



BEYOND WORDS

Sandeep Soni

11/06/2023

WORDS

They can fish

WORDS

The sentence is about the
ability to fish

They can fish

... Or the sentence is about
storing fish

PARTS OF SPEECH

Categories of words that have similar morphological and syntactic properties

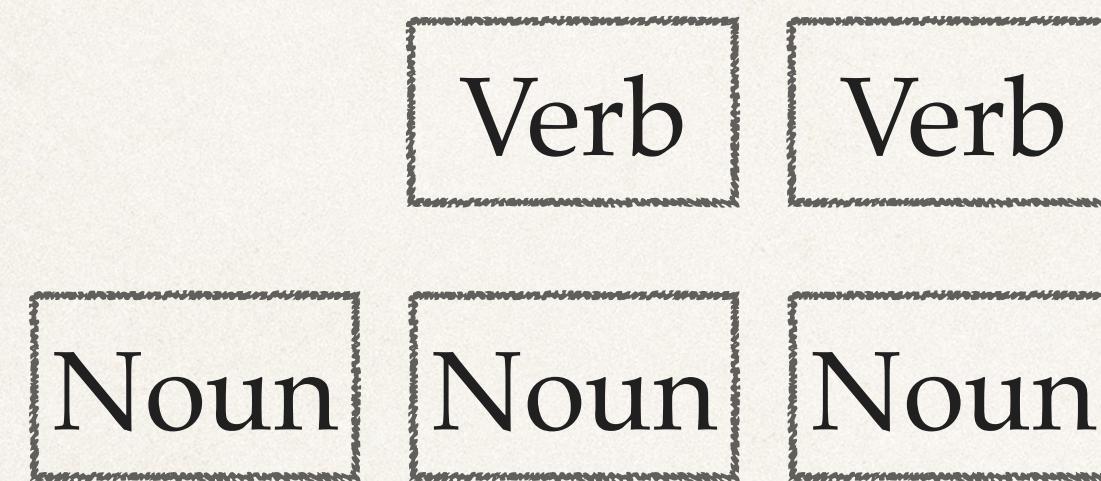
Verb Verb

Noun Noun Noun

They can fish

CONTEXTUAL DEPENDENCE

To find the syntactic role that a word has in a sentence,
more context is helpful



They can fish



They can fish in a cold room

PART OF SPEECH TAGGING

The task of assigning a tag (category) to every word in a sequence.

PRP VB NN IN DT JJ NN

They can fish in a cold room

TAG SET

The number of categories is dependent on our choice,
although there are some core categories

PRP

VB

NN

IN DT

JJ

NN

They can fish in a cold room

Nouns	People, places, things, dates, etc. Depend on quantity (singular Vs Plural)
Adjectives	Properties or qualities of noun phrases
Verbs	Actions, processes. Depends on tense, number, person, etc
Adverbs	Modifiers of verbs; qualify the action (e.g., quickly stopped)
Determiner	Beginning of a noun phrase; emphasizes specificity (e.g., a film)
Prepositions	Indicates spatial/direction/temporal relationship with a noun phrase (e.g., in a room)
Conjunctions	Connectives between phrases, clauses, sentences (e.g., Jack and Jill)
Pronouns	References to noun phrases (e.g., he, she, it)

SEQUENCE LABELING

- Classification: Predict the tag for each word independently
- Structured prediction: Predict the entire tag sequence

The old man the boat

$$f([x_1, x_2, \dots, x_n]) \rightarrow [y_1, y_2, \dots, y_n]$$

Where can part of speech tagging be used?

