

TOWARDS DEVELOPING A MODEL OF RUSSIAN DERIVATION

Natalia Pertsova

npertsova@rambler.ru

Research Scientific Computer Centre of Lomonosov's Moscow State University

The topic of the talk is a model of Russian derivation being developed (with intervals) since 1990s, and an automatic system for Russian derivational analysis and synthesis called RUSLO (*Русское Словообразование*), which is based on the model. The model consists of two main components: formal and semantic.

Formal derivational information includes lists of affixes (prefixes, suffixes and circumfixes), patterns for derivatives and compounds, and some additional rules and restrictions. Among them we shall mention rules of generating those forms of stems which do not coincide with stems in dictionary entries, for example: stems used in indirect cases of nouns: *весн(а) – весенн(ий)*; finite verbal forms: *мы(ть) – мойк(а)*; stems obtained by regular and irregular consonants alternations: *к/ч (рука – ручка)*; ‘ etc.

Additional restrictions include constraints on grammatical characteristics of stem words and derivatives; for instance, stem nouns and derivatives with diminutive suffixes usually have the same gender, and therefore the system is allowed to relate *пирог* with *пирожок* (which are both masculine), but not *вен(а) – веноч* (the first noun is feminine, and the second one is masculine; semantically the two words have nothing in common), etc.

One of the main problems concerning Russian derivation is the fuzziness of morph boundaries. This has already been noticed by the Russian-Polish linguist I. A. Baudouin de Courtenay many years ago. Thus, Baudouin argued that morph boundaries can be more or less fuzzy: in many cases different people, or even the same person in different moments, divide the same word into morphs in different ways. This turns out to be true not only for native speakers of Russian, but for researchers as well. For instance, in the three modern Russian Academic grammars and in the dictionary of Russian morphemes by A. I. Kuznetsova we find cases of different morph division by different authors as well as cases of different morph division by the same authors. In our model, we treat this problem by introducing different types of morph division markers. It helps us to minimize the number of derivational morphs (the total list of noun derivational suffixes from the books mentioned contains about 1000 units, our suffix list is half as much). Some markers represent definite boundaries, while others – fuzzy boundaries; in the latter case the intermediate part is sometimes considered flexible and can be treated as belonging either to an affix, or to a stem.

The semantic component uses a list of regular derivational meanings for Russian. The variety of derivational meanings is divided into 4 classes (with further subdivision):

(1) null: a derivative has the same meaning as a stem word (*мотациклет – мотациклетка*);

(2) syntactic (introduced by E. Kurilovich): only the part of speech, but not meaning changes (*белый – белизна, ходить – хождение*);

(3) lexical: only meaning, but not the part of speech changes (*ваза – вазочка, ходить – приходить, серый – сероватый, роза – розарий, река – заречье*);

(4) lexico-syntactical: both meaning and the part of speech change (*веселый – весельчак, паяц – паясничать, два – удвоить*).

Some derivational meanings (but in fact only the minority of them) coincides with lexical functions from the Meaning \leftrightarrow Text model developed by I. A. Mel'chuk, to which we added some other functions, such as 'Female' (*львица, царица*), 'Descendant' (*львенок, царевич*), or 'Similar' (*ромбويد, бромид*). To describe other meanings we use the notion of an *implied predicate* – a predicate which should be introduced to obtain necessary

3rd Vienna Workshop on Affix Order: Advances in Affix Order Research

derivational meaning, for instance, for ‘object – person’ pairs it may be ‘to produce’ (*канатчик*), ‘to sell’ (*зеленицик*), ‘to obtain’ (*бородач*), ‘to be fond of’ (*книжничик*), etc.