

Morphophonological Challenges for Phase-Based Approaches to Word Formation*

Vera Griбанова
University of California, Santa Cruz

March 14, 2008

0 Introduction

Recent approaches to word formation have productively pursued the idea that syntactic phases can and should be recognized within words (Arad, 2003; Marantz, 2001, 2008). This line of investigation explores the possibility that the morphophonological and morpho-semantic facts that were once accommodated by way of the distinction between lexical and syntactic composition can be treated within syntax alone, with functional heads playing the crucial role of “fixing” both morphophonological and semantic information before further word formation takes place. In this way, the phase boundary replaces the boundary between presyntactic and syntactic composition, simultaneously helping to maintain an important assumption of Minimalism and Distributed Morphology: there is only one generative component of the grammar, and it is the syntactic component. For ease of reference, I call this approach the “phases in words” theory.

The adoption of standard Distributed Morphology assumptions, in conjunction with a particular view of phasehood at the word level, leads to a new and productive way of understanding roots: they are atomic elements, stored as underspecified bundles of semantic and phonological content. Concrete semantic meaning, as well as grammatical category, are assigned only at the point when the root is merged with a functional category in the syntactic derivation. This functional head fixes interpretation of the root, which can in principle give rise to numerous meanings, dependent on the morphosyntactic context in which it appears. Morphophonological patterns are also fixed at the point when the functional head is introduced. Thus, a root cannot be realized independently until it becomes a word via merger of the relevant functional head.

More generally, since this functional category head is a phase, it induces all the processes we typically associate with phases: that is, spell-out and subsequent phonological, syntactic and semantic opacity in further derivation. The introduction of a phase boundary in word formation consequently yields the difference between what was previously understood to be presyntactic composition (now

*Thanks....

corresponding to syntactic composition before the merger of a phase-defining head) and syntactic composition (composition that takes place after the first merger of a functional head). ‘Lower’ word formation from roots exhibits morphophonological and interpretational idiosyncrasy, whereas ‘higher’ word formation is predictable and regular, and may not refer to the internal structure of the root itself. Rather, higher word formation takes as an opaque building block everything below the merged functional head, including not just the root but also any material that was merged prior to the merger of the functional head. This approach yields an automatic distinction between words formed from roots (via the merger of a root with a functional head), as in (1), and words formed from words (via the merger of additional syntactic material above the phase-defining head), as in (2). Words formed from words will be directly semantically dependent on the interpretation that was fixed via the addition of the phase-defining head; words formed from roots can yield unpredictable meanings.



This approach has already been shown to yield productive results in cases for which there is a direct correspondence between semantic and morphophonological idiosyncrasy, e.g. for Hebrew by Arad (2003). In what follows below, however, I argue that the “phases in words” approach is fundamentally not equipped to deal with fairly common (and well-known) cases in which phonological processes and semantic interpretation are mismatched. This is because the notion of “phase” entertained on the approach described above makes too stringent predictions about the correspondence between semantic and phonological “fixedness” after the introduction of the relevant functional head.

Below I examine two cases in which phonological and semantic properties of morphological objects are mismatched. In the first case — Russian prefixation — the phonological and morphosemantic evidence point to two different conclusions (§1). Drawing on recent work on the morphosemantics and morphosyntax of Russian prefixation (Svenonius, 2004a,b; Babko-Malaya, 2003), we can demonstrate that the semantic contribution of the prefix is irregular, and thus on a “phases in words” view should be treated as ‘lower’ material. However, the relevant phonological evidence leads us to a somewhat contradictory conclusion: while several processes reflect a word-external status for prefixes, the addition of syntactic functional material necessitates reference to, and in some cases changes, the phonological profile of the root.

In the second case — Bulgarian definiteness marking — the definite article contributes predictable semantics, and is merged in D according to many standard accounts (Embick and Noyer, 2001). The article is suffixed (post-syntactically, on the view proposed in Embick and Noyer 2001) to the first adjective or noun in the DP, but exhibits word-internal phonology, even though it is merged only after the noun and adjectives in the DP have been formed as words (§2). In both case studies, the evidence presents a clear challenge for accounts in which the phase boundary is invoked to fix semantic, and also morphophonological, properties of the word.

