quite obvious: the flag of the imperfective suffix interacts with the flags from stem and prefix according to well-defined conditions. This is broken down into two steps. First the flags of the prefix and the stem interact. Second, the result is handled over to the flag of the imperfective suffix. There are several kinds of flags triggering different processes and so again there are several possible strategies for an implementation of flags. We will not describe the whole inventory of flags in *xfst* but hope that the following discussion is easy to understand. For detailed information we refer the reader to [1].

Working exclusively with flag diacritics, one could assign every prefix an own flag signaling a positive value, e.g. @P.NA.PLUS@ for the prefix *na* and similarly @P.U.PLUS@ for the prefix *u*. The first P stands for an operation over the feature, in that case positive setting of the feature. NA is the name of the feature and PLUS is its value. The whole expression is surrounded by @-signs.

Stems are also assigned flags. A stem flag checks the value of the prefix flag and resets it only if the prefix combines with this stem for grammatical prefixation. E.g., the grammatical prefix of the stem *pis* is *na* and therefore changes its flag value. This is achieved by the flag @N.NA.MINUS@. In that case, N signals negative resetting of the value of the NA feature such that the value is reset to MINUS. The value of another prefix like *u* is left intact by the stem *pis*.

The imperfective suffix, finally, is also assigned a flag requiring a flag with a value set to PLUS. It has to list all possibilities like @R.NA.PLUS@, @R.U.PLUS@ etc.

A simplified example shall illustrate how this works. The following expressions would lead to possible concatenations in *lexc*² (concatenation is accomplished via continuation classes (linked sublexicons) and here marked with a plus for clarification):

```
na@P.NA.PLUS@+pis@N.NA.MINUS@[yv @R.NA.PLUS@ |
@R.U.PLUS@...]
```

```
u@P.U.PLUS@+pis@N.NA.MINUS@[yv@R.NA.PLUS@ |
@R.U.PLUS@...]
```

In the first case the value of the first flag is set to PLUS. This value is reset by the stem flag to MINUS. As a result, the flag of the imperfective suffix requiring a PLUS value can no longer match with this complex stem. In the second case the value of the first flag is again set to PLUS. This time it is left intact by the

² A language for specifying lexicons, also provided by the XEROX finite-state tools [1].

stem flag and the imperfective suffix can successfully check for a PLUS value. This gives us exactly the right results.

With the sketched solution the imperfective suffix would have attached to it flags which check for every possible prefix. This is expressed by the disjunctive listing of flags for the imperfective suffix, indicated by ”|”. One could also have multiple entries for the imperfective suffix, each one bearing only flag. However, we would like to keep the number of flags and morphemes minimal.

There is a second strategy which uses a combination of flags and continuation classes by doubling the entries for stems. To take the example from above, the prefix *na* takes again the flag @P.NA.PLUS@. The first entry for the stem *pis* has the flag @D.NA.PLUS@ where the D indicates that the feature NA is not allowed to have the value PLUS. Thus the stem *pis* may combine with any prefix except the one it uses for grammatical prefixation, namely *na*. The continuation class of that stem is the imperfective suffix. The second entry for *pis* has the flag @R.NA.PLUS@. Here the R indicates the requirement for a preceding flag with the feature NA set to the value PLUS. The absence of a preceding flag or any other flag setting is forbidden. The continuation class of that stem entry can be anything except the imperfective suffix. The obvious disadvantage of that solution is the increase in network size by doubling information.

