CS395T: Structured Models for NLP
Lecture 8: Trees 2

Greg Durrett

Some slides adapted from Dan Klein, UC Berkeley

Administrivia

- Project 1 due Thursday at 9:30am

Outline

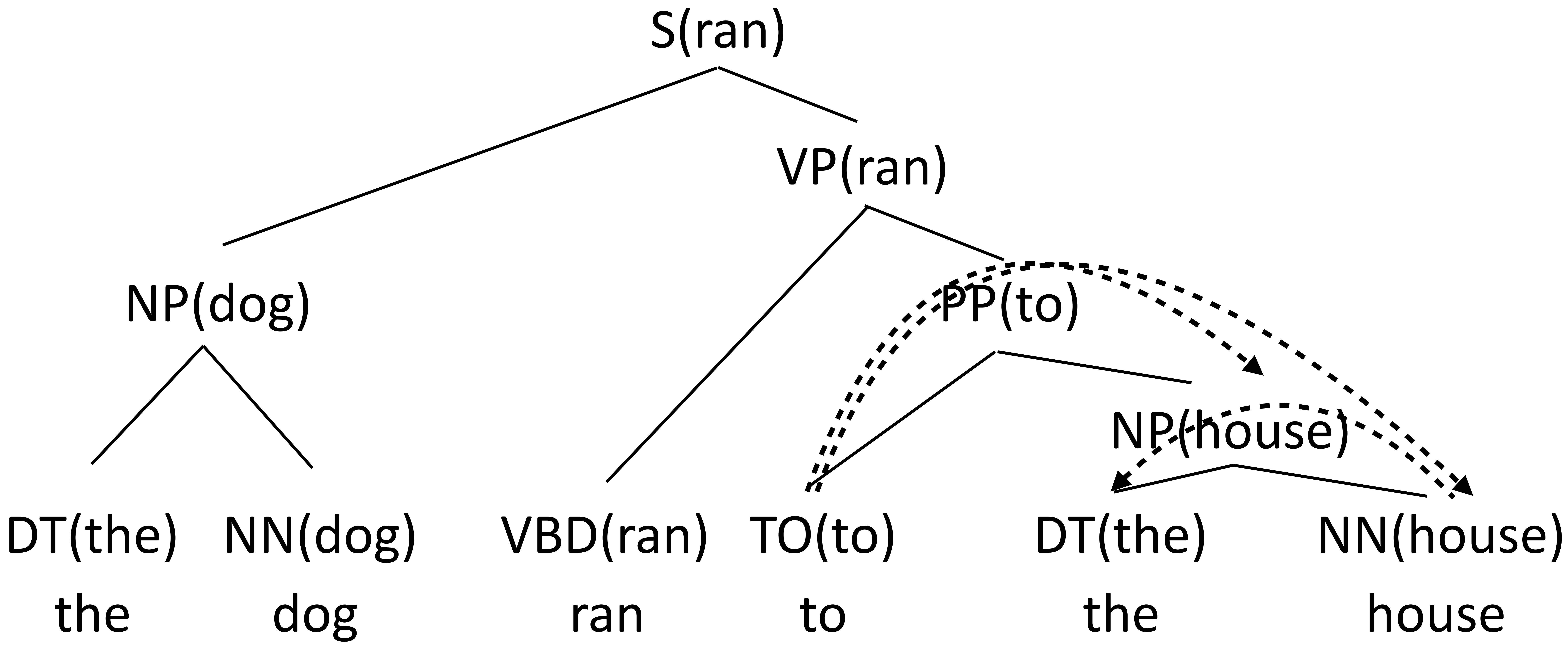
- Lexicalized and state-split constituency parsing (slides from last time)

- Dependency representation

- Contrast with constituency

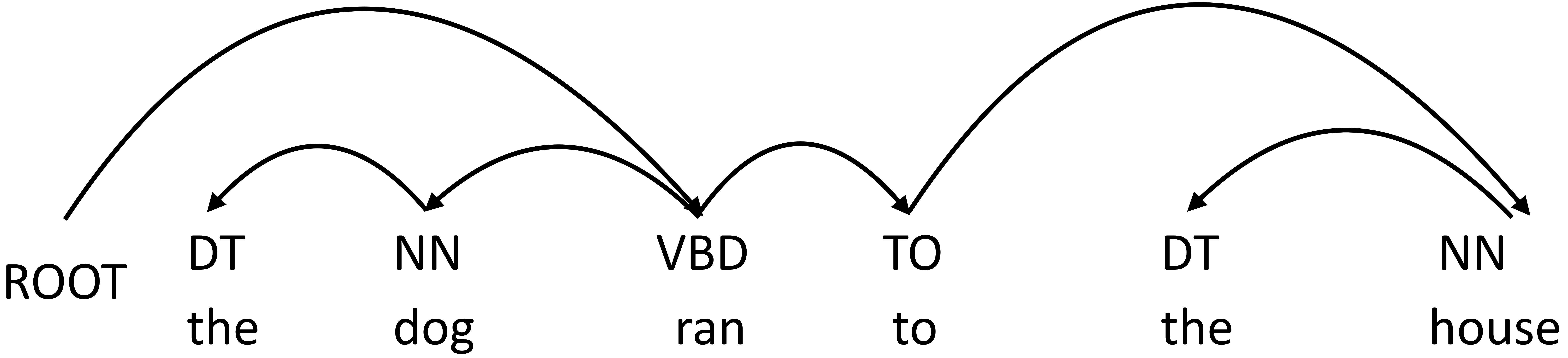
- Projectivity

Lexicalized Parsing



Dependency Parsing

- Dependency syntax: syntactic structure is defined by dependencies
  - Head (parent, governor) connected to dependent (child, modifier)
  - Each word has exactly one parent except for the ROOT symbol
  - Dependencies must form a directed acyclic graph