Sarcasm as Contrast between a Positive Sentiment and Negative Situation

**Ellen Riloff, Ashequl Qadir, Prafulla Surve, Lalindra De Silva,
Nathan Gilbert, Ruihong Huang**

School Of Computing

University of Utah

Salt Lake City, UT 84112

{riloff,asheq,alnds,ngilbert,huangrh}@cs.utah.edu, prafulla.surve@gmail.com

- (a) Oh how I love *being ignored*. #sarcasm
- (b) Thoroughly enjoyed *shoveling the driveway* today! :) #sarcasm
- (c) Absolutely adore it when *my bus is late* #sarcasm
- (d) I'm so pleased mom *woke me up* with *vacuuming my room* this morning. :) #sarcasm

Identified positive sentiment phrases and negative situation phrases using part of speech tags to build a sarcasm recognizer

OTHER SEQUENCE LABELING TASKS

Apple was founded on April 1, 1976, by Steve Jobs in California.

OTHER SEQUENCE LABELING TASKS

Apple was founded on April 1, 1976, by Steve Jobs in California

NAMED ENTITY RECOGNITION

Person	Proper nouns
Date	References to dates
Currency	Named reference to money (e.g., 5 dollars)
Location	Reference to any location (e.g., Rich memorial building)
GPE	Geo political entity (e.g., London)
Organization	Refers to a company or an organization (e.g. United Nations)

BIO NOTATION

B	Beginning of a named entity
I	Inside of a named entity
O	Outside of the named entity

B-ORG O O O B-DATE I-DATE O B-PER I-PER O B-LOC
Apple was founded on April 1976, by Steve Jobs in California

NAMED ENTITY RECOGNITION

Apple **ORG** was founded on April 1, 1976 **DATE**, by Steve Jobs **PERSON** in California **GPE**.

- Named entity recognition can be casted as a sequence labeling tasks
- Entities can be coarse- or fine-grained

PARSING

- Beyond annotating every word with a category or mapping spans to categories, more linguistic structure is enforced by how spans are combined in a sequence

GRAMMAR

$S \rightarrow NP\ VP$

$VP \rightarrow V\ NP \mid VP\ PP$

$PP \rightarrow IN\ NP$

$V \rightarrow eat$

$IN \rightarrow with$

$NP \rightarrow NP\ PP \mid we \mid sushi \mid chopsticks$

These are called the production rules.
Infinite sequences can be generated
by following these rules

