University of Pennsylvania Working Papers in Linguistics

Volume 19
Issue 1 *Proceedings of the 36th Annual Penn Linguistics Colloquium*

Article 19

1-28-2013

Out of Order?: Russian Prefixes, Complexity-based Ordering and Acyclicity

Robert Reynolds
The Ohio State University, reynolds.531@osu.edu

Out of Order?: Russian Prefixes, Complexity-based Ordering and Acyclicity

Abstract
There is a longstanding debate about how to appropriately model the combinability of affixes, especially English suffixes. One widely accepted principle is the notion of so-called selectional restrictions, i.e. grammatical requirements of particular affixes. For example, the suffix -ness can only combine with adjectival bases. Hay (2002) proposed a psycholinguistic approach to affix ordering now known as Complexity-Based Ordering (CBO), which claims that affix order is determined by the parsability of the affixes, i.e. more separable affixes can appear only outside of less separable affixes. Hay shows that this principle accounts for why many grammatical affix combinations are unattested. CBO has since been supported by research of derivational affixes (English prefixes, English suffixes and Russian suffixes). However, as a processing model, CBO should apply very broadly, and in this paper, I discuss some difficulties of reconciling CBO with inflectional affixes. I also examine combinations of Russian prefixes - which have some properties typical of inflection - and show, surprisingly, that they can be ordered with a significantly low number of cycles, as CBO predicts. I discuss alternatives to CBO that explain this phenomenon, and suggest future research to distinguish them.

Out of Order?: Russian Prefixes, Complexity-based Ordering and Acyclicity

Robert Reynolds*

1 Introduction

In natural language, only a small proportion of affix combinations are actually attested. For example, English has two nominalizing suffixes, *-ity* and *-ness*, of which generally only one appears on any given base word: the base word *atomic* can combine with *-ity*, but not *-ness*, but the base word *atomless* cannot combine with *-ity*, and can combine with *-ness*. There is a longstanding debate about what motivates such limitations. For example, one widely accepted principle is that particular affixes can select for a candidate base's morphological, phonological, semantic, or syntactic properties. For example, a selectional restriction of the Russian agentive suffix *-el'* is that it can combine only with bases that are verbs (e.g., *učit'* 'teach' ~ *učitel'* 'teacher', but *sneg* 'snow.NOUN' ~ **snegel'*). While some affixes' selectional restrictions are simple, complex restrictions also exist. For example, the English suffix *-en* selects for monosyllabic adjectival bases that end in an obstruent (e.g., *black ~ blacken*, but *tall ~ *tallen* or *circumspect ~ *circumspecten*). The operation of selectional restrictions is at least tacitly implied in most theories of affix stacking, and some researchers have even claimed that *only* selectional restrictions are needed to account for attested affix combinations (e.g., Fabb 1988, Plag 1996). However, despite the descriptive power of selectional restrictions, it has been argued that selectional restrictions alone are too permissive, allowing for many more combinations than are actually attested (Hay 2003).

Two major approaches have emerged that attempt to be more restrictive than selectional restrictions alone. First, level-ordering or stratum-oriented models (e.g., Kiparsky 1982, Kiparsky 1985) claim that affixes belong to particular layers, or strata, and that the strata are ordered, so that affixes of the first stratum always occur inside of affixes of other strata. This approach has been criticized both for inaccurate predictions, and for not offering a good explanation of why language learners should acquire the distinction between strata.

As an alternative, Hay (2003) proposed an approach that has come to be known as Complexity-based Ordering (CBO), which claims that affixal combinations are restricted by cognitive processing restrictions, and in particular, the parsability of the affix. According to CBO, affixes that are more easily parsed should occur outside of affixes that are less easily parsed, since violations of this order complicate online processing of complex words.

Recent studies in support of CBO have investigated English prefixes (Zirkel 2010), English suffixes (Hay 2002, Hay and Plag 2004, Plag and Baayen 2009) and Russian suffixes (Parker and Sims, in progress). All these studies have investigated affixes that are canonically derivational. It remains to be seen whether affixes with inflectional properties should behave similarly. The present study extends the research of affix ordering and CBO by testing whether Russian prefixes exhibit the acyclic affix ordering that CBO predicts. Russian prefixes are distinct from affixes mentioned above in that some of their properties are more typical of inflection than derivation.

In Section 2, I give an overview of CBO, including its theoretical assumptions. In Section 3, I outline some affix-ordering tendencies of inflectional morphology, and explore whether CBO is likely to predict those tendencies. In Section 4, I outline properties of Russian verbal prefixes that are typical of inflection, and discuss the extent to which CBO applies to affixes with these properties. Sections 5 and 6 contain the methodology and results showing that Russian prefixes are generally hierarchical. In Section 7, I discuss what this result means for CBO, and briefly explore alternative explanations of acyclic affix stacking. Finally, in Section 8, I summarize my conclusions.