ROOT

DT the

NN dog

VBD ran

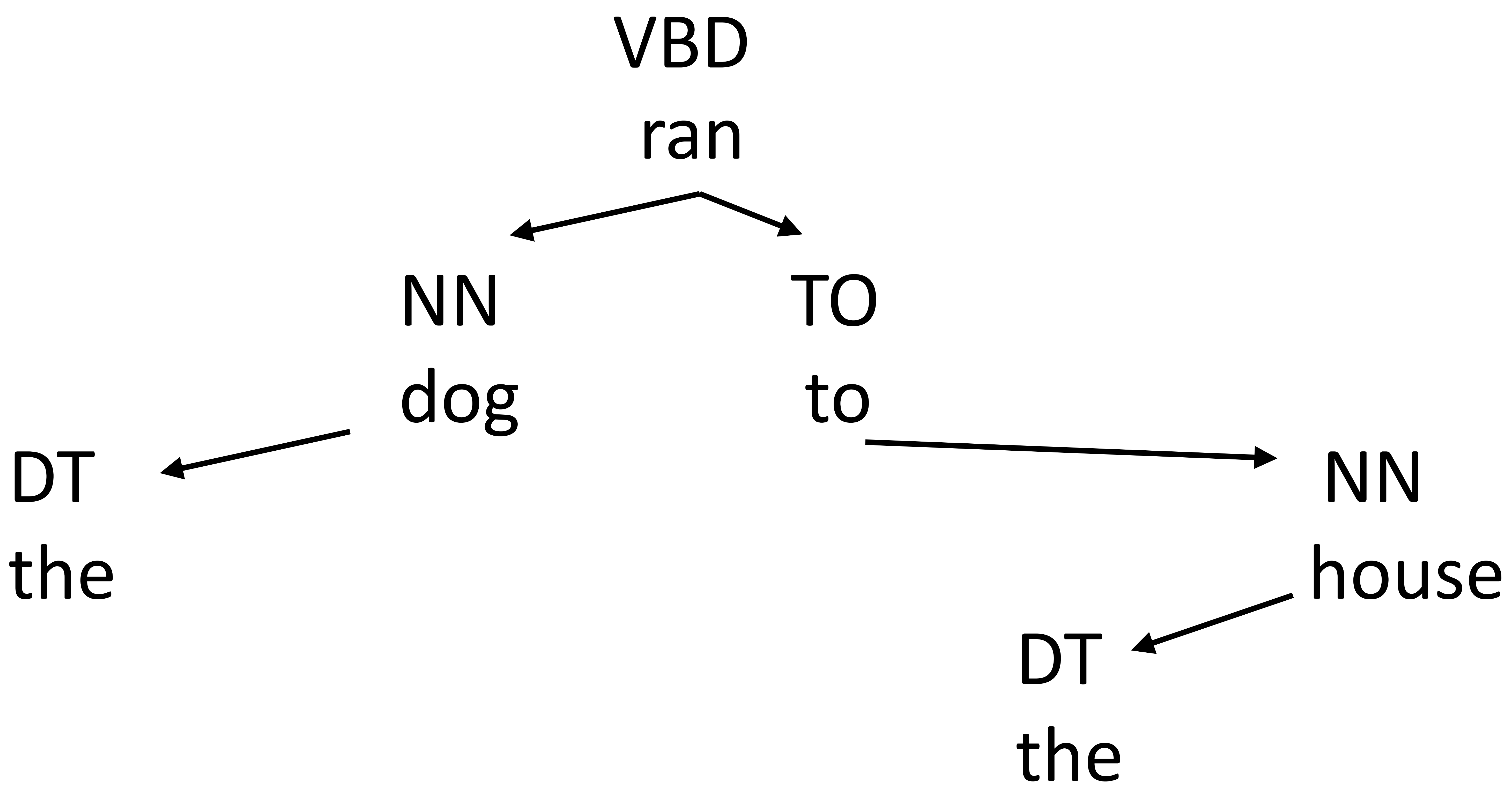
TO to

DT the

NN house

Dependency Parsing

- Still a notion of hierarchy!

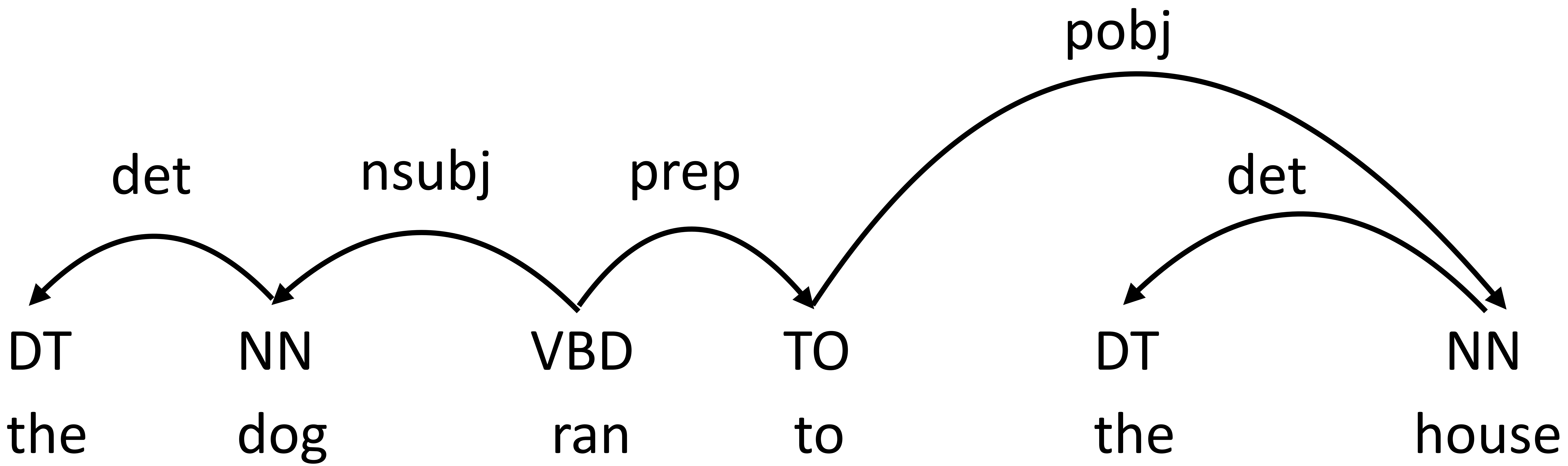


- Can still derive constituents (subtrees)

Dependency Parsing

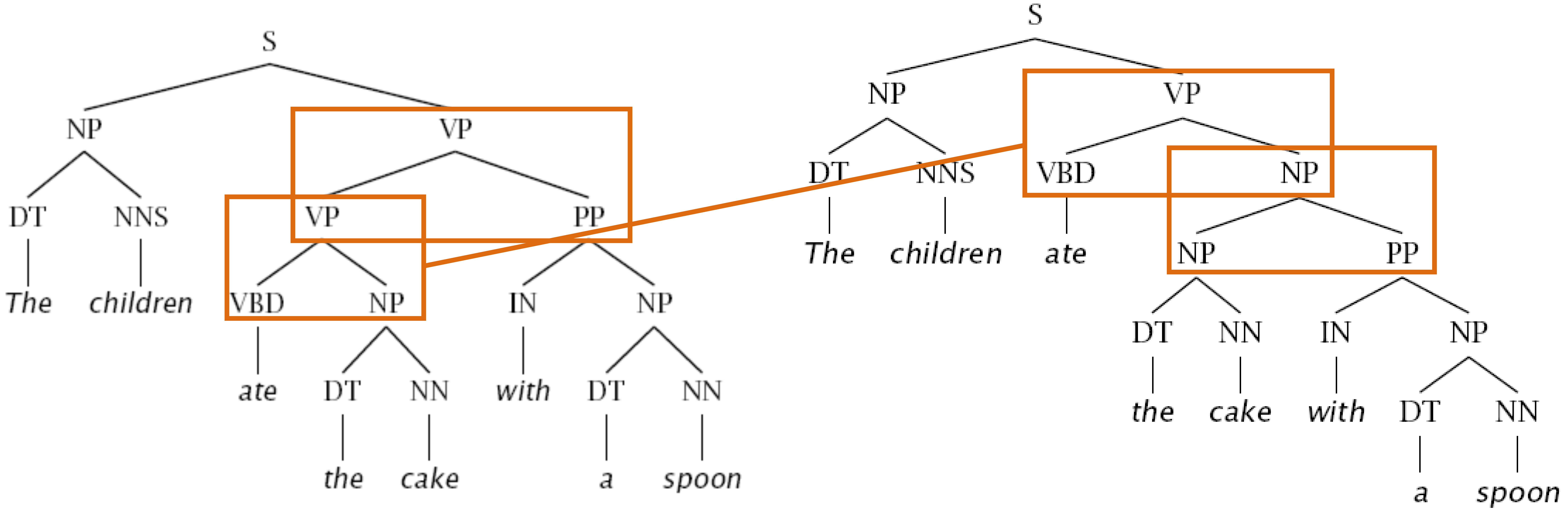
- Can label dependencies according to syntactic function

- Major source of ambiguity is in the structure, so we focus on that more (labeling separately with a classifier works pretty well)



