# Semantic Classes of Czech Verbs

Dana Hlaváčková
hlavack@fi.muni.cz
Karel Pala
pala@fi.muni.cz

Centre of Natural Language Processing
Faculty of Informatics
Masaryk University

# Semantic Classification of Czech Verbs

- Semantic classification of (Czech) verbs is related to language ontologies (verbs vs. nouns)
- Earlier attempts, e.g. Daneš (1987) – 5 main semantic classes
- Vallex (ÚFAL) – 22 syntactic-semantic classes
- Semantic classes of English verbs – Levin: 50 semantic classes
- VerbNet – 274 semantic classes (extension of Levin's classes
- VerbaLex – 82 semantic classes, 258 subclasses

# VerbaLex Semantic Classes

- Inspiration came from Levin and VerbNet

- translation and new Czech synonyms, verbs with prefixes, aspect counterparts

- Main semantic classes (ontology) of Czech verbs        - Social relations                                  - Activities                                          - Processes                                           - Communication                                       - Beginning – ending                                  - Events?                                             - Motion?                                             - Weather?                                            - ?

# Criteria for Semantic Classes

- Meanings of the verbs – what to rely on?

- Existing explanatory dictionaries (introspection)

- Empirical evidence - corpus data

- Using Word Sketch Engine – contexts, clusters

- semantic roles, types and selectional restrictions in VerbaLex

- Subclassification based on the right-side and left-side valency position (*garment, beverage, body part, ...*)

- Lexical sets – a complication?

# VerbaLex and Princeton WordNet

- Links to PWN – 8844 Czech verbs are linked to PWN v.2. via ILI

- Transferring the valency frames to English – can it work?

- Transferring the verb classes – comparison with VerbNet classes

- Unlinkables – prefixed verbs, nouns?

- Some translation equivalents cannot be lexicalized

- Virtual nodes with the description of meaning

# Valency frames and verb classes for Semantic Web?

They should serve as metadata

# Conclusion

Possible applications