2 Complexity-based Ordering (CBO)

2.1 Overview of CBO

CBO is based on a parallel-dual route model of morphological processing. Parallel dual-route

models claim that in perception there is competition between parsing and direct access of complex words in the lexicon (Schreuder and Baayen 1997, Baayen and Schreuder 1999, Bertram et al. 2000, Hay and Baayen 2001, Burani 2003, Clahsen et al. 2003, Jarvikivi 2006). In other words, when a word is perceived, the two possible means of access compete, i.e., whole word access versus decomposition of component parts. The perceived word is ultimately processed by whichever method is faster. According to this position, a morpheme is more likely to be parsed during lexical access if the boundary between itself and its base is strong, i.e., clearly perceived as a boundary. The parsability of an affix from its base is hypothesized to be the result of many factors, including productivity, semantic transparency, relative frequency of base and derivative, affix length, and junctural phonotactics (Ibid., Hay 2002). [1] Some of these factors are generalizations specific to each affix (e.g., productivity), and are therefore invariant from word to word. Some represent a combination of properties of the base and affix (e.g., junctural phonotactics), and still others represent the relationship between the base word and derivative (e.g., semantic transparency, relative frequency). As each factor varies, so does the likelihood of being parsed.

CBO is characterized by the premise that the processing of complex words is simplified by an ordering which puts more easily parsable affixes outside of less easily parsable affixes.

> While some affixes basically tolerate no internal structure, others will tolerate structure to some minimum degree. The degree of internal structure tolerated by an affix is not determined by selectional restrictions, but, rather, by how much structure that affix itself creates. Phrased in terms of processing, an affix that can be easily parsed out should not occur inside an affix that cannot. (Hay 2002:527–528)

In a general sense, CBO predicts that a rank order of affixes, based on average parsability of each affix, will correlate with the order in which those affixes stack. [2] For example, if the hypothetical prefix *A-* is more easily parsable than prefix *B-*, and prefix *B-* is more easily parsable than prefix *C-*, we may rank them as follows: *A-* > *B-* > *C-* > …->*Z*. Given this generalization, CBO claims that combinations *AAA-*, *ABC-*, or *AMZ-* are more easily parsed than combinations *ACB-*, *CBA-*, or *ZMA-* and are therefore more likely to occur in natural language, e.g., *ACB-* is more likely to occur than *AZB-*.

Empirical studies of English affixes have supported the predictions of CBO (Hay and Plag 2004, Plag and Baayen 2009, Zirkel 2010). These studies show that the attested combinations of English affixes can be hierarchically ordered, and that the rank order of affixes correlates significantly with measures of affix parsability. For example, Plag and Baayen (2009) investigated a set of 30 English suffixes. They present the optimal affix ordering as an adjacency matrix (Figure 1), in which a "1" indicates attestation of the affix pair's combination. For example, words with the affix combination *-th-wise*, like *lengthwise*, are represented by the '1' in the top-right corner of the matrix.

In this particular adjacency matrix, the order of the affixes has been arranged so that it is maximally acyclic, or in other words, has the fewest possible attestations below the line. As is evident, there are some exceptions to the optimal hierarchy, but they are mostly shown to be extremely rare and/or old attestations, such as *alchemister* and *evangelistary*. So far, the few studies on CBO that have been published show that its predictions are generally accurate, i.e., that affixes can be ranked with significant acyclicity, and that their ranks correlate with hypothesized factors of parsability.

[1] Junctural phonotactics are a measure of the likelihood of a given phone pair occurring word-internally or spanning a word boundary. Pairs that are more likely to occur spanning a word boundary are said to indicate a strong boundary. For example, *warm-th* is more likely to be parsed than *heal-th* because the boundary [mθ] is stronger than [lθ], i.e., [mθ] is perceived as more likely to be a word boundary than [lθ] (Hay 2002:530).

[2] Strictly speaking, affixes may appear in different orders based on parsability factors that can vary from word to word. In other words, if we abstract a hierarchy based on *average* affix parsability, a hierarchy should emerge, but we should not be surprised to find some violations to the more generalized rank order.