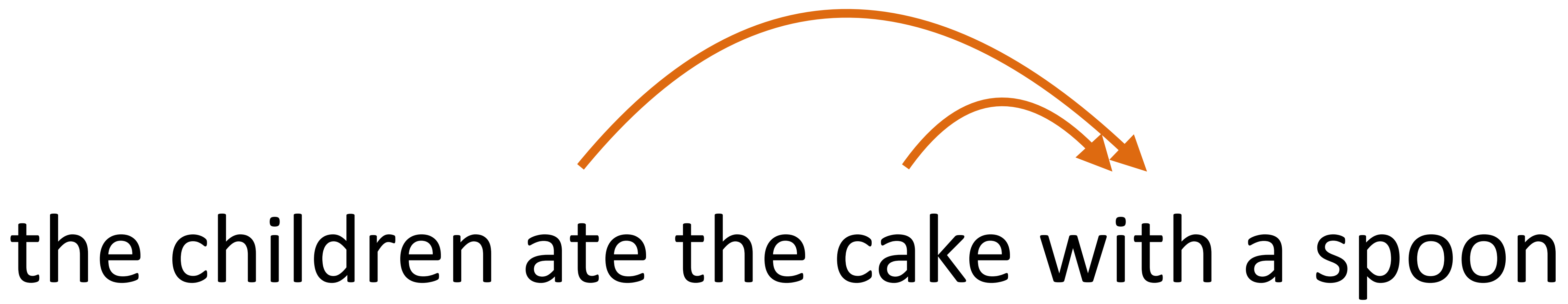
Dependency vs. Constituency: PP Attachment

