Valency Frames as Event Mentions
• valency frames: detailed sense description
• valency frame detection (def.) = detection & sense disambiguation of event mentions
• our system: (auto) valency frame detection for Czech and English
• leveraging common annotation style of Prague Dependency Treebanks
• PDT 2.5: Czech, PCEDT 2.0: Czech & English
• based on Functional Generative Description (Tectogrammatics)

Valency in the Functional Generative Description (Tectogrammatics)
• a specific ability of a word to combine with other lexical units (event participants)
• argument-adjunct distinction:
  - inner participant (argument) × free modifier (adjunct)
  - specific for particular verbs can appear only once
    - ACT, PAT, ADDR...

Valence Frame — one per word sense
• all inner participants + obligatory modifiers
• semantic function (function), obligatoriness, surface form

Annotation of Prague Dependency Treebanks

PDT-Vallex & EngVallex
• lexicons interlinked with treebanks
  - EngVallex: English, based on PropBank
    7,699 frames for 4,337 verbs
  - PDT-Vallex: Czech, created with PDT
    9,191 frames for 5,510 verbs

Automatic Valency Frame Selection
• based on logistic regression
• using the same feature templates for both languages:
  - word forms
  - morphology
  - formemes
  - a-layer dependency labels
  - neighborhood (topological & syntactic)
• trained on PDT 2.5 (Czech) & PCEDT 2.0 (English)
• backoff to valency lexicons
• use: included in the automatic t-layer annotation pipeline

Results
• unlabeled = detecting that a valency frame should be selected
• labeled = with disambiguation, i.e., finding the correct frame ID

Conclusions
• first results in valency frame selection on Prague Dependency Treebanks using automatic annotation
• encouraging, given the high granularity of the lexicons
• detecting frames — easier for Czech, selecting the correct one — English better
• problems:
  - idiomatic expressions
  - frames with identical set of members

Supported by the Grant No. GPP465/16/0337 of the Grant Agency of the Czech Republic and the project LH2293 of the Ministry of Education, Youth and Sports of the Czech Republic, and by the Charles University SVV project 244 104. Data provided by LINDAT/DEBAN, Infrastructure Project MZOV CR No. LM20160047.