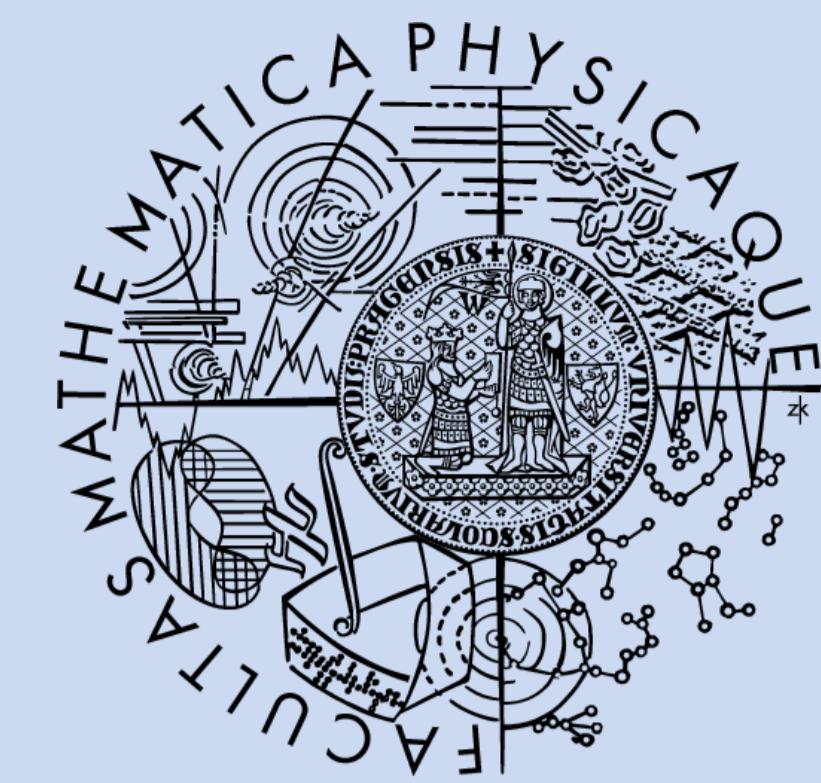
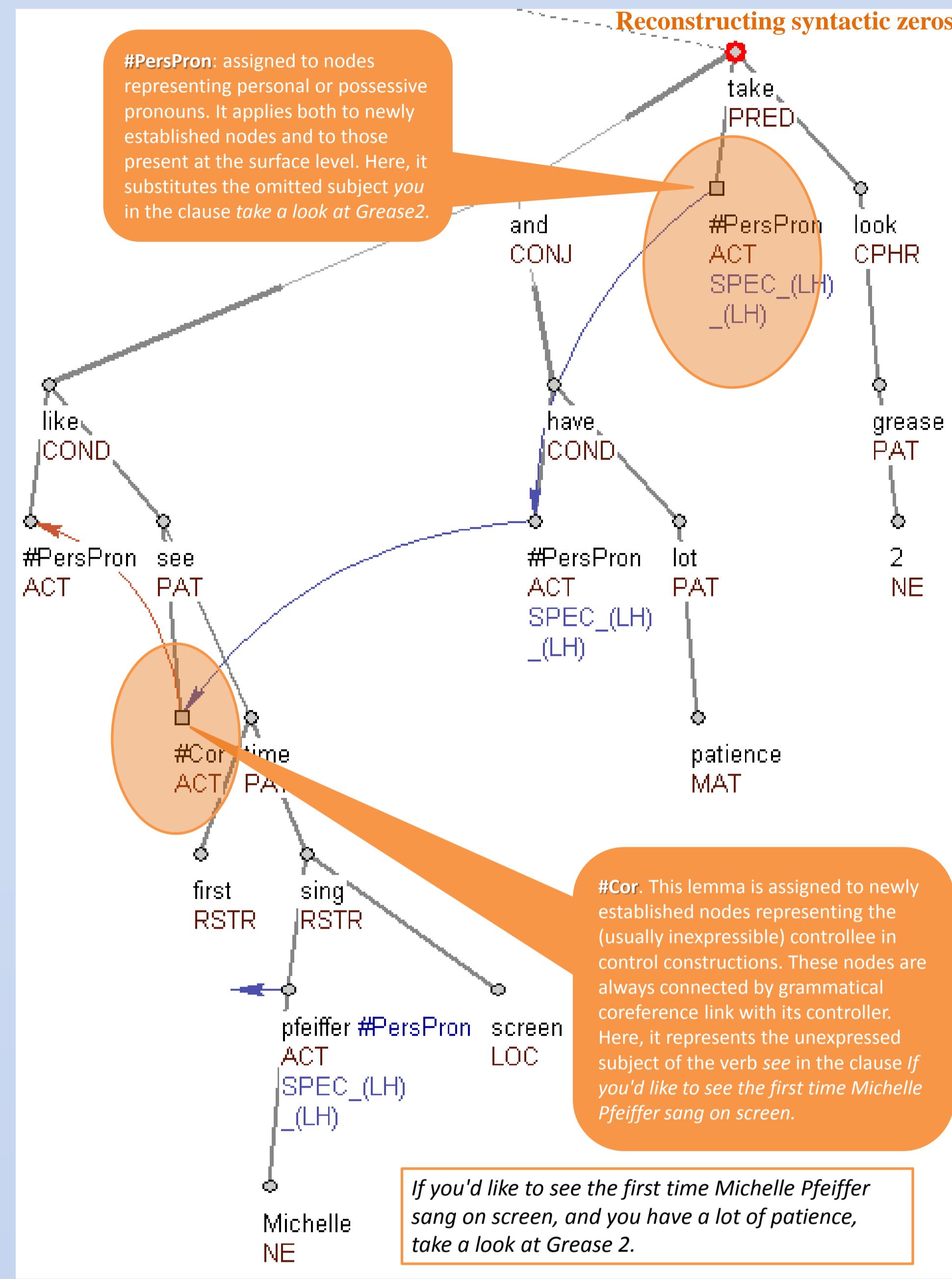