- Constituency: several rule productions need to change



Dependency vs. Constituency: PP Attachment

- Dependency: one word (with) assigned a different parent

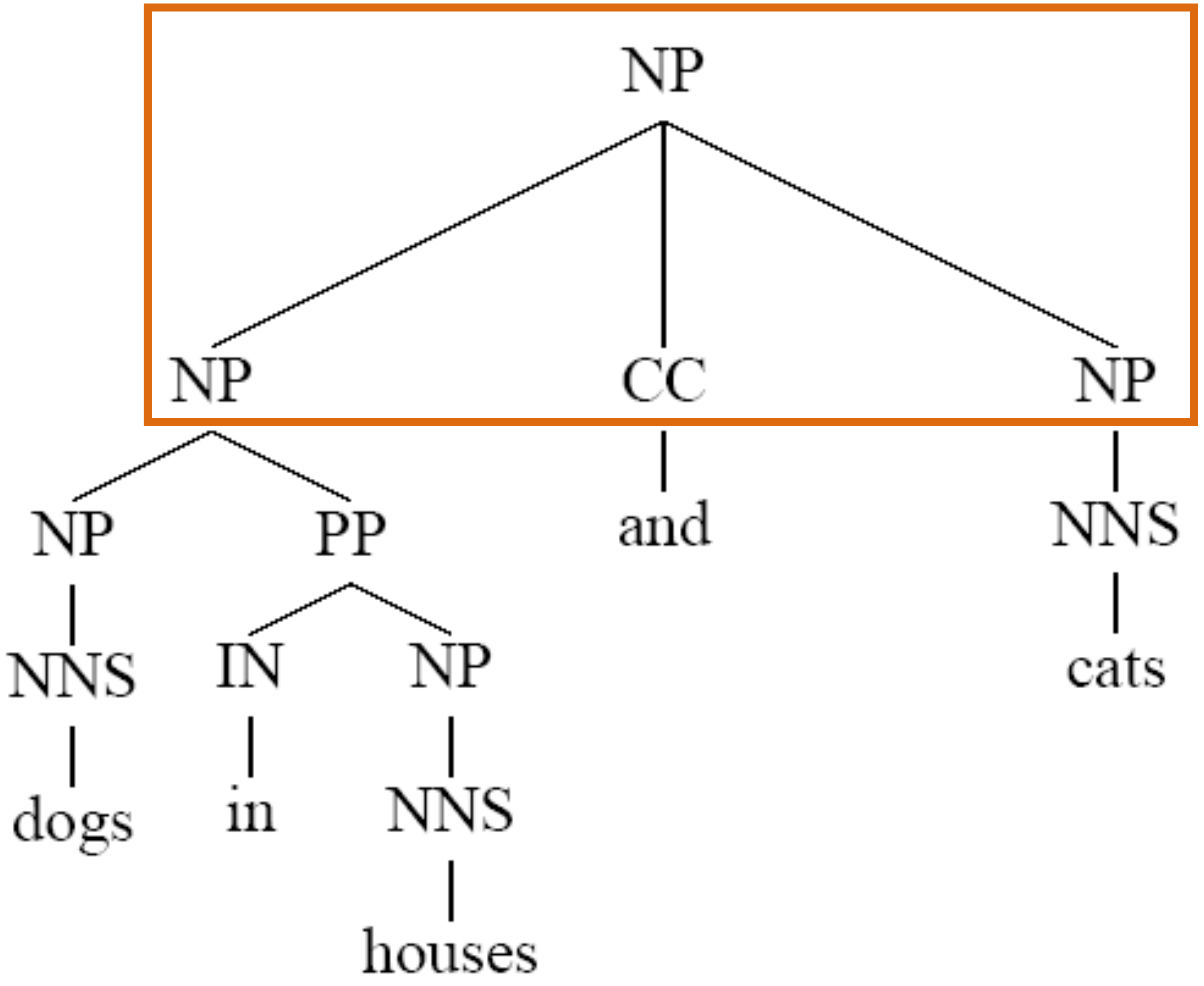


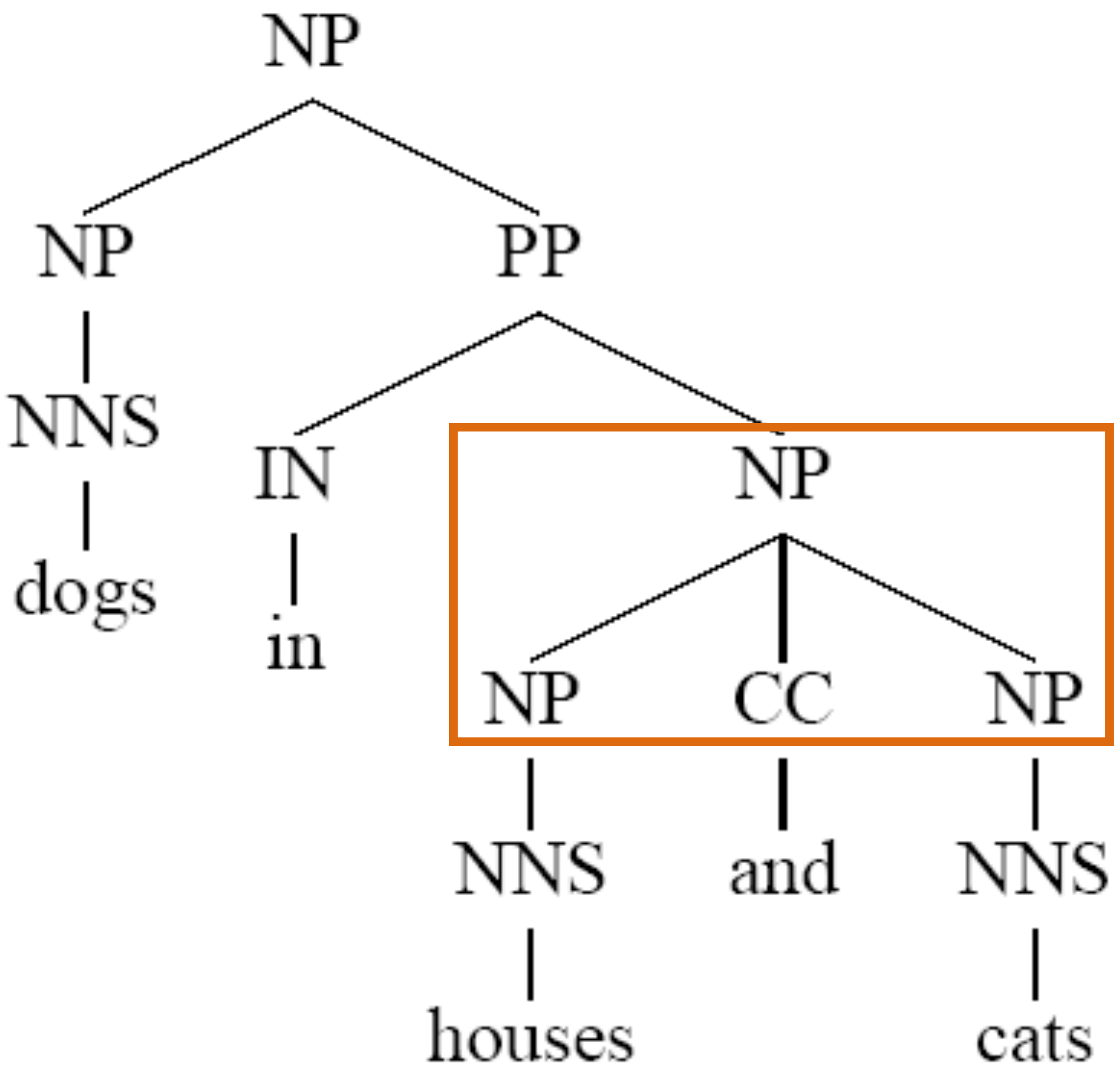
the children ate the cake with a spoon

- More predicate-argument focused view of syntax
- “What’s the main verb of the sentence? What is its subject and object?” — easier to answer under dependency parsing

Dependency vs. Constituency: Coordination

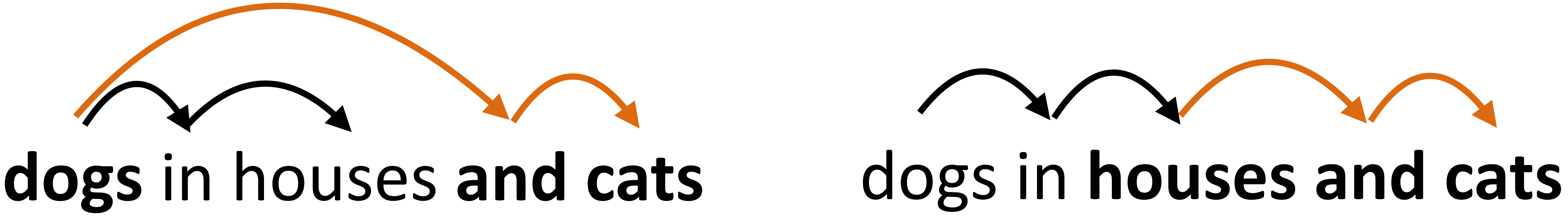
- Constituency: ternary rule NP -> NP CC NP





Dependency vs. Constituency: Coordination

- Dependency: first item is the head



dogs in houses **and cats**

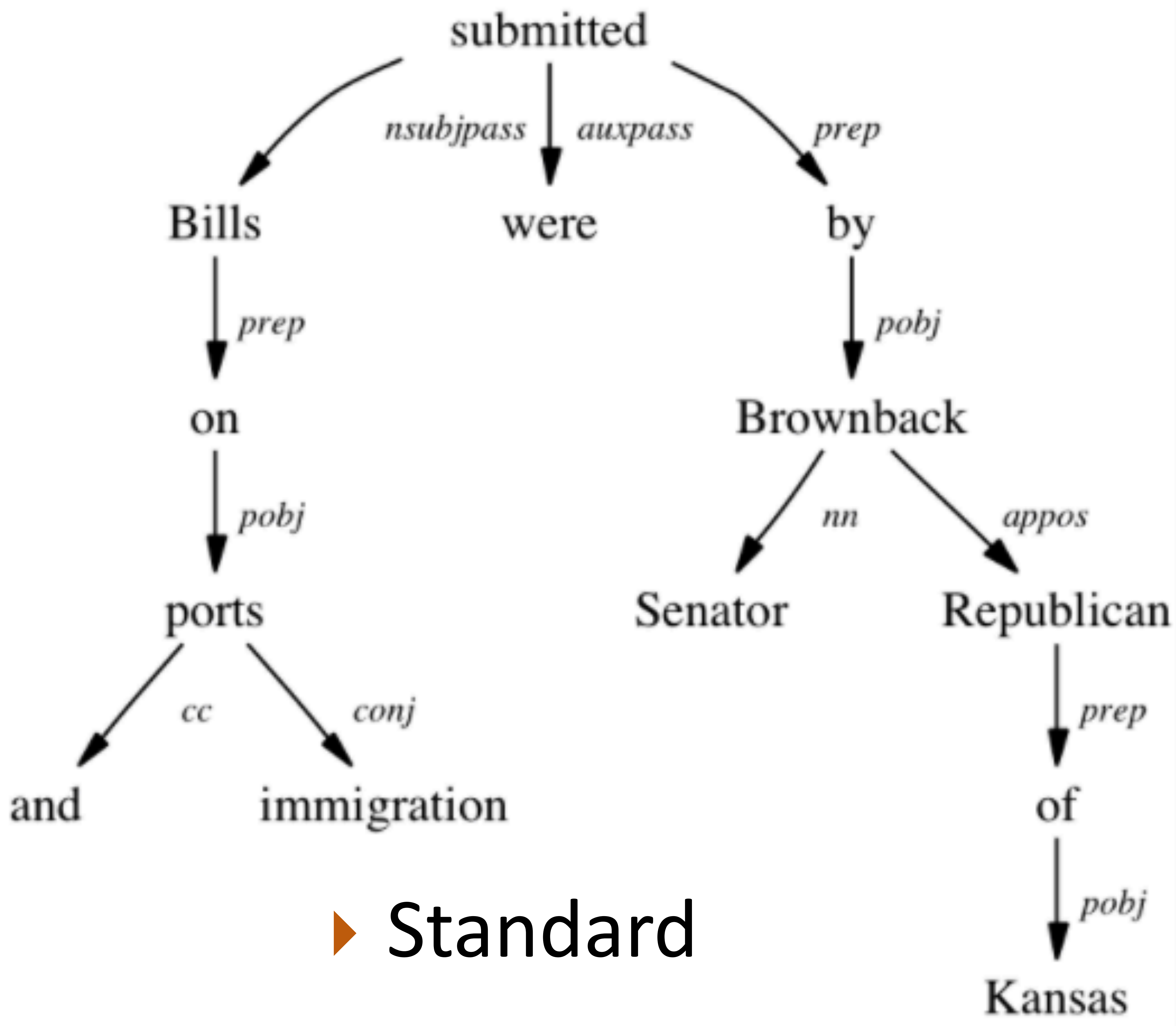
dogs in **houses and cats**

- Coordination is decomposed across a few arcs as opposed to being a single rule production as in constituency
- Can also choose *and* to be the head
- Both cases: headword doesn't really represent the phrase

