



# Electronic Tools and Resources for Multi-Word Unit detection and research in Serbian

Jelena Mitrovic, University of Belgrade

- Serbian is one of the under-resourced languages when it comes to NLP – many resources and tools are still being developed
- **Electronic MWUs dictionary** – morphological dictionary with complex prepositions, conjunctions, interjections, complex adjectives e.g. **mrtav pijan** ‘dead drunk’ and complex nouns e.g. **nemasno mleko u prahu** ‘fat free powdered milk’
- **Serbian WordNet** – percentage of MWUs approximately **32.5%**



- **Ontology of Rhetorical Figures for Serbian** – unambiguous formal description of 98 rhetorical figures in Serbian
- Human annotation of rhetorical figures is not precise enough due to their large number and similarities that exist – that is why an ontology is very helpful
- Many rhetorical figures are MWUs