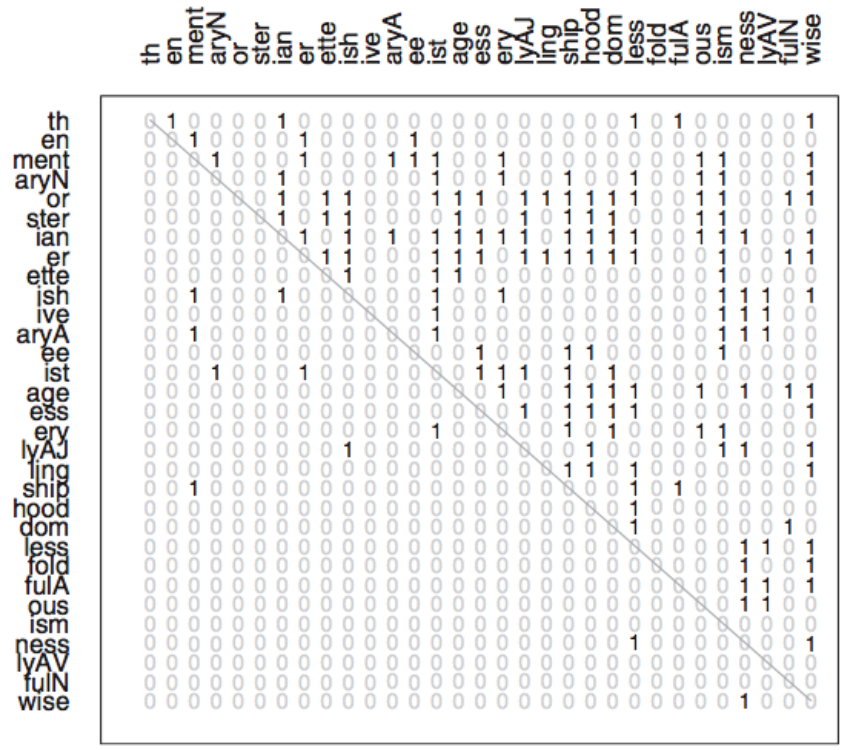

Figure 1: Adjacency matrix of English suffix ordering hierarchy (Plag and Baayen 2009:122).

2.2 Theoretical Assumptions of CBO

Because it is based on the parsability of form, CBO is forced to make some theoretical assumptions about how to represent words and affixes. More specifically, CBO implicitly assumes a
lexical-incremental model of morphology (Stump 2001:2), i.e., one in which affixes are meaning-bearing lexical entries. This brings with it the assumption that morphology is basically agglutinative, i.e., that adding an affix's form adds its inherent function or meaning. However, among other scholars, Beard (e.g., 1987, 1995) has documented many cases in which the form and meaning of morphemes do not match. Form-meaning parallels are perhaps more characteristic of derivational morphology, and since previous studies addressing the CBO hypothesis have focused on derivational phenomena, it has been easy (and probably reasonable) for this work to ignore the issues raised by Beard.

Since CBO was first formulated, its applicability to inflectional morphology has been left an open question (Hay 2003:25–26).[3] A processing account of word structure should apply very broadly, but canonical inflection is substanstially different from canonical derivation. In the following section, I discuss the problems of CBO's interpretability in the context of inflection.

3 CBO and Inflection

3.1 Affix Ordering in Inflection

Affix ordering in canonical inflection is determined by morphosyntactic function. In other words,

[3] Some scholars reject the possibility of a clear differentiation between inflection and derivation (for discussion, see Bybee 1985, Stump 1998:§2.2, and Booij 1998:§3–§4). As will be discussed in Section 4, Russian prefixes are themselves outstanding evidence that the task of categorizing affixes is fraught with complications.

affixes of a given function always appear in the same position relative to other parts of the word. Swahili verbal prefixes offer a good example of this. Swahili is a pro-drop language with person-number agreement for both subject and object (Deen 2001, 2006):

(1) Swahili verbal prefixes (modified from Stump 1992:129, 138)
   a. *ni-* *ta-* *wa-* *penda*
      1SG.SBJ - FUT - 3PL.OBJ - love
      'I will love them'
   b. *wa-* *ta-* *m-* *penda*
      3PL.SBJ - FUT - 3SG.OBJ - love
      'they will love him'
   c. *wa-* *ta-* *wa-* *penda*
      3PL.SBJ - FUT - 3PL.OBJ - love
      'they will love them'

In example (1), we see that each morphosyntactic function is expressed in a particular position relative to other prefixes: Subject - Tense - Object - Root. The affix *-wa-* expresses third person plural agreement for object (1a), subject (1b), and both subject and object (1c). Furthermore, these same examples illustrate that the very same person-number formant can appear in different positions, sometimes even appearing more than once in the same word. In summary, inflectional affixes *of the same morphosyntactic function* consistently appear in the same position relative to other affixes.[4]

3.2 CBO and Inflection

The potential for incompatibility of inflectional affixes and CBO stems from the facts that (i) CBO is rooted in the parsability of individual affixes' *form*, as stored in the lexicon, and (ii) the position of inflectional morphs is rooted in morphosyntactic *function*. For CBO to be accurate, the parsabilities of the affixes of each morphosyntactic function would have to be discrete from the parsabilities of all the affixes of neighboring morphosyntactic functions. Although such a state is conceivable, there is no reason to believe that such a coincidence should be so universally constant across languages. Indeed, revisiting the Swahili data presented in (1), the same formant can appear in different positions even within the same word. This suggests a parsability distribution in which prefixes of various functions overlap in their parsability. To the extent that CBO relies on the *form* of individual inflectional affixes to measure their parsability, and thus predict their ordering, it allows far more affix combinations in canonical inflection than are attested.