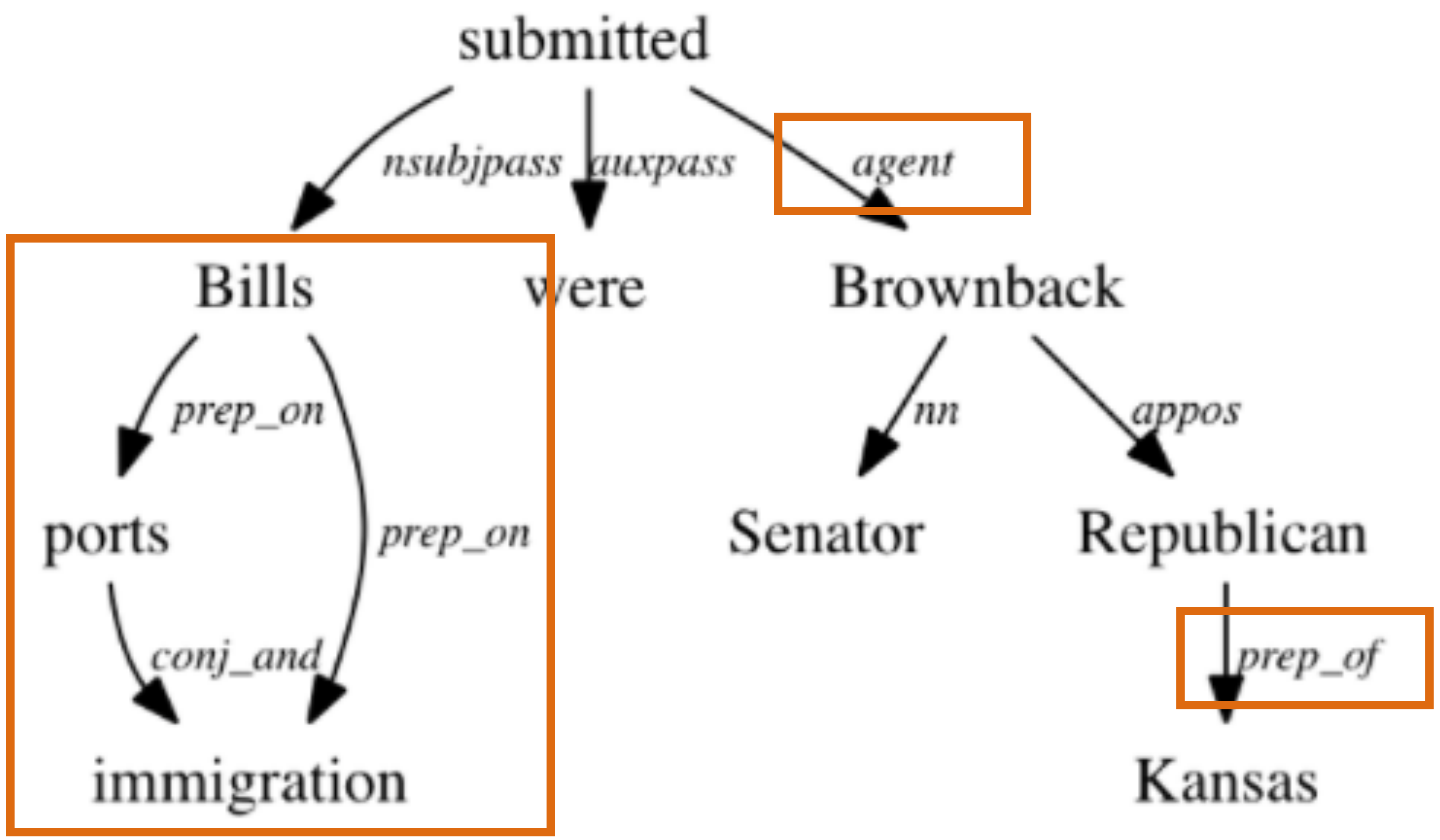
Stanford Dependencies

- Designed to be practically useful for relation extraction

Bills on ports and immigration were submitted by Senator Brownback, Republican of Kansas



- Standard



- Collapsed

Dependency vs. Constituency

- Dependency is often more useful in practice (models predicate argument structure)
- Slightly different representational choices:
  - PP attachment is better modeled under dependency
  - Coordination is better modeled under constituency
- Dependency parsers are easier to build: no “grammar engineering”, no unaries, easier to get structured discriminative models working well
- Dependency parsers are usually faster
- Dependencies are more universal cross-lingually

Universal Dependencies

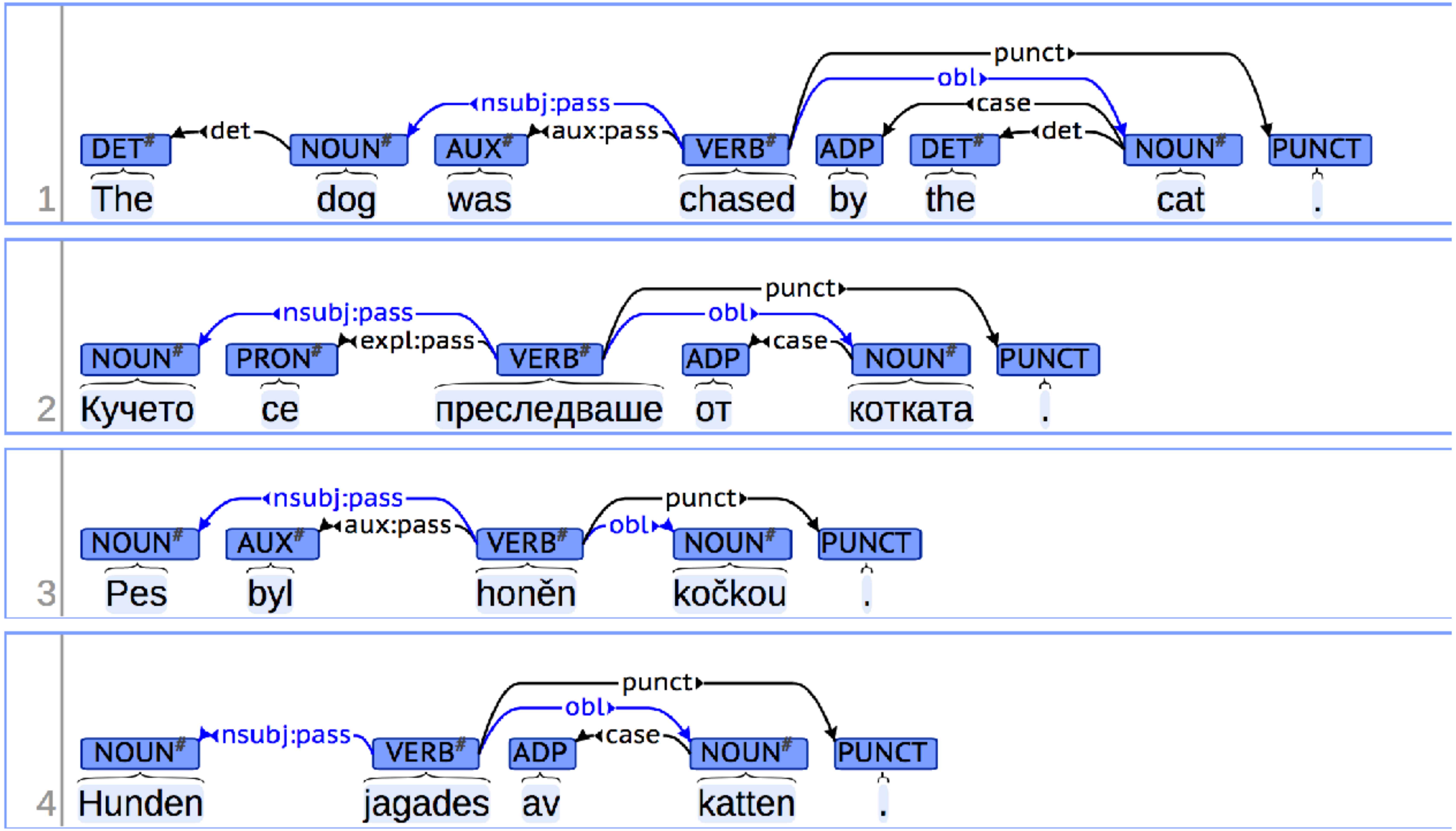
- Annotate dependencies with the same representation in many languages