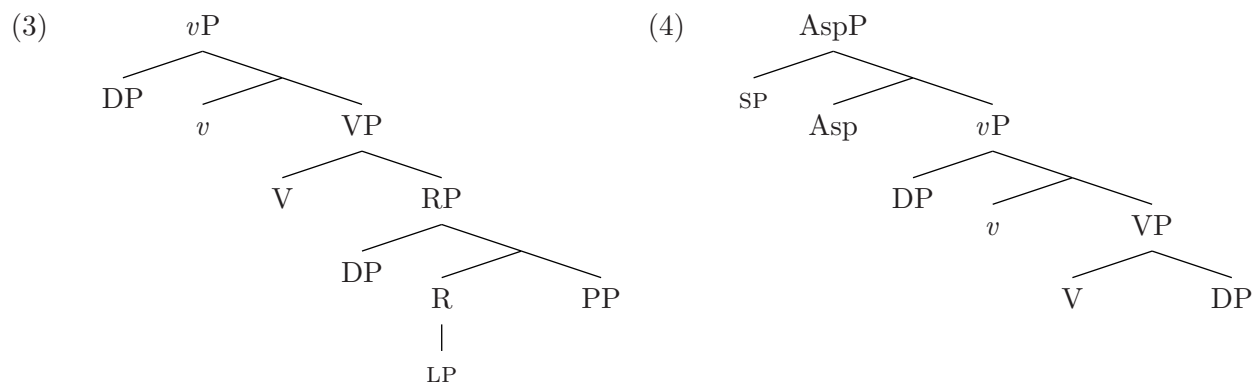
1 Russian Prefixation

Arad (2003) cites Russian prefixation paradigms as an example of the “fixing” of semantic content after a certain point in a word’s derivation: “the root may acquire many interpretations in combinations with different prefixes, but once its meaning has been assigned, it is retained throughout all further derivation” (Arad, 2003, 775). It is true that Russian roots acquire different interpretations when joined with certain prefixes, but in order to evaluate the extent to which this evidence support’s Arad’s overall proposal, we need some more empirical background.

1.1 The Prefixes Subdivide

There is substantial evidence to suggest that Russian prefixes, which in the verbal domain are widely associated with perfectivity, fall into two semantically and syntactically distinct, homophonous subclasses: lexical (LP) and superlexical (SP) (Isačenko, 1960; Svenonius, 2004a,b; Babko-Malaya, 2003, *inter alia*). SP contribute predictable (quantificational or adverbial) meanings, they are incompatible with secondary imperfectivization, they stack outside LP in prefix stacking, they can co-occur, and they do not change the argument structure of the verb. By contrast, LP contribute idiosyncratic meanings, are compatible with secondary imperfectivization, appear closest to the verbal stem in prefix stacking, cannot co-occur, and can change the verb’s argument structure (see Svenonius 2004b for a thorough investigation of the various properties that distinguish the two classes).

A current approach to distinguishing between the two prefixal classes (Svenonius, 2004a,b) has accounted for their differing characteristics structurally. The analysis assumes a currently standard views of the architecture of the verbal complex, with a transparent correspondence between syntactic and morphological structure. LP attach within *vP* as R[esultative] heads of small clauses (3), parallel to proposals for the German verb-particle construction (Ramchand and Svenonius, 2002). SP attach above *vP* (4).¹



Recent work in this area has corroborated the idea that SP are in fact merged in the functional

¹Following Svenonius 2004c, I assume tentatively that *v* hosts the verb’s theme vowel, which determines numerous properties of the verb (among them argument structure, allomorphic selection, etc.).

layer of the Russian clause, in part by invoking evidence from V-stranding VPE (Griбанова, 2008). As for LP, there is little to no evidence that it forms the head of a constituent RP in Russian, since traditional tests for constituency (Pseudogapping, Coordination, and Right Node Raising) all yield inconclusive or even negative results. In general, many of the properties of LP, in particular its contribution to the verb’s argument structure and the fact that it facilitates idiosyncratic meaning, suggest that LP might be best analyzed along the lines of Arad’s proposal — that is, as joined with the root before the addition of the category-defining functional head, rather than as the head of a separate projection that is complement to V.

1.2 LP According to Arad 2003

If we pursue this line of investigation a bit further, we will see that it yields some interesting and even elegant results. It so happens that the nominal prefix-root combinations are in general typical of LP, and not of SP. For example, consider the Root [bOr],² which combines with numerous LP in both the nominal and verbal domain (notice that there is no extra ‘nominal’ morphology below, only the prefix and the root).