4 Russian Verbal Prefixes

Russian prefixes play an essential role in word formation, yet share important properties of canonical inflection. All Russian verbs express one of two aspects, imperfective or perfective, so that most verbs are paired with an aspectual counterpart that shares the same lexical meaning. For example, the English verb *to read* corresponds to the Russian aspectual pair *čitat'*$_{\text{IMPF}}$ and *pročitat'*$_{\text{PF}}$. For many verbs, the imperfective and perfective verbs are differentiated by the presence or absence of a prefix: the unprefixed, simplex verb is imperfective, and the prefixed verb, perfective. Although prefixed verbs usually carry the exact same lexical meaning as their imperfective counterparts, some prefixed verbs have a slightly different meaning than their base. Usually the difference in meanings is predictable, especially in newly coined or infrequent words. For example, in the word *perečitat'*$_{\text{PF}}$ 'reread', the prefix *pere-* carries the adverbial meaning 'again'. However, we also find that words – especially those with relatively high frequency – sometimes have unpredictable, noncompositional meanings, such as *perežit'*$_{\text{PF}}$ 'to worry' (cf. *žit'*$_{\text{PF}}$ 'to live') or *perevarit'* 'to digest' (cf. *varit'* 'to boil').

Despite their role in word formation, Russian prefixes have some properties typical of

[4] There are uncommon instances (e.g., Fula, Quechua, and Swazi) of functions reversing their order when particular combinations of morphosyntactic values are expressed (Stump 1994:164–166).

inflection. For example, it is widely claimed that Russian verbs can have a maximum of two prefixes (e.g., Istratkova 2005, Romanova 2005). This stacking limitation is generally very robust, with only rare exceptions.[5] Such a limitation is typical of inflection, since layered morphology (typical of derivation) allows extensive stacking, as in *com-part-ment-al-iz-ation*. Further, as I will show, the meaning of a prefix is dependent on its position relative to other prefixes, much like the Swahili agreement prefixes in (1).

Beginning with Isačenko (1960), scholars have articulated two distinct categories of Russian verbal prefixes: lexical (also referred to as internal), and superlexical (also referred to as external) (Babko-Malaya 2003, Svenonius 2004, Istratkova 2005, Richardson 2007). Many of these studies are based on the Mirror Principle, which is that morphological derivations must directly reflect syntactic derivations (and vice versa). As their names imply, internal (lexical) prefixes, if present, are VP-internal, and external (superlexical) prefixes, if present, are VP external.

1. Internal and external prefix stacking
   a) [external prefix] + [internal prefix] + [root] + [suffixes]
   b) *[internal prefix] + [external prefix] + [root] + [suffixes]
   c) *[internal prefix] + [internal prefix] + [root] + [suffixes]
   d) [external prefix] + [external prefix] + [root] + [suffixes]
2. Relative ordering of external and internal prefixes:
   a) govorit'$_{\text{IMPF}}$ 'to speak'
   b) *u*-govorit'$_{\text{PF}}$ 'to persuade' (internal prefix)
   c) pere-*u-govorit'*$_{\text{PF}}$ 'to persuade again' (external + internal prefix)
   d) **u-pere-govorit'*$_{\text{PF}}$ (internal + external prefix)

Svenonius (2004:193–195) provides lists of both kinds of prefix, as shown in Table 1.[6] Note that phonologically, superlexical prefixes are a subset of lexical prefixes (Ramchand 2008:1694). Table 1 shows that the prefixes' meanings are conditioned by their position, unlike typical word-forming morphemes. For example, in the combination *pere-za-*, the prefix *pere-* cannot mean 'across' because it is in the external position.

In summary, there are three properties that Russian verbal prefixes share with canonical inflection. First, the meaning of the prefixes is dependent on their position. Second, the type of information expressed in each position is qualitatively different. Internal affixes express direction, and external affixes express manner or degree. Third, there is a strong limitation on the number of prefixes on a verb, i.e., more than two prefixes is not allowed.

In Section 3, I argued that CBO is unlikely to make correct predictions with inflectional affixes. However, Russian prefixes are canonically neither derivational nor inflectional, but share properties of both. For this reason, it is unclear whether we should expect to find the acyclic ordering that CBO predicts. In the following sections, I present data that indicate that Russian prefixes can be ordered with significant acyclicity.

[5] Examples of Russian verbs with potentially more than two prefixes include the following:

do-ras-pro-stranit'	'to finish spreading/passing out'
pere-o-s-myslit'	'to reconsider'
pere-ras-pre-delit'	'to redistribute'
pred-ras-po-ložit'	'to predispose'

These verbs raise an important question about what counts as a prefix at all. It should be stressed that even if an affix may be identified in a given word, it does not necessarily follow that it is stored or processed as such. For example, in *pred-ras-po-ložit'* 'to predispose' the prefix *po-* is problematic because there is no word **ložit'* from which *po-ložit'* could be derived. Based on criteria of semantic transparency, it is unlikely that *položit'* can ever be parsed into component morphemes, but rather is processed whole, and the more likely parse is *pred-ras-položit'*, a verb with only two prefixes. A similar analysis may be applied to *do-ras-pro-stranit'* 'to finish spreading out' (**stranit', *prostranit'*), as well as *pere-ras-pre-delit'* 'to redistribute' (**predelit'*).

