

Frequencies of occurrence of entries and subcategorization frames in *LGLex* lexicon with IRASubcat

Elsa Tolone & Romina Altamirano

FaMAF, Universidad Nacional de Córdoba, Argentina

elsa.tolone@univ-paris-est.fr;

romina.altamirano@gmail.com

SUMMARY

We present a method for enlarge a lexicon (with frequencies information), that is useful for parsing and others NLP applications. We show an example enlarging the verbal *LGLex* lexicon of French, using several corpora extracted from the evaluation campaign for French parsers Passage. To do that, we use the results of the FRMG parser with IRASubcat, a tool that automatically acquires subcategorization frames from corpus in any language and that also allows to complete an existing lexicon. We obtain the frequencies of occurrence for each input and each subcategorization frame for 14,068 distinct lemmas.