SYNTAX TREE

$S \rightarrow NP\ VP$

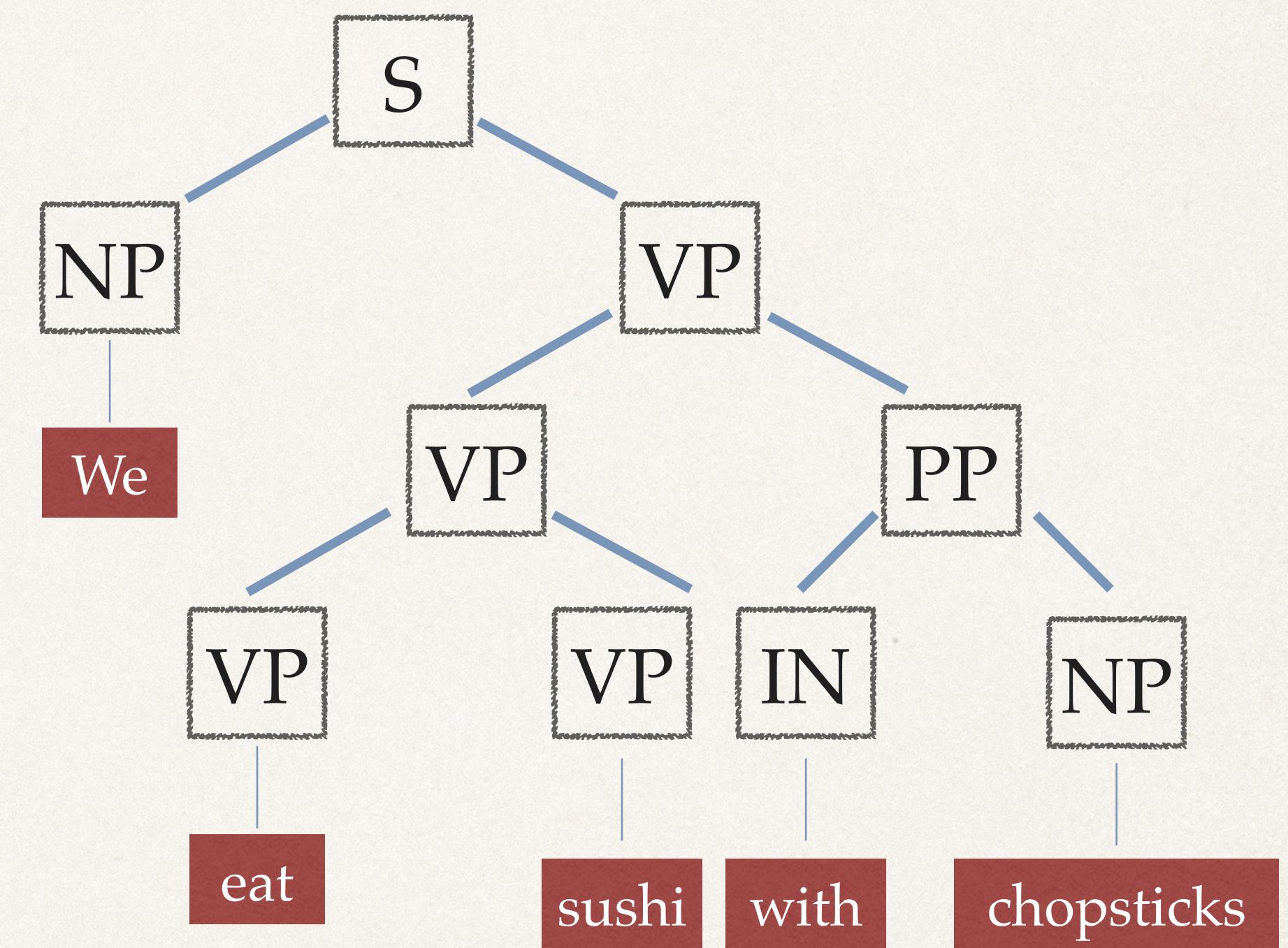
$VP \rightarrow V\ NP \mid VP\ PP$

$PP \rightarrow IN\ NP$

$V \rightarrow eat$

$IN \rightarrow with$

$NP \rightarrow NP\ PP \mid we \mid sushi \mid chopsticks$



PARSING

$S \rightarrow NP\ VP$