English

Bulgarian

Czech

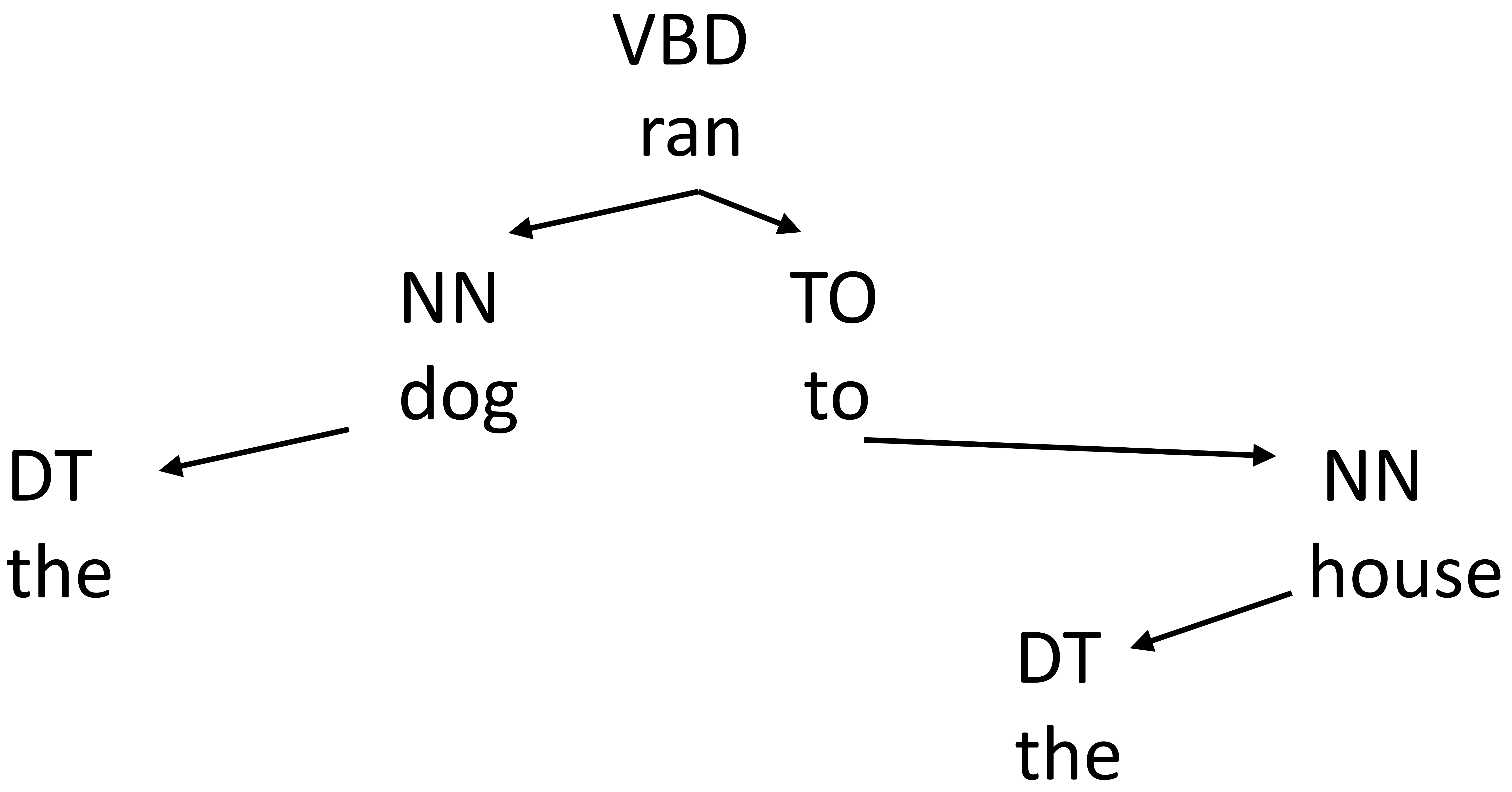
Swiss



http://universaldependencies.org/

Projectivity

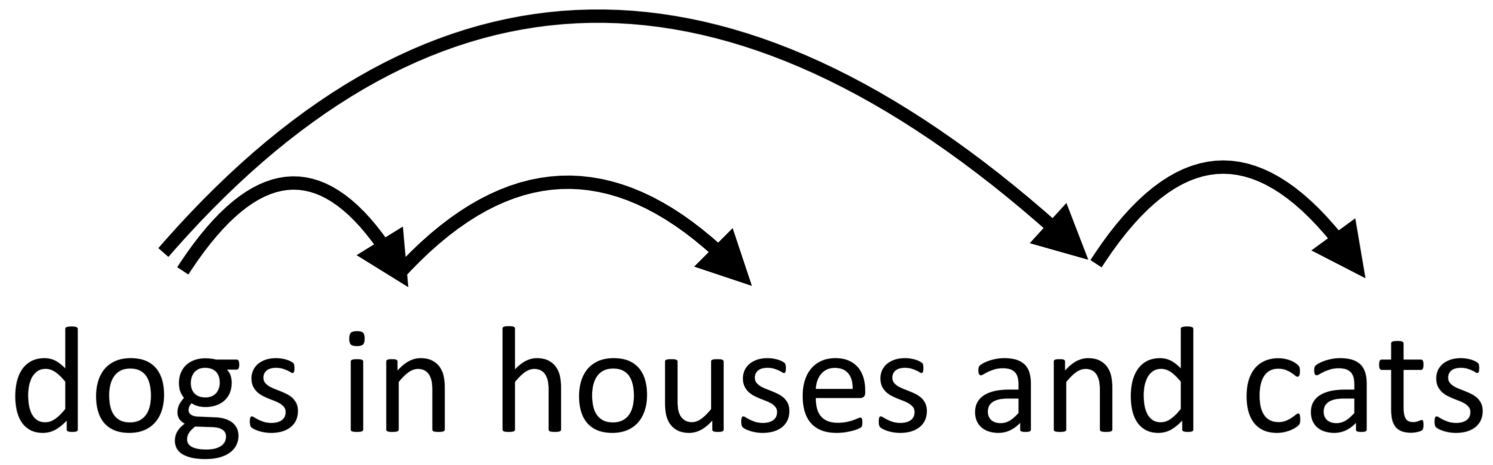
- What conditions have to hold for things to be tree-shaped?