There is also a third solution with single entries for all morphemes and a minimal number of flags used that is already sketched in [1]. In that analysis all morphemes in the overgenerating *lexc* grammar have a special formal marking. Rewrite rules check the markings of the morphemes and change them into flags in special contexts. To take a concrete example, the prefix *na* has the notation naPLEX, the prefix *u* has the form uPLEX, and the prefix *s* has the form sPLEX. PLEX is a dummy meaning lexical prefix. Stems also have some special formal marking which indicates the prefix that is used for grammatical prefixation. E.g., the stem *pis* has the form pisNA, indicating that the prefix *na* is used for grammatical prefixation. Similarly, the stem *sluš* is notated as slushU. Finally, the imperfective suffix also has an additional formal marking, namely IMPRPLEX. All these additional formal markings are just mnemonic placeholders and could be transformed into a more sophisticated notation according to the linguist’s needs. When the *lexc* grammar is compiled it contains all conceivable combinations of prefixes, stems and suffixes, among them unwanted combinations. In a next step simple rewrite rules delete the special formal marking of the prefix if the prefix happens to have the same form as the extra formal marking of the stem. Rewrite rules are given here in the *xfst* formalism. They denote regular relations which can later be composed with the lexicon.

The form naPLEX, e.g., is replaced by simple na if somewhere after the prefix the form NA is found:

```
define rule1 [n a P L E X -> "na" || _ $[N A] ];
```

All that is left to do now is to transform all special formal markings which have not been deleted into flag diacritics. This is also accomplished by rewrite

rules which are ordered after the prefix rules.

```
define rule2 [P L E X -> "@P.LEX.LEX@"] ;
define rule3 [R L E X -> "@R.LEX.LEX@"] ;
```

In effect, only prefixes which combine with stems in lexical prefixation can now have a flag which is required by the imperfective suffix. The imperfective suffix cannot combine with grammatically prefixed stems because in the complex stems the prefixes do not have the required flag. The long distance dependency between prefix and suffix is thus resolved into a local binary interaction between only two flags. This is similar to the resolution of long distance dependencies in phonology. A phoneme can be conceived as a multi-tiered representation with each feature belonging to its own tier [11]. Two sounds, although separated from each other by intervening sounds, can exert an influence on each other because their features are neighbours on one of these tiers.

Our implementation has several advantages. From a linguistic point of view, it comes close to the theoretical analysis. There are two stages of morphological stem formation and each time only two features are checked. It is not the case that one affix has to have access to the inner structure of a complex stem. In order to achieve these theoretical goals, in our analysis flags are not part of the lexicon but instantiated in a later step by rewrite rules.

From an implementational point of view, we avoid the use of filters and the resulting increase in network size. We also use a minimal amount of information by severely restricting the number of flags and continuation classes. As is to be expected, each time a stem is added to the lexicon the amount of states and arcs is more increased in the hybrid solution than in the solution with only flags. The opposite is true with the number of paths. As far as the mini lexicon and the addition of a handful of stems is concerned, the overall increase in network size is the same for both solutions.

So far we have shown how to block a special case of suffixation. The long distance dependency was resolved into two local phenomena. First, the concatenation of prefixes and stems was checked for some kind of pattern matching. Second, the concatenation of the complex stem and the imperfective suffix was restrained by flag diacritics.

In the next paragraph a long distance dependency is again resolved into a purely local phenomenon.

3 Russian Deverbal Nominalization with /nie/

In [12] a special case of blocking in Russian is described. Almost all simple or complex Russian verb stems as described above can combine with the deverbal nominalization suffix *nie*. See the following examples:

9	<i>pisát'</i> (write-ipf.)	<i>pisánie</i> (writing-N.)	RES/CEN
	<i>pét'</i> (sing-ipf.)	<i>pénie</i> (singing-N.)	SE/CEN
	<i>sobráť'</i> (collect-pf.)	<i>sobránie</i> (meeting-N.)	RES/SE
	<i>starát's'a</i> (try-ipf.)	<i>staránie</i> (endeavour-N.)	SE
	<i>raspisát'</i> (write out-pf.)	<i>raspisánie</i> (timetable-N.)	RES
	<i>spisát'</i> (write off-pf.)	<i>spisánie</i> (writing off-N.)	CEN
	<i>zatverdét'</i> (harden-pf.)	<i>zatverdénie</i> (hardening-N.)	RES/CEN
	<i>izgotóvit'</i> (manufacture-pf.)	<i>izgotovlénie</i> (manufacture-N.)	SE/CEN