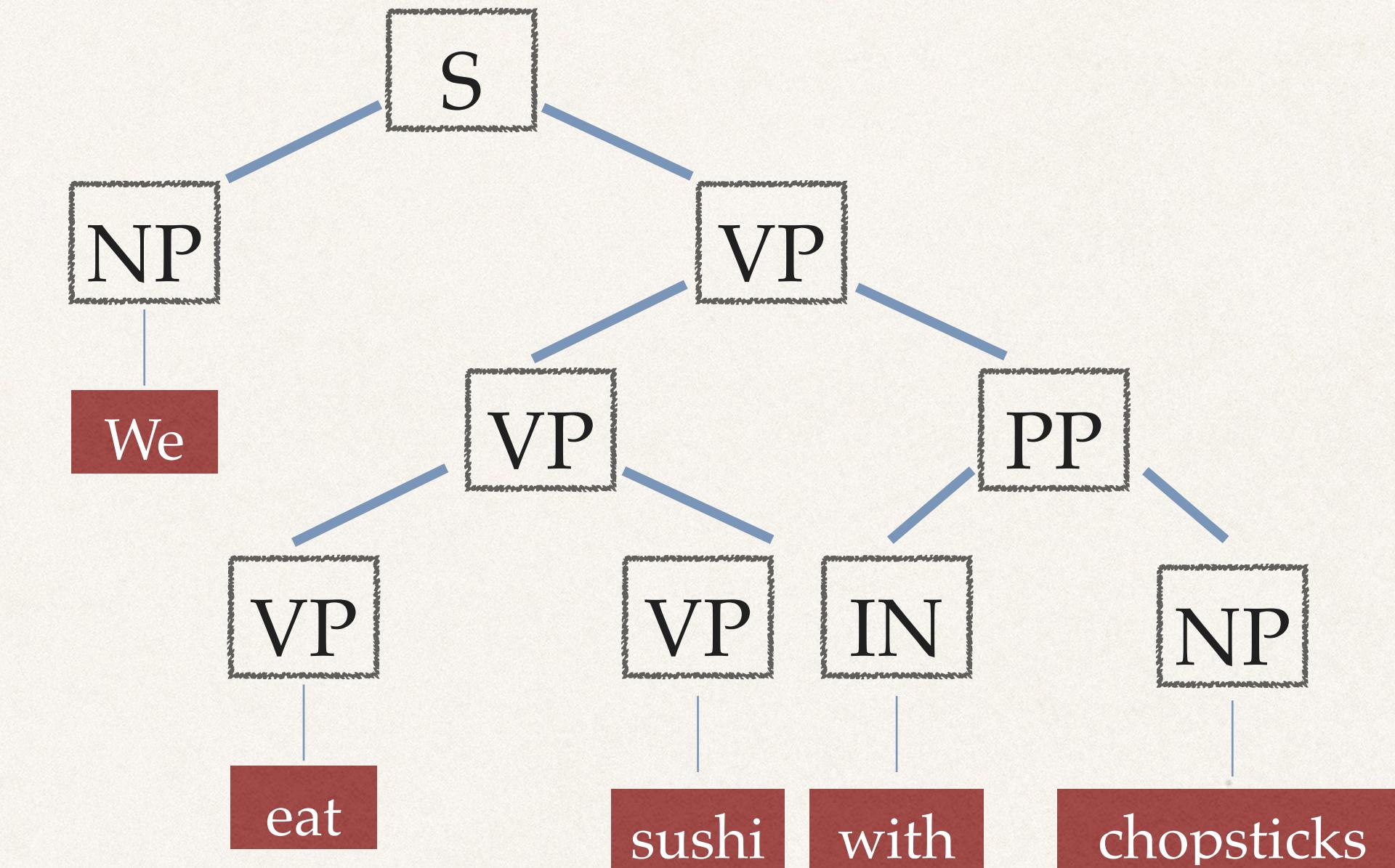
$VP \rightarrow V\ NP \mid VP\ PP$

$PP \rightarrow IN\ NP$

$V \rightarrow eat$

$IN \rightarrow with$

$NP \rightarrow NP\ PP \mid we \mid sushi \mid chopsticks$

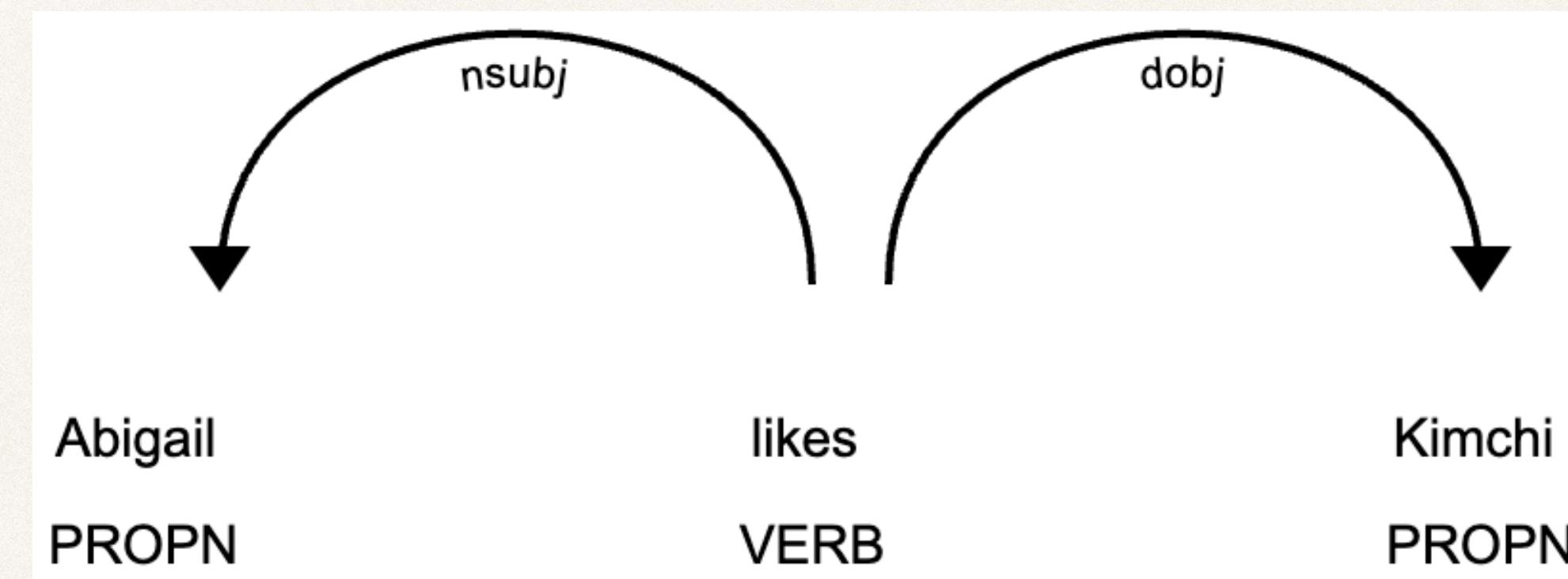


The goal of parsing is to learn the production rules and infer the most likely parse tree for a given sequence

DEPENDENCY PARSE

- Instead of decomposing a sequence into constituent subtrees, one can markup the sequence by labeling:
 - Important constituents (heads)
 - Their modifiers (dependents)
 - The type of relationship between the heads and dependents