[6] Some prefixes are missing from his lists, including *bez-* 'without', *vz-* 'up', *nad-* 'over', *o-* 'about', *pre-* 'trans-', and *pred-* 'fore-'.

Superlexical prefixes		Lexical prefixes	
do-	completive	*v-*	'in'
za-	inceptive	*vy-*	'out'
na-	cumulative	*do-*	'up'
na-	saturative	*za-*	'onto'
ot-	terminative	*iz-*	'out of'
pere-	repetitive	*na-*	'on'
pere-	excessive	*ot-*	'away'
po-	delimitative	*pere-*	'across'
po-	attenuative	*po-*	'along'
po-	distributive	*pod-*	'under'
pro-	perdurative	*pri-*	'by'
		pro-	'about'
		ras-	'around'
		s-	'from'
		u-	'at'

Table 1. Superlexical and lexical prefixes (Svenonius 2004a)

5 Methodology

All thirty-two prefixes listed in the *Dictionary of Russian morphemes* (Kuznetsova and Efremova 1986) were considered for analysis.[7] This means that the set of affixes under investigation represents a more or less complete set of native Russian prefixes. Russian data were taken from Tixonov's *Orthographic-morphemic dictionary* (2002), which contains approximately 100,000 words, parsed into component morphemes.[8] A computer script was used to collapse allomorphy and identify all words with two or more prefixes. These words were then examined by hand to remove false hits.[9] In all, 985 lexemes with stacked prefixes were identified. Eleven prefixes were not found in any stacking environment, leaving 21 prefixes for further analysis.[10] Of these 21 prefixes, 3 were found only in initial position, 18 were found in both positions, and none were found exclusively in second position.[11] Of the 484 hypothetically possible prefix combinations, 153 (31.6%) are attested.

6 Results

One unique feature of these data, unlike previous CBO-related studies of affix order, is that there are twenty pairs of prefixes with reciprocal combination, e.g., *po-raz-* and *raz-po-*. These combinations are highlighted with dark boxes in Figures 2 and 3. This means that no ordering of the prefixes can result in less than twenty rank order violations, which represents 13% of all

[7] Allomorphs were counted as one prefix, e.g., *raz-* = *ras-* = *raz''-* = *ros-*

[8] The subtitle of the dictionary ("100,000 words") is slightly misleading. The dictionary frequently includes several entries from the same verbal lexeme. For example, past passive participles are frequently included as separate headwords.

[9] For example, the root *voz-* 'take by vehicle' was frequently mistaken as the prefix of the same spelling, as in *v-voz-it'* 'to import'. All adverbs were also removed, including those of the type *snačala* 'from the beginning' and *vdogonku* 'in pursuit of', because what appears to be a prefix is actually a lexicalized prepositional phrase.

[10] Despite being stacked recursively on itself, the prefix, *pra-*, 'grand-'(father/mother) was also removed because, not surprisingly, it did not stack with any other prefix.

[11] Since superlexical prefixes are a subset of lexical prefixes, we should expect that all prefixes be found closer to the stem, but only superlexical prefixes should stack on other prefixes. However, in stacking environments, not even one of the 21 prefixes considered in this study is found only in second position (i.e., closer to root). Eighteen prefixes are found in both positions, and three are attested only in first position (i.e., attached to other prefixes). This is precisely the opposite of what the widely-accepted Mirror Principle approach predicts.

attested combinations. Two prefixes (*s-* and *po-*) participate in 60% of the reciprocal combinations. Seven prefixes, *iz-*, *vy-*, *v-*, *pre-*, *do-*, *nad-*, and *pred-*, do not participate in any reciprocal combinations. Only one prefix, *s-*, stacks on itself (e.g., *so-s-vodničat'* 'to get [two people] together').

In order to identify whether the Russian prefix data are significantly acyclic, I attempted to identify an optimally acyclic ordering for the Russian prefix data, i.e., an ordering that yields the fewest possible rank order violations, i.e., attestations "above the diagonal". Although identifying optimal orders is computationally intractable, approximations can be computed. To this end, I wrote a simple computer algorithm, named aveBestRank, to approximate the optimal ordering.[12]

The aveBestRank algorithm applied to the Russian prefix data yielded an ordering with 26 violations. Since 20 of these violations are unavoidable reciprocal combinations, six of these violations could hypothetically be avoided in a different ordering. In fact, adjusting aveBestRank's order by hand, I was able to reduce the number of violations to 23 (Figure 3). Each number represents the type count of words with a given prefix combination. For example, the bottom-left cell indicates that there are 13 headwords in Tixonov's dictionary that begin with the prefix combination *pred-u-*.

