

Title: Survival of the Frequent: How Affixes Select for Small Stems

Speaker: [Maria Gouskova](#) (NYU)

Time: Friday November 3rd, 12 – 1:30 pm

Location: Emerson 305

Abstract:

Selection for size (usually in terms of syllables) is a commonplace kind of phonological subcategorization, but such restrictions remain poorly understood. Restrictions might be noted but left unformalized, or framed in terms that do not withstand scrutiny (e.g., trochees are often implicated in English syllable-counting selection, but they are different trochees depending on the affix, and with some unhandled exceptions). I argue that this lack of interest has allowed an interesting correlation to escape notice: size selection is often, if not always, correlated with lexical or semantic selection. In this talk, I demonstrate this on a couple of previously unnoticed cases of syllable-counting selection in Russian, where one affix selects for maximally disyllabic stems, and another selects for maximally trisyllabic ones. Neither affix's selection can be correlated with any kind of foot in Russian. But the affixes have highly specific semantics: one refers to external (typically inalienable) attributes, and the other to baby animals and humans. Semantically broad affixes such as deadjectival nominalizers and the Russian counterparts of "-ist" and "-ism" are, by contrast, non-selective. I show quantitative evidence that for the selective affixes, the sublexicon of stems cannot have arisen by chance; randomly drawing similarly-sized sublexicons from a morphosyntactically similar pool does not result in the same syllable count distributions. For non-selective affixes, by contrast, the syllable count distributions in their sublexicons look just like what would result from chance selection.

My argument is that syllable counts are not part of the affixal realization rules at all; the generalizations are emergent and extracted by statistical comparison between the lexicon and the sublexicon. I show that when the affixes are extended productively in hapax legomenon use in a corpus, the distribution mimics the sublexicon in the case of selective affixes, but the reference lexicon in the case of non-selective ones. I conclude that the right approach to size selection is sublexical generalization, not realization rules with subcategorization frames.