EXAMPLE

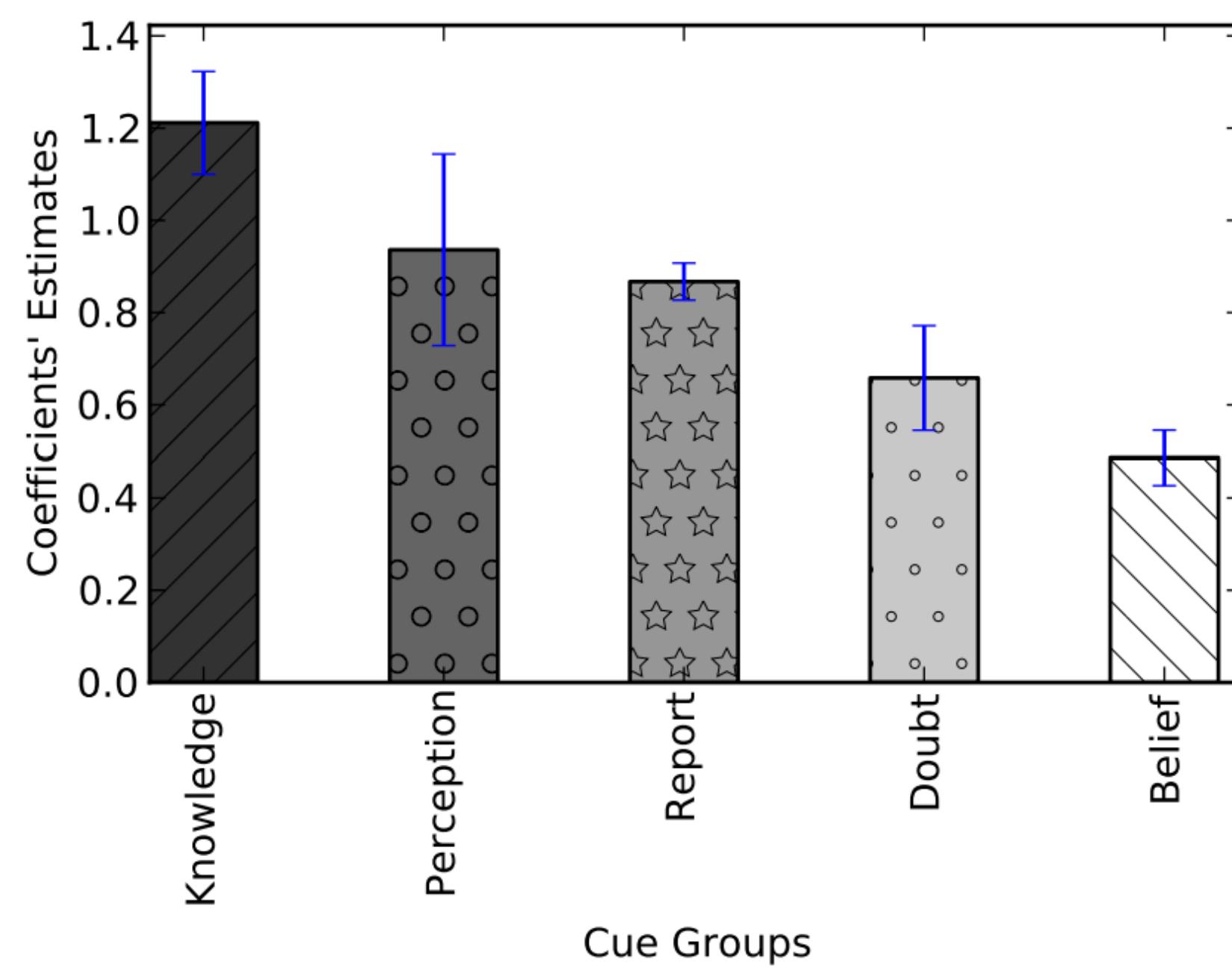
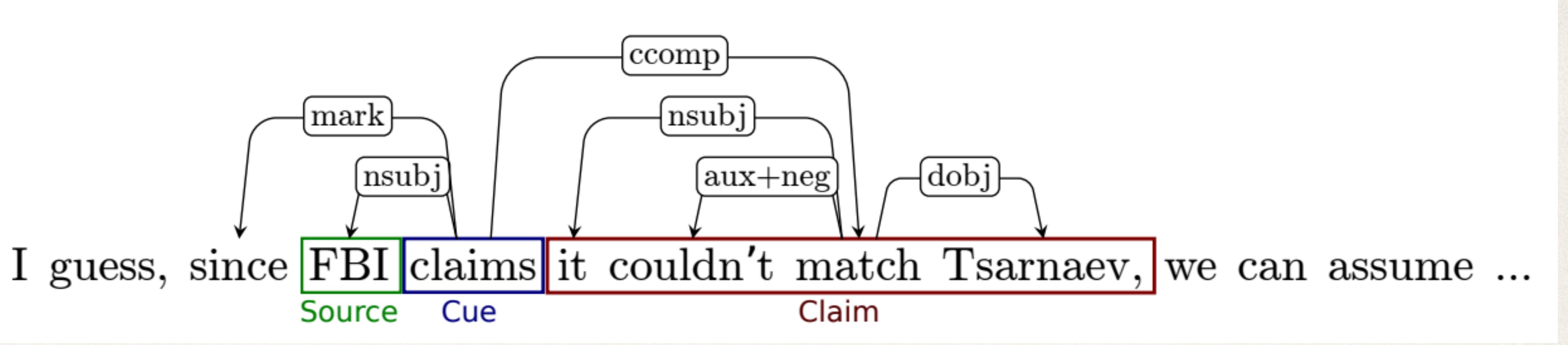