	u	raz	iz	pro	s	vy	na	za	v	o	vz	ot	pri	po	pod	pre	pere	bez	do	nad	pred
u	---	3			1																
raz	19	---		1	9					13				16							
iz	3	4	---				4														
pro			9	---	1			1			11										
s	2	6	3	11	3			1				7	9	6			2				
vy				6	2	---															
na					5	2	---						1		2						
za	1	3		6	2			---					2								
v		5						1	---												
o	8	2		6	20			4		---			3	8				189			
vz					5					5	---			5							
ot					5					1		---									
pri	28				13		9	2	12	13		4	---								
po	6	23	4	2	14	17	23	10	2	14	2	2	5	---				2			
pod		5			9	1	12	5			2	1		1	---						
pre	8			6						2	15				6	---					
pere	18	7	4		13	9	12	19	4	44	1				5		---				
bez	8	10	4	7	14	3	3			7	2	9	9	8		1	3	---			
do	3	2			3	4			4				2	1					---		
nad																	1			---	
pred	13	6	1		1	2		1		8	7	2		6							---

Figure 3. Adjacency matrix of Russian prefix stacking (aveBestRank, then adjusted by hand)

To determine how well random chance can account for the degree to which the prefixes can be acyclically ordered, I follow the methodology used in recent studies of affix combinations and CBO (Plag and Baayen 2009, Zirkel 2010). Plag and Baayen generated 10,000 random datasets, ordered them using an approximation algorithm, and tracked how many rank order violations the random datasets achieve. If attested data have significantly fewer violations than the randomly generated data, then the degree to which they can be hierarchically ordered is not the result of chance.

For the Russian prefix data, I generated 100,000 random datasets with the same number of affixes (21) and the same number of attested combinations (153) as the attested data. For each random dataset, I recorded the number of rank order violations in the aveBestRank ordering. As mentioned earlier, the number of rank order violations in the aveBestRank ordering of the Russian data is 26. The minimum score from 100,000 random datasets was 28, with two datasets achieving that score. The results of this simulation show that the likelihood that chance can account for the acyclic ordering of the Russian prefix data is less than 1 in 100,000.

7 Discussion

In the last decade, a handful of studies of affix ordering and CBO have demonstrated that affix

[12] Unlike Plag and Baayen 2009 and Zirkel 2010, I do not use the *graphviz* approximation algorithm, but instead use aveBestRank, which outperformed graphviz in my tests.

combinations are more-or-less hierarchical, and that affixes' ranks in these hierarchies correlate with measures of parsability, such as productivity, phonotactics, relative frequency, etc. (e.g., Hay and Plag 2004, Plag and Baayen 2009, Zirkel 2010, Parker and Sims in progress). All these studies have explicitly focused on derivational affixes, and have left alone the question of whether CBO is applicable to inflectional morphology. In the present article, I have argued that CBO makes inaccurate predictions of inflectional affix order, because of the mismatch between the foundations of each: according to CBO, affix ordering is rooted in the *form* or *identity* of affixes, but inflectional affixes are generally ordered by morphosyntactic *function*. Yet despite this conclusion, Russian prefixes – which have some properties typical of inflection – have been shown to be significantly acyclic. It should be stressed that no attempt has been made to correlate the optimally acyclic ordering with measures of parsability, since the needed resources for such an analysis are not available for Russian.

Since this result is not necessarily expected – given the inflectional properties of Russian prefixes – it is possible that some principle other than CBO can explain the hierarchical ordering that we find in Russian prefixes. Although further research is needed, we may speculate about what principle or combination of principles could potentially explain these data. All of the principles discussed below may contribute to affix ordering in varying degrees for different languages.

First, acyclic affix orderings can simply be an accident of history. We know that once established, derivational affixes generally become less productive over time (e.g., Bauer 2001:7). The time at which an affix becomes productive or unproductive has a direct effect on its position relative to other affixes. The earlier that an affix has become unproductive, the closer it will be found to the stem, and the more recently an affix has become productive, the greater the potential for being found outside of other older affixes. To take an analogy from archaeology, affix orderings in the lexicon could be viewed as a stratigraphic record of when affixes were productive, with older affixes being found deeper in the structure and newer affixes closer to the surface. Note that this does not imply that there was ever any synchronic motivation for the relative positions of affixes, but is merely a function of the time at which an affix was productive. The natural rise and fall of affix productivity results in more acyclicity than if productivity were eternally static, since some affixes systematically occur either before or after others, depending on when they were productive. Support for this explanation comes from the fact that earlier studies of CBO have shown high correlations between measures of productivity and ranks within the optimally acyclic orderings. However, one problem with this view is that it oversimplifies what it means for an affix to be productive. For example, competing English suffixes *-ical* and *-ic* both produce adjectives, and in general *-ic* is far more productive. But in the context of words that end in *-ology* (e.g., *biology*), *-ical* is far more productive than *-ic* (Lindsay 2011). Although their domains are distinct, both suffixes are robustly productive. This example is strong evidence that there are synchronic motivations of affix ordering that cannot be explained by a simple historical account of this sort.