The type of the resulting nominalization (adopted from [12] and given at the end of the line) is not predictable and can be a complex event nominal (CEN, e.g. *spisánie*), a simple event nominal (SE, e.g. *staránie*), a result nominal (RES, e.g. *raspisánie*), and one nominalization can also have different types. We will not provide more details about the semantic characteristics and the tests for different types of nominalizations because they are not crucial for the following discussion.

In [12] two generalizations are mentioned that apply to that kind of word formation. First, if nominalization applies the secondary imperfectivization of a lexically prefixed verb then the type of the nominalization is always the same, namely a complex event nominal. Second, if the secondary imperfectivization of a lexically prefixed word does not use the allomorph *yv* but instead the allomorph *a* then the deverbal nominalization suffix *nie* cannot be attached to that verb. As examples for that blocking they cite the following data from lexically prefixed verbs:

10	proclaim:	<i>provozglasít'</i> (pf.)	<i>provozglašát'</i> (ipf.)
	proclamation:	<i>provozglašénie</i>	* <i>provozglašánie</i>
	visit:	<i>posetít'</i> (pf.)	<i>poseščát'</i> (ipf.)
	visit (N):	<i>poseščénie</i>	* <i>poseščánie</i>
	inform:	<i>soobščít'</i> (pf.)	<i>soobščát'</i> (ipf.)
	communication:	<i>soobščénie</i>	* <i>soobščánie</i>

consolidate:	<i>ukrepít'</i> (pf.)	<i>ukrepl'át'</i> (ipf.)
consolidation:	<i>ukreplénie</i>	* <i>ukrepl'ánie</i>
destroy:	<i>razrušít'</i> (pf.)	<i>razrušát'</i> (ipf.)
destruction:	<i>razrušénie</i>	* <i>razrušánie</i>
destroy:	<i>razorít'</i> (pf.)	<i>razor'át'</i> (ipf.)
destruction:	<i>razorénie</i>	* <i>razor'ánie</i>
resolve:	<i>postávit'</i> (pf.)	<i>postavl'át'</i> (ipf.)
resolution:	<i>postanovlénie</i>	* <i>postanovl'ánie</i>

Again this is a case of long distance dependency. The suffix *nie* has to have access to the inner structure of the morphologically complex verb.

In [12] this problem is discussed from the point of view of word syntax. According to their analysis such an approach in combination with locality conditions on affixation has to abuse feature marking and percolation conventions to “permit a purely morphological feature to percolate from the root to the top of the tree”. Even then it does not explain blocking effects of deverbal nominalizations in Russian verbal morphology. It is argued that the generalization can only be stated by a morphological rule of referral.

In [13] it has already been shown how morphological frameworks using rules of referral can be formulated in a finite-state framework. However, we would like to suggest another and much more simple analysis of the facts.

4 Renalysis of the Long Distance Dependency

A simple solution could again use flag diacritics. However, there is a much simpler approach. The crucial assumption in [12] is: “There is no purely phonological restriction which will account for the lack of **razrušánie*, **ukrepl'ánie* [...]”

But let us have a closer look at the data. The first thing to note is that if a verb contains stems with palatalized allomorphs in its paradigm then the deverbal suffix *nie* is always attached to such a stem allomorphs:

- 11 *izgotóvit'* (manufacture-pf.)
izgotól'u (manufacture-1.Sg.pres.pf.)
izgotovlénie (manufacture-N.)
- 12 *provozglasít'* (proclaim-pf.)
provozglašú (proclaim-1.Sg.pres.pf.)
provozglašénie (proclamation-N.)