LP	Verb	Noun
<i>razO-</i>	razobrat ^j , razbirat ^j ‘take apart’ (PFV/IMPF)	razbor ‘analysis’
<i>pri-</i>	pibrat ^j , pribirat ^j ‘tidy up’ (PFV/IMPF)	pribor ‘device’
<i>na-</i>	nabrat ^j , nabirat ^j ‘accumulate, dial up’ (PFV/IMPF)	nabor ‘set (of objects)’
<i>vi-</i>	vibrat ^j , vibirat ^j ‘choose’ (PFV/IMPF)	vibor ‘choice’
<i>podO-</i>	podobrat ^j , podbirat ^j ‘pick up’ (PFV/IMPF)	podbor ‘collection’
<i>sO-</i>	sobrat ^j , sobirat ^j ‘collect’ (PFV/IMPF)	sbor ‘gathering’
<i>otO-</i>	otobrat ^j , otbirat ^j ‘choose’ (out of a group) (PFV/IMPF)	otbor ‘selection’

Figure 1: Verbal and nominal LP-stem combinations

It may be useful to further note that the nominals in the figure above appear without a ‘theme’ vowel, which by assumption is provided by the functional head *v*. This vowel is therefore predicted to be absent in the nominal versions of LP-root combinations.

Curiously, there *is not*, contrary to Arad’s claim, always a direct semantic correspondence between related nominal and verbal LP-root combinations. For example, while *razobrat^j* ‘take apart’ and *razbor* ‘analysis’ are clearly semantically related, the same cannot be said of *pibrat^j* ‘tidy up’ and *pribor* ‘device’. Such facts can still be made to fall out of a phase-based view of word formation, if *razobrat^j* is a denominal verb, formed on the basis of the nominal *razbor* (and therefore semantically dependent). By contrast, *pibrat^j* and *pribor* would be formed independently of one another, with the same combination of [pri] and [bOr] yielding varied meanings when dominated by either an *n* head or a *v* head.

With respect to LP’s semantic properties, then, we can with some confidence conclude that the approach outlined in Arad 2003 would probably fare very well. Turning to the phonological facts,

²Capitalized vowels represent jers, which are present underlyingly and alternate with zero in certain morphological contexts. Note that underlying jers are present both root-internally and prefix-finally.

however, this same approach is faced with some serious challenges.

1.3 Morphophonological Difficulties

Perhaps the first sign that a phase-based approach will face challenges in accounting for the Russian data is that SP and LP are homophonous, and exhibit no differences at all with respect to phonological processes such as jer realization, word-final devoicing, palatalization, and hiatus resolution (for extensive discussion, I refer the reader to Matushansky 2002 and Griбанова 2007). If SP are indeed merged in a functional Asp projection, it seems surprising that they would be identical phonologically to LP, which on the “phases in words” approach would be merged before the first category-defining head. In fact, certain pieces of this evidence point to the conclusion that both groups of prefix are, for the purposes of phonology, added at a relatively late stage of morphological derivation, in both nominal and verbal domains. Consider the palatalization pattern for prefixes, which mimics the pattern we find across word-boundaries, rather than word-internally.

Word-internally and across certain stem-suffix boundaries, consonants palatalize to conform in backness to following high and mid front vowels (as in (5)).³

- (5) /obide/ → [obʲidʲe] ‘offense.DAT’
 /alʲt + ist/ → [alʲtʲist] ‘viola player’

Across word boundaries, however, the strategy for resolving backness mismatches in similar environments changes: instead of palatalizing word-final (underlyingly unpalatalized) consonants, the backness quality of following underlyingly high front vowels is altered (I follow the traditional literature in calling this process *retraction*). This results in a velarized consonant followed by a [+back] vowel (6).

- (6) /ugol ivana/ → [ugolʲivana] (*ugolʲivana) ‘Ivan’s corner’
 /sad iriny/ → [satʲiriny] (*satʲiriny) ‘Irina’s garden’⁴

Both LP and SP follow the patterns in (6), with retraction being the strategy for backness mismatch resolution.

- (7) /ot + iskatʲ/ → [otʲiskatʲ] (*otʲiskatʲ) ‘find.INF’

The crucial property of the process described in (6) is that it operates only across certain kinds of boundaries (word boundaries at least). Given that, it is unsurprising that it should take place between SP and the root, since they are separated, on the proposal under discussion here, by the phase boundary defined by *v*. What is unexpected, however, is that retraction also takes place

³In fact, the pattern is more complicated than what is described here, in particular with respect to palatalization across stem-suffix boundaries. A more detailed discussion appears in Blumenfeld 2003.