A second possible explanation of acyclicity is that a synchronically motivated ordering at an earlier stage in the language – whether morphological, phonological, semantic, or otherwise – can become fossilized in individual lexical items, after which the grammar of the language changes. Unlike the accidental rise and fall of affix productivity discussed above, this presupposes an earlier synchronic motivation of the ordering. If the original synchronic system was acyclic, then the remnant fossils should obviously exhibit the same acyclicity. For example, Papke (2010) investigated preverb ordering of Classical Sanskrit, which shows strong acyclic tendencies similar to Russian prefixes. After investigating many potential explanations of their ordering, including CBO, she concluded that the preverb ordering of Sanskrit is not due to any synchronic motivation, but is a fossilized part of the lexicon, inherited from the grammar of Proto-Indo-European, whose preverb ordering was originally "based on semantic requirements of the verbal complex" (Papke 2010:163). Perhaps the most persuasive evidence for this conclusion is a strong correlation between the preverb ordering of Classical Sanskrit and Old Irish, two languages that are only very distantly related.

One final principle, which I will call the Enhanced Planning Hypothesis (EPH), was suggested by Plag and Baayen (2010:145), who claim that an acyclic ordering "afford[s] enhanced anticipation (in comprehension) and enhanced planning (in production)". As one progresses down the ranks of an affix hierarchy, the number of affixes that can potentially follow decreases,

because higher-ranking affixes can be ruled out as potential sequents. On the other hand, if there are rank order violations, they introduce recursive feedback loops that make it impossible to rule out higher-ranking affixes as potential sequents. In other words, the predictive power of a particular affix is maximized when all of the affixes' combinations follow a hierarchy. As cycles are introduced, that predictive power decreases rapidly. Although computationally appealing, it is unclear by what mechanism an acyclic arrangement should come to exist in the first place.

So far, the corpus methodologies used in studies of affix ordering have not been adapted to differentiate historical vestiges from active synchronic constraints. Because of this, it has been difficult – or impossible – to differentiate between potential factors or causes of affix ordering. Although there must be synchronic restrictions on affix combinations, future research must find ways to be sensitive to diachronic factors like those discussed here, in order to adequately isolate potential synchronic factors, such as CBO and EPH.

8 Conclusions

When Hay (2003) first proposed CBO, she explicitly left the question of whether CBO applied to inflectional morphology open, and researchers of affix ordering and CBO have so far ignored inflectional affixes altogether. In this article, I have argued that when applied to inflectional affixes, CBO is unlikely to make accurate predictions. This conclusion casts doubt on CBO more generally, because its cognitive foundations imply that it should apply very broadly. My treatment of this question has assumed that the parsing of inflectional affixes is based on the same factors as derivational affixes, such as phonotactics, productivity, and relative frequency. However, it is not unreasonable that the parsing of inflectional affixes is fundamentally different – and more language-specific – than derivation. Inflectional affixes are typically far more productive and far more frequent than derivational affixes (Gaeta 2008). Because of this, they may enjoy privileged parsing, qualitatively different from derivational affixes.

In this article, I also examined whether the ordering of Russian prefixes exhibits the same kind of acyclicity observed in English affixes and Russian suffixes. I found that Russian prefix combinations are inherently acyclic, or in other words, the degree to which they can be ordered acyclically cannot reasonably be accounted for by chance. This result should not necessarily be taken as support for CBO, since it was not possible to test for correlation with measures of parsability. Regardless of how one explains the appearance of acyclicity, with Russian prefixes being added to a growing list of languages with acyclic affix ordering, this issue merits more attention from a typological point of view. Is acyclic affix ordering a language universal?

Many factors can potentially contribute to the development of constraints on affix ordering in a language. Although all of the principles discussed above – and surely others – may explain some parts of Russian prefix ordering to a certain degree, future research should include methods that can better isolate these (and possibly other) factors using historical or behavioral data.

References

Baayen, Harald, and Robert Schreuder. 1999. War and peace: Morphemes and full forms in a noninteractive activation parallel dual-route model. *Brain and Language* 68 (1):27–32.

Babko-Malaya, Olga 2003. Perfectivity and prefixation in Russian. *Journal of Slavic Linguistics* 11 (1 March):5–36.

Beard, Robert. 1987. Morpheme order in a lexeme/morpheme-based morphology. *Lingua* 72 (1) (May):1–44.

Beard, Robert. 1995. *Lexeme-morpheme base morphology: a general theory of inflection and word formation*. SUNY Press.

Bertram, Raymond, Robert Schreuder, and Harald Baayen. 2000. The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26 (2):489–511.