How dependency trees and tectogrammatics help annotating coreference and bridging relations in Prague Dependency Treebank



Anna Nedoluzhko Jiří Mírovský

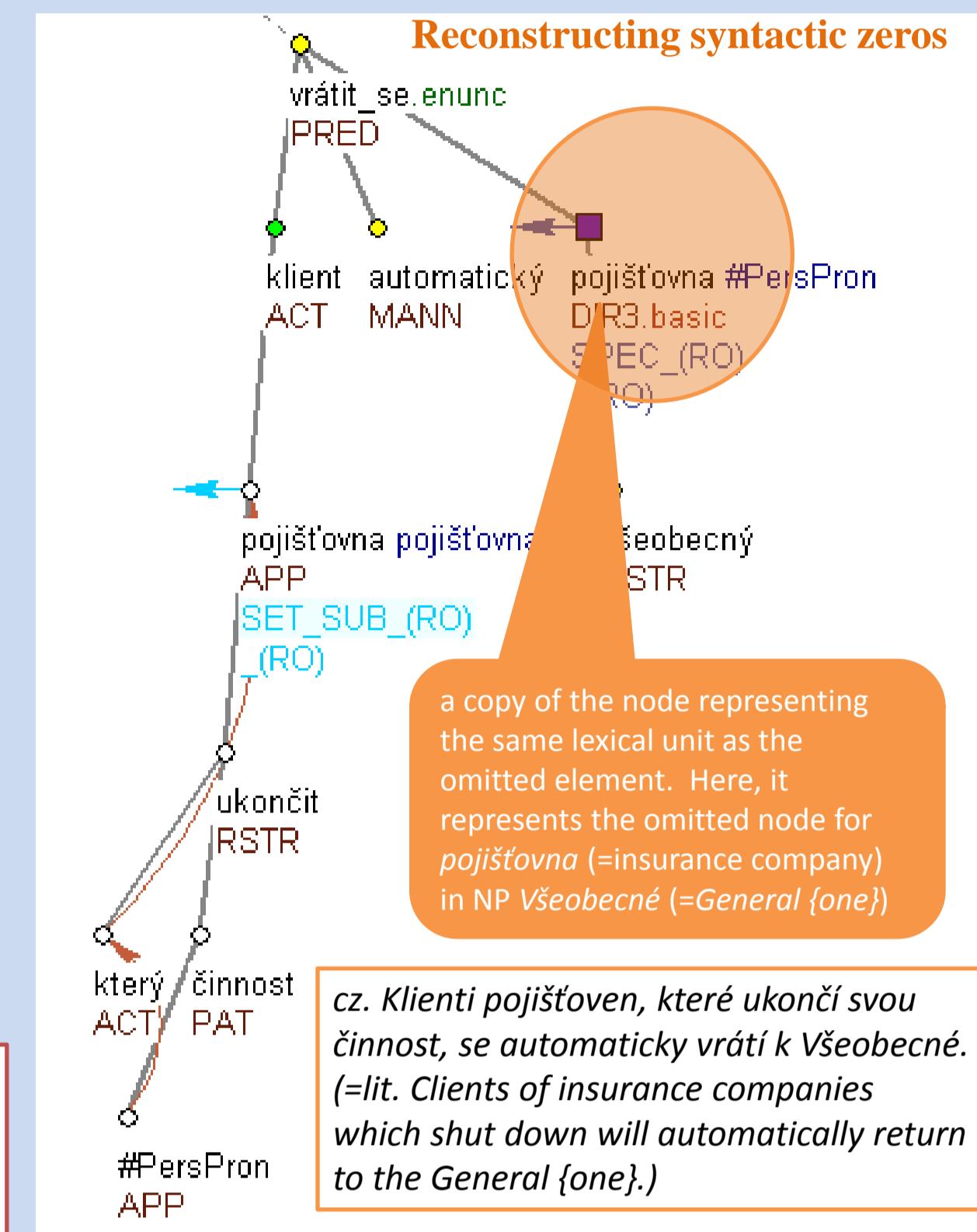


Coreference and bridging relations annotated in PDT:

- grammatical coreference
 - coreference of relative pronouns,
 - coreference of reflexive pronouns
 - arguments of verbs of control,
 - arguments in constructions with reciprocity and verbal complements
- textual coreference (pronominal and nominal, including zero anaphora)
 - coreference of specific noun phrases (type SPEC),
 - coreference of generic noun phrases (type GEN)
 - bridging relations (mark some associative semantic relations between non-coreferential entities)
 - PART-OF (e.g. room - ceiling),
 - SUBSET (e.g. students - some students),
 - FUNCT (e.g. state - president),
 - CONTRAST (e.g. this year - last year),
 - ANAF (e.g. rainbow - that word, conference - at that time)
 - REST (e.g. grandfather - grandson, enterprise - entrepreneur; etc.)

Benefits of dependency trees:

- automatic extraction of markables (MIN-IDs and maximal scope)
- syntactic zeros are reconstructed in tectogrammatical structure
 - #Perspron: personal or possessive pronouns
 - #Cor: controller in control constructions
 - #Qcor: valency modification in constructions with quasi-control, e.g. *He offered Jan {#QCor} protection*.
 - #Rcp: participants that are left out as a result of reciprocation, e.g. *The lovers kissed {#Rcp.PAT}*.
 - a copy of the node representing the same lexical unit as the omitted element
- non-referring expressions:
 - appositions
 - coordinative constructions
 - verbal complements



Syntactic zeros that cannot be linked by coreference/bridging relations:

- #Gen for a general participant (*Houses are built {#Gen.ACT} from bricks*).
- #Unsp for valency modifications with vague semantic content (*U Nováků (#Unsp.ACT) dobré varí. =They cook well at Novák's*)
- #EmpNoun for non-expressed nouns governing syntactic adjectives, which are not the case of textual ellipsis (*Příšl jen mladší. (#EmpNoun.ACT) mladší. (=lit. Come only (#EmpNoun) younger)*)
- #Obflm for obligatory adjuncts that are absent at the surface level (*Ta vypadá. (#Obflm.MANN) (=lit. That.fem looks; meaning: She looks awful/so strange...)*)