The next thing to note is that the suffix *nie* is always added to a stem ending either in a or e, even though the stem to which the deverbal nominalization is added might never realize that vowel elsewhere in the paradigm:

- 13 *izgotóvit'* *izgotovlénie*, **izgotovlínie*
ukrepít' *ukreplénie*, **ukreplínie*

Our assumption then is that the vowel preceding *nie* in fact belongs to the nominalizing suffix. It is attached to a stem that does not have an thematic vowel and is realized as *a* after hard consonants and as *e* after soft consonants. The only morphological requirement for the application of the nominalization suffix is to take a special palatalized stem form if there is one in the paradigm. Everything else is governed by phonology:

- 14 /Vnie/ → [enie] / C[soft] _
 /Vnie/ → [anie] elsewhere

This analysis can be stated in a more sophisticated phonological theory. Using underspecified feature structures, one possibility would be to say that the underlying vowel needs only to be specified for [LOW]. A postlexical rule deletes this feature in the context of a preceding soft consonant which is always specified for [HIGH] (see [14]). A vowel with no feature specification at all will per default be as realized as coronal [e]. With only specified for [LOW] after hard consonants, a redundancy rule will realize this vowel as dorsal [a].

There is one slight complication with sibilants in Russian. In Russian all consonants can have soft (with secondary palatalization) and hard variants (without secondary palatalization). The sibilants *š*, *ž* and the dental affricate *ts*, however, do not have both variants but always surface as hard consonants. On the other hand, in certain phonological contexts which are sensitive to the softness of the consonant the hard sibilants behave like soft consonants. This is true for the above mentioned rules of vowel alternation and for similar rules of stress-sensitive vowel neutralization. There is thus evidence that underlyingly they are soft. With that assumption one could explain why the form *provózglašénie* is encountered instead of *provózglašánie*. Again only a detailed phonological analysis will lead to these generalizations. The exact details of the Russian phonological system are quite elegant and straightforward but need not to be copied one by one into the *xfst*-framework. It suffices to know that the generalization to be captured is phonological. With these observations at hand one could now explain the following blocking effect:

- 15 *razrušít'* (destroy-pf.) *razrušénie* (destruction)
razrušát' (destroy-ipf.) **razrušánie*
- 16 *razorít'* (destroy-pf.) *razorénie* (destruction)
razor'át' (destroy-ipf.) **razor'ánie*

The reason why a form like **razoránie* is never encountered as opposed to the form *razorénie* is that there is a simple case of phonological neutralization at work. The underlying stem used for the formation of *razorénie* is /razor'/ which ends in a soft consonant. The vowel of the deverbal suffix undergoes a simple

assimilation – after soft consonants it is fronted and surfaces as [e] whereas elsewhere it surfaces as [a].

This phonological generalization holds also for verbs from the consonantal class which do not have imperfective suffixes but signal the secondary imperfectivization via the whole stem form:

- | | | | | |
|----|-----------------|------------------|------------------|------------------|
| 17 | <i>sobrát'</i> | (collect-pf.) | <i>sobírat'</i> | (collect-ipf.) |
| | <i>sobránie</i> | (collection) | <i>sobiránie</i> | (collecting-N) |
| 18 | <i>výžat'</i> | (wring out-ipf.) | <i>vyžímat'</i> | (wring out-ipf.) |
| | not attested | | <i>vyžimánie</i> | (wringing out-N) |

The phonological generalization also applies to unprefixated verbs:

- | | | | | |
|----|---------------|--------------|----------------|-------------|
| 19 | <i>pisát'</i> | (write-ipf.) | <i>pisánie</i> | (writing-N) |
| | <i>bít'</i> | (beat-ipf.) | <i>bijénie</i> | (beat-N) |

The last piece of evidence comes from semantics. Normally, nominalizations with *nie* formed from perfective verbs do not show a predictable pattern of nominalization type, as pointed out in [12]. It is therefore interesting to note that nominalizations of lexically prefixed verbs where the secondary imperfectivization uses the allomorph *a* almost always have a complex event reading besides a simple event reading or result reading. The explanation is easy in our analysis – since the nominalization of these verbs can have two potential underlying stems (perfective and imperfective) and since the secondary imperfectives show a regular pattern of nominalization type this generalization is preserved independent of the phonological neutralization.

