

Lennart Lönngren

## WORD FREQUENCY IN SLAVIC, ESPECIALLY RUSSIAN

### Abstract

Frequency dictionaries exist in several Slavic languages, but the bulk of this work has been done for Russian. In this brief outline two Russian frequency dictionaries are compared: “Z” (Zasorina’s dictionary from 1977) and “U” (the Uppsala corpus dictionary from 1993).

The corpora underlying these dictionaries are of the same size and have approximately the same composition of text genres. However, the presentation of the data is very different. In Z, there is one alphabetical list comprising all words occurring in the corpus, and one shorter list with the ten thousand most frequent words. In U, the longest list corresponds to the shorter list in Z. In addition, U gives short separate lists, based on modified frequency, for words and word forms, and also for words occurring frequently in informative prose and rarely in literary texts and vice versa.

A concluding appendix presents some high frequency data extracted from U.