Burani, Cristina, and Anna Thornton. 2003. The interplay of root, suffix and whole-word frequency in processing derived words. Ed. R. H. Baayen and R. Schreuder. *Morphological structure in language processing* 151:157–232.

Clahsen, Harold, Ingrid Sonnenstuhl, and James P. Blevins. 2003. Derivational morphology in the German mental lexicon: A dual mechanism account. *Trends in linguistic studies and monographs* 151:125–156.

Deen, Kamil Ud. 2006. Subject agreement in Nairobi Swahili. In *Selected proceedings of the 35th annual conference on African linguistics*, 225–233.
Fabb, Nigel. 1988. English suffixation is constrained only by selectional restrictions. *Natural Language & Linguistic Theory* 6 (4) (November):527–539.
Gaeta, L. 2008. On the double nature of productivity in inflectional morphology. *Morphology* 17 (2) (January):181–205.
Hay, Jennifer, and Harald Baayen. 2001. Parsing and productivity. *Yearbook of morphology*: 203–235.
Hay, Jennifer. 2002. From speech perception to morphology: Affix ordering revisited. *Language* 78 (3) (September):527–555.
Hay, Jennifer. 2003. *Causes and consequences of word structure*. Routledge, August 19.
Hay, Jennifer, and Ingo Plag. 2004. What constrains possible suffix combinations? On the interaction of grammatical and processing restrictions in derivational morphology. *Natural Language and Linguistic Theory* 22 (3):565–596.
Isačenko, Aleksandr. 1960. *Grammaticeskij stroj russkogo jazyka. Morfologija. Castj vtoraja [Grammatical composition of Russian. Morphology. Part 2]*. Bratislava: Vydavatelstvo Slovenskej Akadémie vied.
Istratkova, Vyara. 2005. On multiple prefixation in Bulgarian. *Nordlyd* 32:301–321.
Järvikivi, Juhani, Raymond Bertram, and Jussi Niemi. 2006. Affixal salience and the processing of derivational morphology: The role of suffix allomorphy. *Language and cognitive processes* 21 (4):394–431.
Kiparsky, Paul. 1982. Lexical phonology and morphology. *Linguistics in the morning calm* 2:3–91.
Kiparsky, Paul. 1985. Some Consequences of Lexical Phonology. *Phonology Yearbook* 2:85-138.
Kuznecova, Ariadna I., and Tat'jana F. Efremova. 1986. *Slovar' morfem russkogo jazyka*. Moscow: Russkij jazyk.
Lindsay, Mark. 2011. *Self-organization in the English suffix system* presented at the Quantitative Measures in Morphology and Morphological Development, January 15, University of California, San Diego. http://www.chd.ucsd.edu/seminar/documents/san_diego.pdf.
Papke, Julia K. P. 2010. Classical Sanskrit preverb ordering: a diachronic study. PhD, Columbus, OH: The Ohio State University. http://etd.ohiolink.edu/view.cgi/Papke%20Julia.pdf?osu1268028740.
Plag, Ingo. 1996. Selectional restrictions in English suffixation revisited: a reply to Fabb (1988). *Linguistics* 34 (4):769–798.
Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin: Walter de Gruyter.
Plag, Ingo, and Harald Baayen. 2009. Suffix ordering and morphological processing. *Language* 85 (1):109–152.
Ramchand, Gillian. 2008. Perfectivity as aspectual definiteness: Time and the event in Russian. *Lingua* 118 (11):1690–1715.
Richardson, Kylie R. 2007. *Case and aspect in Slavic*. Oxford University Press.
Romanova, Eugenia. 2005. Superlexical versus lexical prefixes. *Nordlyd* 32:255–278.
Russian National Corpus. http://www.ruscorpora.ru/en/index.html.
Schreuder, Robert, and Harald Baayen. 1997. How complex simplex words can be. *Journal of Memory and Language* 37 (1):118–139.
Spencer, Andrew. 1991. *Morphological theory: An introduction to word structure in generative grammar*. Wiley-Blackwell.
Stump, Gregory. 1992. Position classes and morphological theory. *Yearbook of morphology*: 129–180.
Stump, Gregory. 1996. Template morphology and inflectional morphology. *Yearbook of morphology*: 217–241.
Stump, Gregory. 2001. *Inflectional morphology: A theory of paradigm structure*. 1st ed. Cambridge University Press.
Svenonius, Peter. 2004. Slavic prefixes and morphology. An Introduction to the Nordlyd volume. *Nordlyd* 32:177–204.
Tixonov, Alexander N. 2002. *Morfemno-orfografičeskij slovar': Okolo 100 000 slov*. Moscow: OOO Izdatel'stvo Astrel'. http://slovari.yandex.ru/~книги/Морфемно-орфографический словарь/.
Zirkel, Linda. 2010. Prefix combinations in English: structural and processing factors. *Morphology* 20 (1) (February): 239–266.

Department of Slavic and East European Languages and Literatures
The Ohio State University
Columbus, OH 43210
reynolds.531@osu.edu