A Collocation Extraction Tool for Romanian
Violeta Seretan, Eric Wehrli, Luka Nerima (University of Geneva),
Amalia Todirascu (University of Strasbourg)

- Aims: corpus-based extraction of collocations for the Romanian language to complete an existing dictionary
  - Method: syntactically-bound lexical combinations extracted from parsed data and statistical measures
    - FIPS parser (Wehrli et Nerima, 2015) for Romanian (Seretan et al., 2010)
    - Extension of the multilingual collocation extraction tool, FipsCo (Seretan, 2008) to Romanian, applied to Europarl corpus (2011)
  - Evaluation: 1 annotator, top 1000, precision 67.7%
  - Comparison with a Romanian collocation dictionary (Todirascu et al, 2008)