“Are Russian prefixes out of order?: Complexity-Based Ordering and position classes”
Robert Reynolds, The Ohio State University

In natural language, only a very small proportion of possible affix combinations are actually attested (e.g. *atomicity, but *atomicness; or atomless–atomlessness, but *atomlessness). The last several decades have seen a number of attempts to identify general principles governing affix combinability (e.g. Kiparsky 1985, Fabb 1988, Plag 1996, Hay 2003, Plag & Baayen 2009). A relatively recent processing-based proposal, Complexity-Based Ordering (Hay 2003), has been supported in studies of canonical derivational affixes, e.g. English prefixes and suffixes, and Russian suffixes. However, a processing account should apply very broadly, and this study is a first effort at exploring how to interpret Complexity-Based Ordering with inflectional affixes. Various principles have been proposed for governing the order of inflectional affixes including syntax and semantic scope (Baker 1985, Rice 2000). In this presentation, I assume that in at least some languages, they are conditioned by language-specific position classes (Stump 1992). Based on data from Russian prefix combinations, I argue that the primary evidence which has been cited as evidence of Complexity-Based Ordering – acyclic ordering of affixes – is likely the result of both historical processes and cognitive processing. Historical factors are usually ignored in the corpus-based research of affix ordering, conflating possible cognitive effects with historical.

Hay (2003) proposed a processing approach to affix ordering, Complexity-Based Ordering, which claims that, in addition to selectional restrictions, affixal combinations are restricted by cognitive processing restrictions, and in particular, the parsability of the affix: an affix can only combine with affixes that are less parsable than itself. For example, in addictiveness, -ness is quantifiably more easily parsed than -ive, so the combination is allowed. This ordering principle leads to the prediction that affix combinations are acyclic (i.e. A>B, B>C, but not C>A), a prediction that has been supported by studies of English and Russian derivational affixes. However, by examining a range of possible inflectional scenarios, I demonstrate that Complexity-Based Ordering is unlikely to make accurate predictions for inflectional affixes, as well as other affixes associated with position classes.

Russian verbal prefixes have two position classes (e.g. Isačenko 1960), so the presence of an acyclic ordering in Russian prefixes is not expected. Some prefixes can appear in both positions, but with distinct meanings in each position. For example, in the internal position, pere-means ‘across’ (e.g. perexodít ‘to go across’), but in the external position, it means ‘excessively’ (e.g. pere-vy-polnit ‘to execute beyond expectations’). In order to measure the degree to which the order of Russian prefixes is acyclic, I examine the stacking of 32 Russian prefixes in Tixonov’s Morpho-orthographic Dictionary of Russian (Figure 1). By comparing these data with 100,000 randomly generated datasets, I show that Russian prefixes are significantly acyclic (p<0.01). Since Complexity-Based Ordering is not a viable explanation of this acyclicity, I discuss alternative explanations, both cognitive and historical. For example, Plag & Baayen (2009) speculate that an acyclic ordering of affix combinations might arise because it maximizes the ability to predict the next affix, by ruling out ‘higher-ranking’ affixes.

Complexity-Based Ordering has hit upon a striking feature of affix stacking – acyclic ordering – which has, I argue, both historical and cognitive factors. Although historical factors of word structure are sometimes mentioned in the affix-stacking literature, corpus methodologies have not been adapted to disambiguate cognitive and historical effects. I discuss how the use of historical corpora can contribute to future processing-based research of affix stacking.
Figure 1. Type-frequency adjacency matrix of Russian prefixes in an approximated optimally acyclic order. Boxed cells indicate combinations that form reciprocal pairs. For example, both u-raz- (3 types) and raz-u- (19 types) are attested, so both cells are boxed.

References