In sum, with another morphological segmentation the effect of blocking turns into a case of phonological neutralization.

5 Conclusion

We have demonstrated how linguistic analyses can be simplified when seen from different perspectives. It remains an interesting but open question whether similar phenomena can also be reanalyzed in this way and what repercussions a special linguistic theory has for engineering aspects.

6 Acknowledgements

I would like to thank Annette Hautli for her inspiring suggestions and her help with the formalization of the analysis. Miriam Butt kindly provided me the opportunity to engage more in finite-state research. Special thanks to Thomas Mayer for discussing the topics with me and for his patient help with the formatting. Finally, I would like to thank two anonymous reviewers for their helpful comments on an earlier version of this paper. Of course, I take the full responsibility for all possible errors.

This research was partially supported by the SFB 471 "Variation und Entwicklung im Lexikon" at the University of Konstanz and university research funds for Aditi Lahiri.

References

1. Beesley, K.R., Karttunen, L.: *Finite State Morphology*. Stanford, California: CSLI Publications (2003)
2. Lehmann, V.: *Aspekt*. In Jachnow, H., ed.: *Handbuch der sprachwissenschaftlichen Russistik und ihrer Grenzdisziplinen*. Harrassowitz, Wiesbaden (1999) 214–242
3. Timberlake, A.: *A Reference Grammar of Russian*. Cambridge University Press (2004)
4. Isacenko, A.: *Die russische Sprache der Gegenwart – Formenlehre*. fourth edn. Max-Hueber-Verlag, München (1995)
5. Spencer, A.: *Morphological Theory*. Blackwell Publishers, Oxford (1998)
6. Breu, W.: Interactions between lexical, temporal and aspectual meanings. *Studies in Language* **18**(1) (1994) 23–44
7. Breu, W.: Zur Position des Slavischen in einer Typologie des Verbalaspekts (Form, Funktion, Ebenenhierarchie und lexikalische Interaktion). In Breu, W., ed.: *Probleme der Interaktion von Lexik und Aspekt (ILA)*. *Linguistische Arbeiten*, Tübingen (2000) 21–54
8. Williams, E.: On the notions 'lexically related' and 'head of a word'. *Linguistic Inquiry* **12** (1981) 245–274
9. Zwicky, A.: Heads. *Journal of Linguistics* (21) (1985) 1–20
10. Beesley, K.R.: Constraining separated morphotactic dependencies in constraining separated morphotactic dependencies in finite-state grammars. In Karttunen, L., Oflazer, K., eds.: *FSMNLP'98. Proceedings of the International Workshop on Finite State Methods in Natural Language Processing*, Ankara, Turkey, Bilkent University (1998)
11. Lahiri, A.: Phonology: Structure, representation, and process. In Wheeldon, L., ed.: *Aspects of language production*. Psychology Press, Hove (2000) 165–225
12. Sadler, L., Spencer, A., Zaretskaya, M.: A morphomic account of a syncretism in Russian deverbal nominalizations. In Booij, G., van Marle, J., eds.: *Yearbook of Morphology*. Kluwer Academic Publishers, Dordrecht (1996) 181–215
13. Karttunen, L.: Computing with realizational morphology. In Gelbukh, A., ed.: *Computational linguistics and intelligent text processing*. Volume 2588 of *Lecture notes in computer science*. Springer, Berlin/Heidelberg (2003) 205–216
14. Lahiri, A., Evers, V.: Palatalization and coronality. In Paradis, C., Prunet, J.F., eds.: *Phonetics and Phonology 2*. Academic Press, Inc. (1991) 79–100