- What is happening? Something is being liked
- Who is liking? Abigail (the subject)
- What is being liked? Kimchi (the object)

Core dependents of clausal predicates	Non-core dependents of clausal predicates	Special clausal dependents
<p><i>Nominal dep</i> <i>Predicate dep</i></p> <p><u>nsubj</u> <u>csubj</u></p> <p>↳ <u>nsubj:pass</u> ↳ <u>csubj:pass</u></p> <p>↳ <u>nsubj:outer</u> ↳ <u>csubj:outer</u></p> <p><u>obj</u> <u>ccomp</u></p> <p><u>iobj</u></p>	<p><i>Nominal dep</i> <i>Predicate dep</i> <i>Modifier word</i></p> <p><u>obl</u> <u>advcl</u> <u>advmod</u></p> <p>↳ <u>obl:nmod</u> ↳ <u>advcl:relcl</u></p> <p>↳ <u>obl:tmod</u></p>	<p><i>Nominal dep</i> <i>Auxiliary</i> <i>Other</i></p> <p><u>vocative</u> <u>aux</u> <u>mark</u></p> <p><u>discourse</u> ↳ <u>aux:pass</u></p> <p><u>expl</u> <u>cop</u></p>
<p>Noun dependents</p> <p><i>Nominal dep</i> <i>Predicate dep</i> <i>Modifier word</i></p> <p><u>nummod</u> <u>acl</u> <u>amod</u></p> <p>↳ <u>acl:relcl</u></p> <p><u>appos</u> <u>det</u></p> <p>↳ <u>det:predet</u></p> <p><u>nmod</u></p> <p>↳ <u>nmod:nmod</u></p> <p>↳ <u>nmod:tmod</u></p> <p>↳ <u>nmod:poss</u></p>	<p>Compounding and unanalyzed</p> <p><u>compound</u> <u>flat</u></p> <p>↳ <u>compound:prt</u> ↳ <u>flat:foreign</u></p> <p><u>fixed</u> <u>goeswith</u></p>	<p>Coordination</p> <p><u>conj</u> <u>cc</u></p> <p>↳ <u>cc:preconj</u>.</p>
<p>Case-marking, prepositions, possessive</p> <p><u>case</u></p>	<p>Loose joining relations</p> <p><u>list</u> <u>parataxis</u> <u>orphan</u></p> <p><u>dislocated</u> <u>reparandum</u></p>	<p>Other</p> <p><i>Sentence head</i> <i>Punctuation</i> <i>Unspecified dependency</i></p> <p><u>root</u> <u>punct</u> <u>dep</u></p>