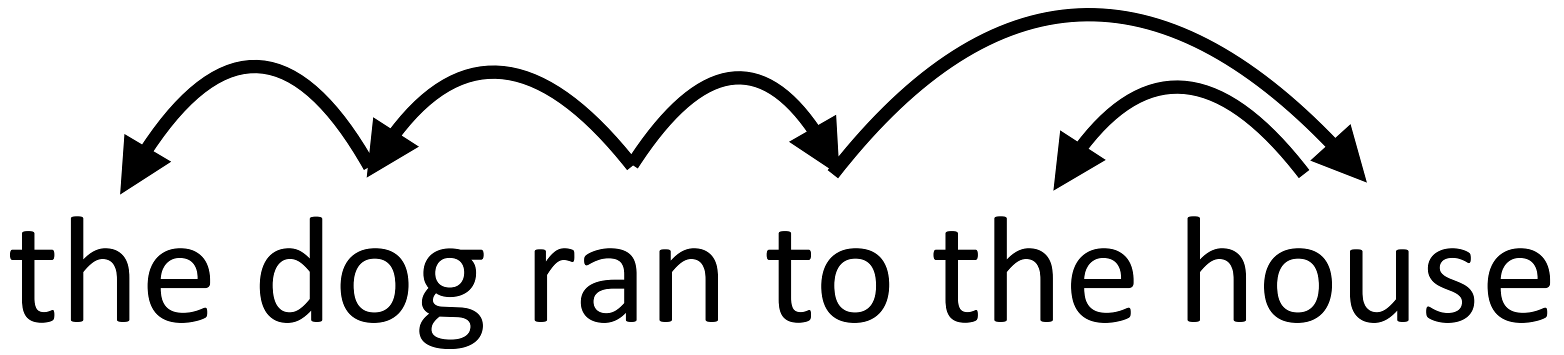
- Any subtree is a contiguous span of the sentence <-> tree is *projective*

Projectivity

- Projective <-> no “crossing” arcs

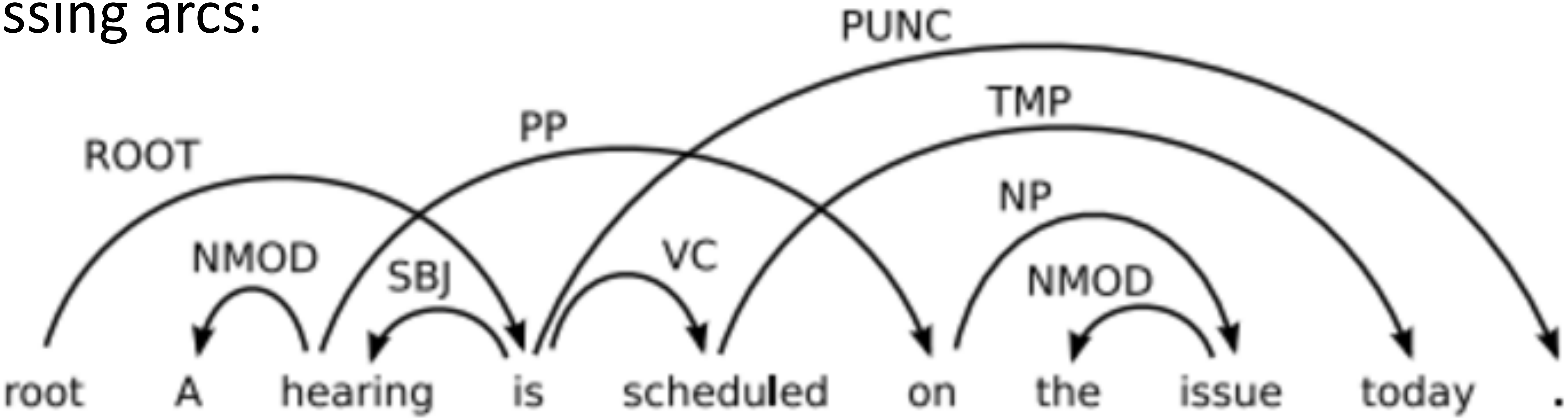


dogs in houses and cats



the dog ran to the house

- Crossing arcs:

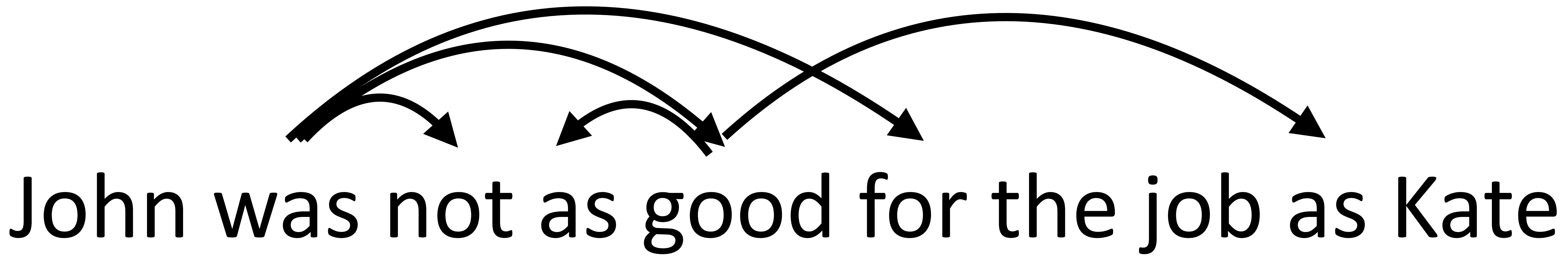


- Extraposition: *A hearing* ***on the issue*** *is scheduled today .* is projective

credit: Language Log

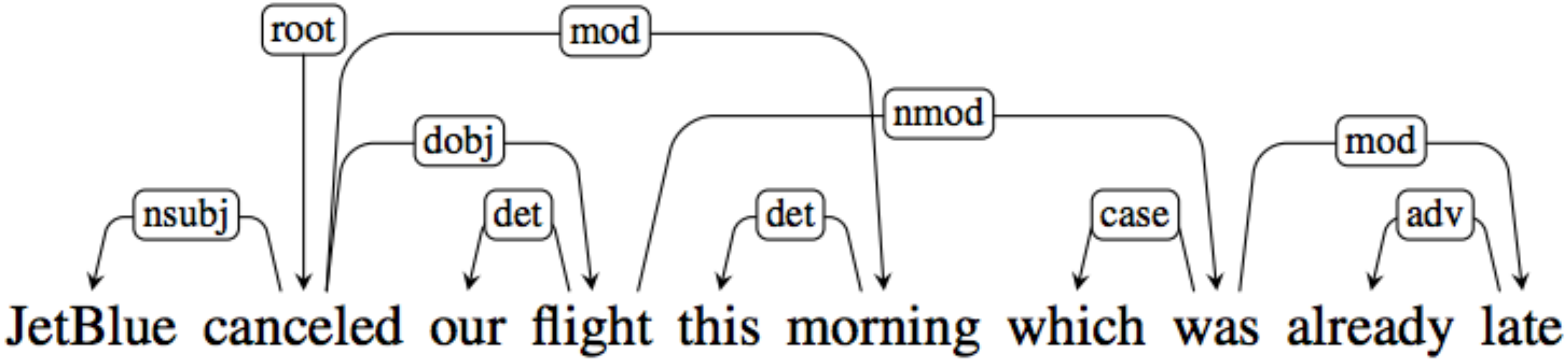
Projectivity

- More extraposition



John was not as good for the job as Kate

- Time expressions can go a lot of places in sentences!