⁴Though word-final devoicing has no direct relevance to the discussion here, I include it in the transcriptions.

between LP and the root. On the view being considered here, LP combines with the root within the spell-out domain of v and the presence of the kind of major boundary that would trigger retraction is unexpected.

On a theory that designates the category-defining head as the boundary for morphophonological as well as semantic idiosyncracies, we might expect LP to differ from SP both semantically and phonologically; but the evidence above suggests that this is not the case. LP behave, at least with respect to certain phonological processes, as if there is some sort of prosodic boundary between LP and the root; we might not expect this on a view in which the LP and the root are both within the spell-out domain introduced by the category-defining head.

A much more problematic piece of phonological evidence comes from the somewhat complicated pattern of jer realization found in prefixed verbs. Recall that in many cases both the root and the prefix contain an underlying jer vowel that is either realized or deleted, depending on the morphophonological context in which it appears. Most contemporary views of jer phonology posit that a jer can be shown to be present underlyingly if morphologically related variants of a word appear with and without the relevant vowel ([o] or [e]). It is usually assumed that the phonological computation either deletes the jer in the surface form or realizes it as a full vowel (Yearley, 1995).

The problem we are discussing is not new, but in the context of the “phases in words” theory of word formation, it gains a new relevance. Perhaps we can start by illustrating it from the point of view of Pesetsky (1979), who described a mismatch between the phonological and semantic bracketing of prefixed verbs. He noted that the phonology of prefixed verbs in which both the root and the prefix contain a jer suggests the bracketing in (8a), but the semantics of these same verbs suggests the bracketing in (8b).

- (8) a. ⟨prefix ⟨⟨root⟩ inflection⟩⟩
 b. ⟨⟨prefix ⟨root⟩⟩ inflection⟩

The concrete case at issue here is one in which inflectional information triggers jer realization in the root or in the prefix, as in (9). Here we have the feminine and masculine third person paradigm of the verb ‘to burn up’, composed of the verb root /žOg/ and the LP sO -, both with a jer as a nucleus:

- (9) /sO + žOg + la/ → [sožgla] ‘she burned (something) up’
 /sO + žOg/ → [sžog] ‘he burned (something) up’

A few points should be made here. First, given the framework we are exploring here, we do not expect the addition of verbal inflection (presumably housed in T) to have any impact on the realization or deletion of jers below the phase level. But it appears that the type of inflection can trigger jer realization in the prefix, which should be out of the question, since the prefix is by hypothesis ‘trapped’ in the lower phase. Thus, the old problem gains yet more relevance: given the semantic facts, what we expect is, roughly speaking, the bracketing in (8b), with the prefix and root dominated by a little head v . If that were the case, the internal structure (underlying jers included) of the prefix and root would be phonologically inaccessible to inflection, resulting in the

ungrammatical form in (10); yet, inflectional features have the potential to trigger realization of underlying jers in the root and the prefix.

- (10) $\langle\langle sO \langle žOg \rangle \rangle M \rangle \rightarrow [sžog + 0] \rightarrow [sžog]$ (masculine)
 $\langle\langle sO \langle žOg \rangle \rangle F \rangle \rightarrow [sžog + la] \rightarrow *[sžogla]$ (feminine)

A second piece of evidence points to essentially the same problematic conclusion: higher, functional material, when added to the verbal complex, can trigger jer realization or deletion inside the root, long after the root should have become phonologically opaque to new material in the syntactic derivation. For example, the addition of the secondary imperfective suffix — which has been analyzed as instantiating an Asp head between T and *vP* (Jabłońska, 2004; Svenonius, 2004a,b) — to the verbal complex can trigger jer realization prefix-finally and root-internally.⁵

- (11) a. $podobrat^j$ ‘pick up.PFV’ / $podbirat^j$ ‘pick up.2IMPF’
 $otobrat^j$ ‘pick out.PFV’ / $otbirat^j$ ‘pick out.2IMPF’
 b. $podognat^j$ ‘chase up (to).PFV’ / $podgonjat^j$ ‘chase up (to).2IMPF’
 $otognat^j$ ‘chase away.PFV’ / $otgonjat^j$ ‘chase away.2IMPF’
 c. $soslat^j$ ‘send away.PFV’ / $ssilat^j$ ‘send away.2IMPF’
 $otoslat^j$ ‘send away.PFV’ / $otsilat^j$ ‘send away.2IMPF’
 d. $razorvat^j$ ‘tear apart.PFV’ / $razrivat^j$ ‘tear apart.2IMPF’
 $oborvat^j$ ‘tear off.PFV’ / $obrivat^j$ ‘tear off.2IMPF’