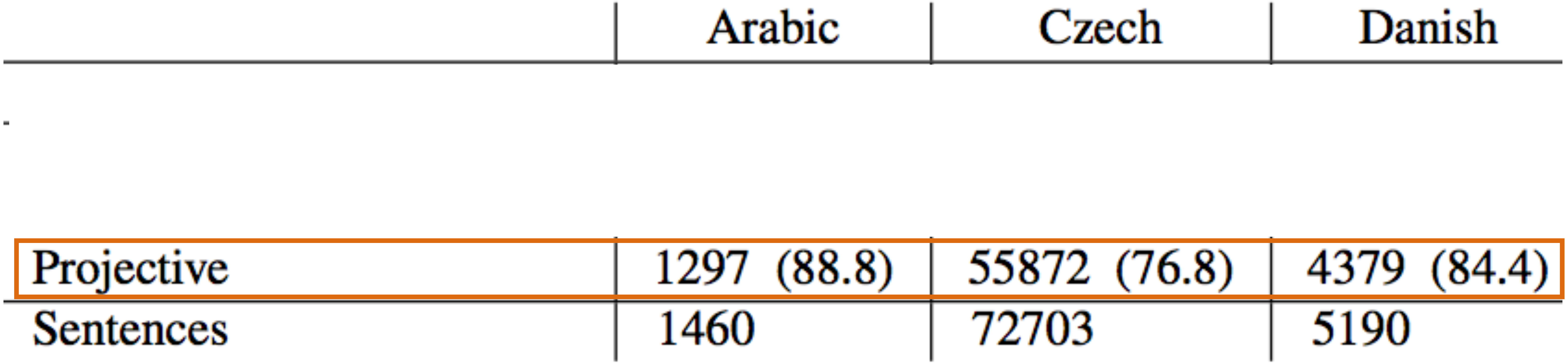


JetBlue canceled our flight this morning which was already late

Gomez-Rodriguez et al.; Jurafsky+Martin

Projectivity

- Number of trees produceable under different formalisms



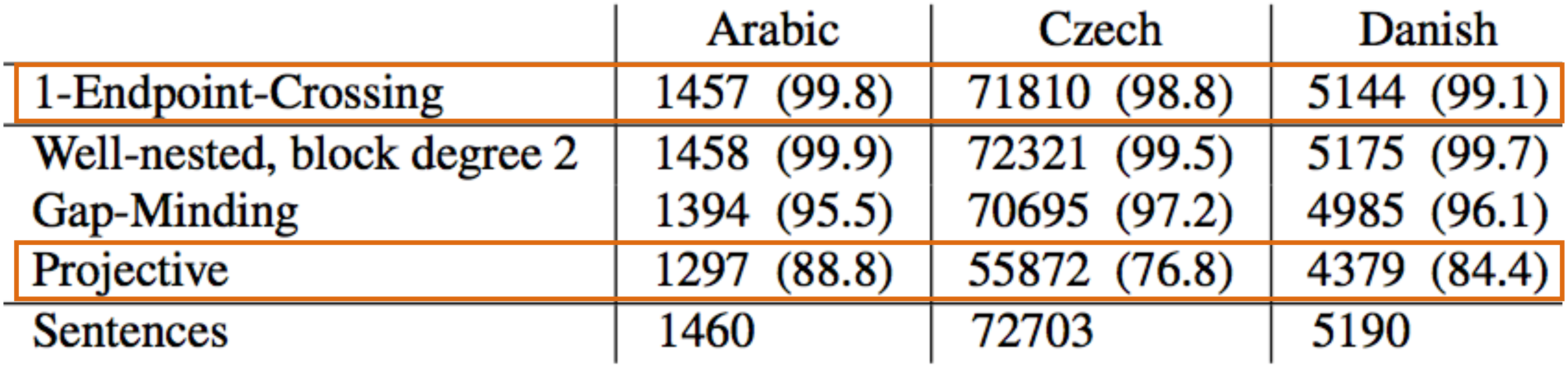
	Arabic	Czech	Danish
Projective	1297 (88.8)	55872 (76.8)	4379 (84.4)
Sentences	1460	72703	5190

- Many trees in other languages are nonprojective

Pitler et al. (2013)

Projectivity

- Number of trees produceable under different formalisms



	Arabic	Czech	Danish
1-Endpoint-Crossing	1457 (99.8)	71810 (98.8)	5144 (99.1)
Well-nested, block degree 2	1458 (99.9)	72321 (99.5)	5175 (99.7)
Gap-Minding	1394 (95.5)	70695 (97.2)	4985 (96.1)
Projective	1297 (88.8)	55872 (76.8)	4379 (84.4)
Sentences	1460	72703	5190

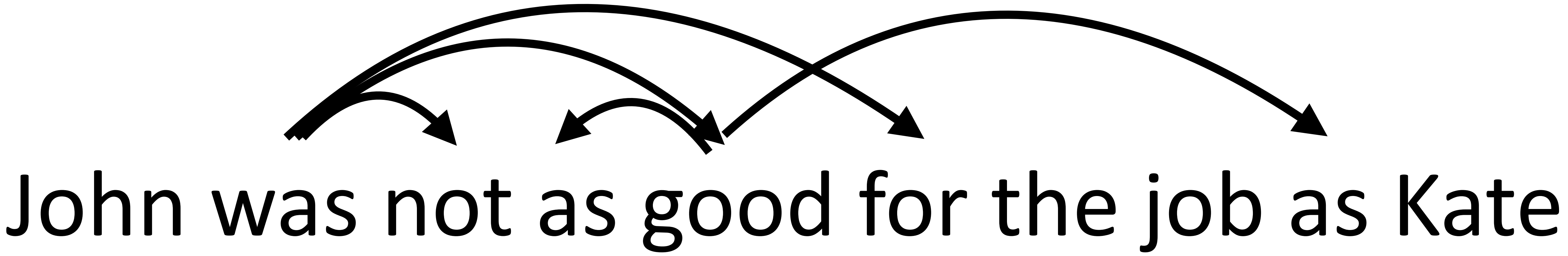
- Many trees in other languages are nonprojective

- Some other formalisms (that are harder to parse in), most useful one is 1-Endpoint-Crossing

Pitler et al. (2013)

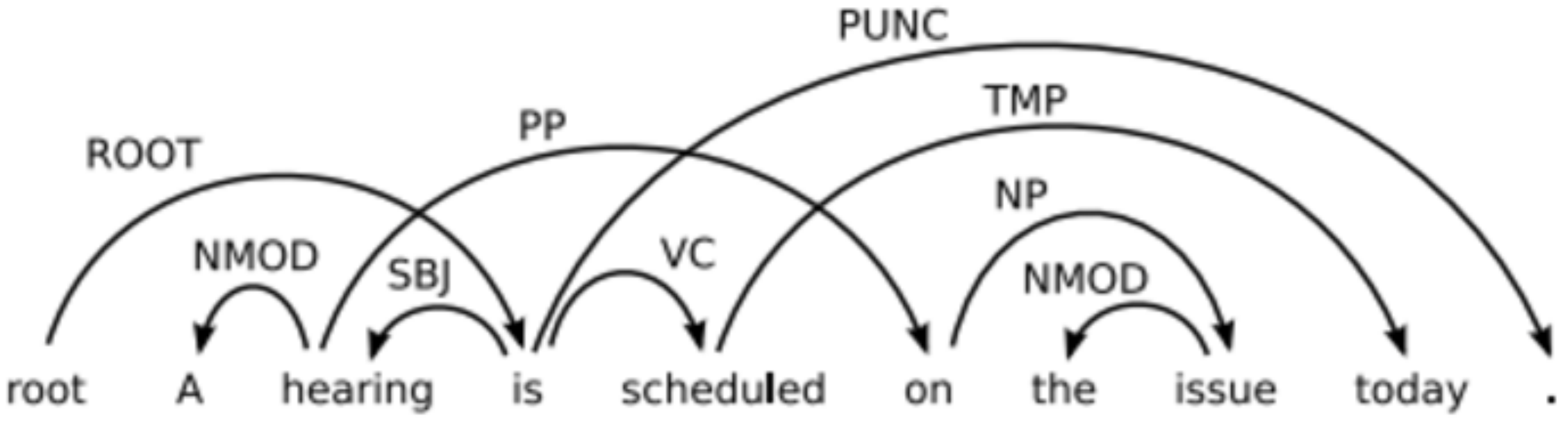
Projectivity

- 1-Endpoint-Crossing: for any edge, all edges that cross it share an endpoint



John was not as good for the job as Kate

- True



- False:
  hearing -> on

- Captures most cases, still efficient parsing algorithms