Since 2IMPF in this case is expressed by a vowel alternation in the verbal stem, this alternation must also be the source of the jer alternation found prefix-finally. This pattern indicates that the prefix and 2IMPF undergo phonological evaluation simultaneously. If the prefix and root are dominated by a phase-defining head, however, neither root-internal nor prefix-final jer realization should be triggered by the addition of 2IMPF to the verbal complex.

Thus, while the semantic facts are nicely accounted for by analyzing the prefix-root sequence as combining before the addition of a category-defining head, in the rich domain of phonological processes pertaining to Russian prefixation, the account consistently predicts incorrect results.

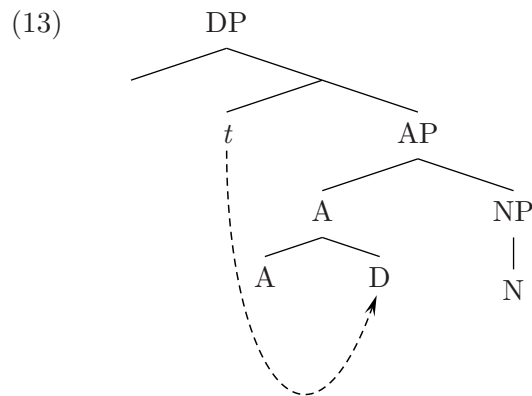
2 Bulgarian Definiteness

A second example of a semantics-phonology mismatch comes from Bulgarian definiteness marking (DEF), which surfaces as a suffixed article on the first noun or adjective in a DP, achieving a sort of second-place effect within the constituent.

⁵In a number of the cases below the root-internal jer undergoes a further phonological change, so that it is realized as [i/i] instead of [o] or [e]. It is standardly accepted that these are still jers, however, because in most cases other morphological variants of the same root can be found in which the jer vowel surfaces unaltered in quality.

- (12) a. kniga-ta
book-DEF
‘the book’
- b. interesna-ta kniga
interesting-DEF book
‘the interesting book’
- c. xubava-ta interesna kniga
nice-DEF interesting book
‘the nice interesting book’
- d. *xubava interesna-ta kniga
nice interesting-DEF book

A standard Distributed Morphology account of DEF takes it to instantiate a D head, which lowers to its ultimate position in the DP via a postsyntactic *Lowering* operation (Embick and Noyer, 2001).⁶



It is almost trivial to point out that DEF contributes no semantically idiosyncratic properties, and this is reflected in Embick and Noyer’s (2001) analysis of it as a functional element heading DP. Yet the placement of DEF, and its morphophonological properties, seem to reflect a status that the “phases in words” theory would predict is impossible. DEF in Bulgarian is realized as a suffix, and correspondingly exhibits word-internal behavior with respect to phonological rules like word-final devoicing, which is bled when DEF is attached.

- (14) a. bratovčed → [bratovčɛt] ‘cousin’
b. muž → [maš] ‘husband’
- (15) a. bratovčed-ət → [bratovčɛdət] ‘the cousin’
b. muž-ot → [mažot] ‘the husband’

⁶I adopt this particular account because it is proposed in the Distributed Morphology framework, and therefore should in principle be consistent with the assumptions of the “phases in words” theory. However, it should be noted that the Bulgarian data present challenges to the “phases in words” theory independent of this particular analysis.

Other prosodically dependent elements, such as possessive clitics, do not participate in the same process. In (16b), word-final devoicing still occurs, despite the addition of the clitic.

- (16) a. mǎž [mǎš] ‘husband’
 b. mǎž ì [mǎši] ‘her husband’

If roots must be dominated by a category-defining head before they can become words, then in both the case of adjectival and nominal definite suffixation, the suffix must attach to an element that should (on these assumptions) be already opaque. That is, the noun or adjective that hosts DEF will have achieved word-hood before DEF is lowered postsyntactically from D, and therefore we expect DEF to behave just like other prosodically dependent clitics (e.g. as in ((16))).

A similar conclusion can be drawn from the fact that certain kinship terms (but not all: see 17d)) forbid the realization of DEF in certain (otherwise definite) configurations (cf. 17a,b with 17c).

- (17) a. majka mu
 mother his.CL
 ‘his mother’
 b. *majka-ta mu
 mother-DEF his.CL
 c. xubava-ta mu majka
 pretty-DEF his.CL mother
 ‘his pretty mother’
 d. djado-to mu
 grandfather-DEF his.CL
 ‘his grandfather’

With the exception of certain kinship terms, possessive clitics appear adjacent to DEF (see Franks 2001). By extension, we can be reasonably sure that the kinship term in (17a) appears where DEF would normally be required, especially given (17c). The illicitness of overt DEF inflection in (17b) looks conspicuously like a lexical gap, and these sorts of morphological idiosyncrasies are characteristic, according to the “phases in words” theory, of material that is merged before the addition of the category-defining functional head.

We thus have another case of semantic and phonological mismatch, in which the Bulgarian definite article behaves phonologically like it is part of the “lower” domain, when most analyses place it in a “higher” domain in accordance with its semantic characteristics.

3 Remarks

The “phases in words” theory — proposed by Marantz (2001, 2008) and adopted by Arad (2003) for a very successful exploration of Hebrew word formation — makes strong predictions with respect

to cyclic phonological effects and semantic dependence in morphological word formation. But this result is advantageous *only* in cases where evidence from these two domains matches. The theory is conversely poorly equipped to deal with cases of semantics-phonology mismatch of the sort described in §1 and §2. Without a significant attempt to accommodate semantics-phonology mismatches, “phases in words” makes predictions that are not borne out by the facts.

References

- Arad, Maya. 2003. Locality constraints on the interpretation of roots. *Natural Language and Linguistic Theory* 21:737–778.
- Babko-Malaya, Olga. 2003. Perfectivity and prefixation in Russian. *Journal of Slavic Linguistics* 11(1):5–36.
- Blumenfeld, Lev. 2003. Russian palatalization in Stratal OT: Morphology and [back]. In W. Browne, J.-Y. Kim, B. H. Partee, and R. Rothstein (Eds.), *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics 11*, 141–158, University of Massachusetts, Amherst. Michigan Slavic Publications.
- Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.
- Griбанова, Vera. 2007. Russian prefixes, prepositions and palatalization in Stratal OT. In *Proceedings of the West Coast Conference on Formal Linguistics 26*, Somerville. Cascadilla Press.
- Griбанова, Vera. 2008. Distinguishing modes of composition: The morphosyntax of Russian verbal prefixation. Unpublished ms., UCSC.
- Isačenko, Aleksandr V. 1960. *Grammatičeskij Stroj Russkogo Jazyka. Morfologija. Čast' Vtoraja*. Bratislava: Vydavatel'stvo Slovenskej Akademie.
- Jabłońska, Patrycja. 2004. When the prefixes meet the suffixes. *Nordlyd* 32(2):363–401.
- Marantz, Alec. 2001. Words. In *Proceedings of WCCFL 20*, University of Southern California.
- Marantz, Alec. 2008. Phases and words. Unpublished ms., NYU.
- Matushansky, Ora. 2002. On formal identity of Russian prefixes and prepositions. *MIT Working Papers in Linguistics* 42:217–253.
- Pesetsky, David. 1979. Russian morphology and Lexical Phonology. Unpublished ms., MIT.
- Ramchand, Gillian, and Peter Svenonius. 2002. The lexical syntax and lexical semantics of the verb-particle construction. In L. Mikkelsen and C. Potts (Eds.), *Proceedings of WCCFL 21*, 387–400, Somerville, MA. Cascadilla Press.
- Svenonius, Peter. 2004a. Russian prefixes are phrasal. *Formal Descriptions of Slavic Languages* 5.
- Svenonius, Peter. 2004b. Slavic prefixes inside and outside VP. *Nordlyd* 32(2):205–253.
- Svenonius, Peter. 2004c. Slavic prefixes and morphology: An introduction to the Nordlyd volume. *Nordlyd* 32(2):177–204.

Yearley, Jennifer. 1995. Jer vowels in Russian. In J. N. Beckman, L. W. Dickey, and S. Urbanczyk (Eds.), *Papers in Optimality Theory*, 533–571. University of Massachusetts, Amherst: GLSA.

Vera Gribanova
Department of Linguistics
University of California Santa Cruz
gribanov@ucsc.edu
<http://people.ucsc.edu/~gribanov>