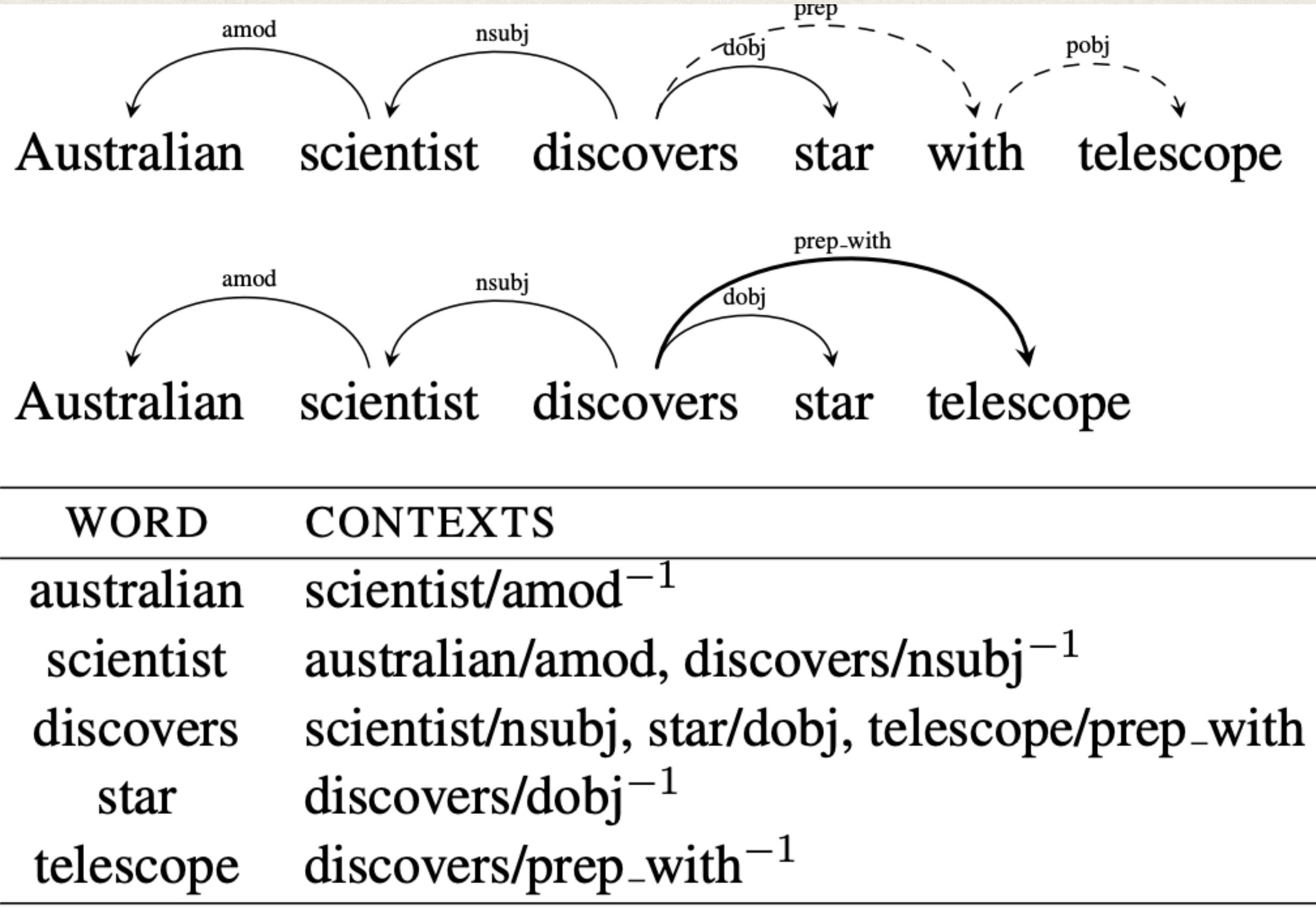
There are many syntactic dependencies that can be identified (example taken from universal dependencies treebank)

What can we do with dependency parse?



Using the dependency parse, we can identify the different parts of a sentence

- Sandeep Soni, Tanushree Mitra, Eric Gilbert, and Jacob Eisenstein. 2014. [Modeling Factuality Judgments in Social Media Text](#). In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 415–420, Baltimore, Maryland. Association for Computational Linguistics.



florida	gainesville fla jacksonville tampa lauderdale	fla alabama gainesville tallahassee texas	texas louisiana georgia california carolina
---------	---	---	---

Similar words (last column is based on dependent parsing contexts)

COREFERENCE RESOLUTION

- Monica told Rachel that she upset Phoebe
- What does she refer to?

COREFERENCE RESOLUTION

- Reference resolution is the problem of linking spans of text within the same sequence that refer to the same entity
- Identifying the syntactic and semantic roles can help in reference resolution

What can we do with if we knew how to resolve references?

Description of women, as a percentage of characterization,
broken out by author gender

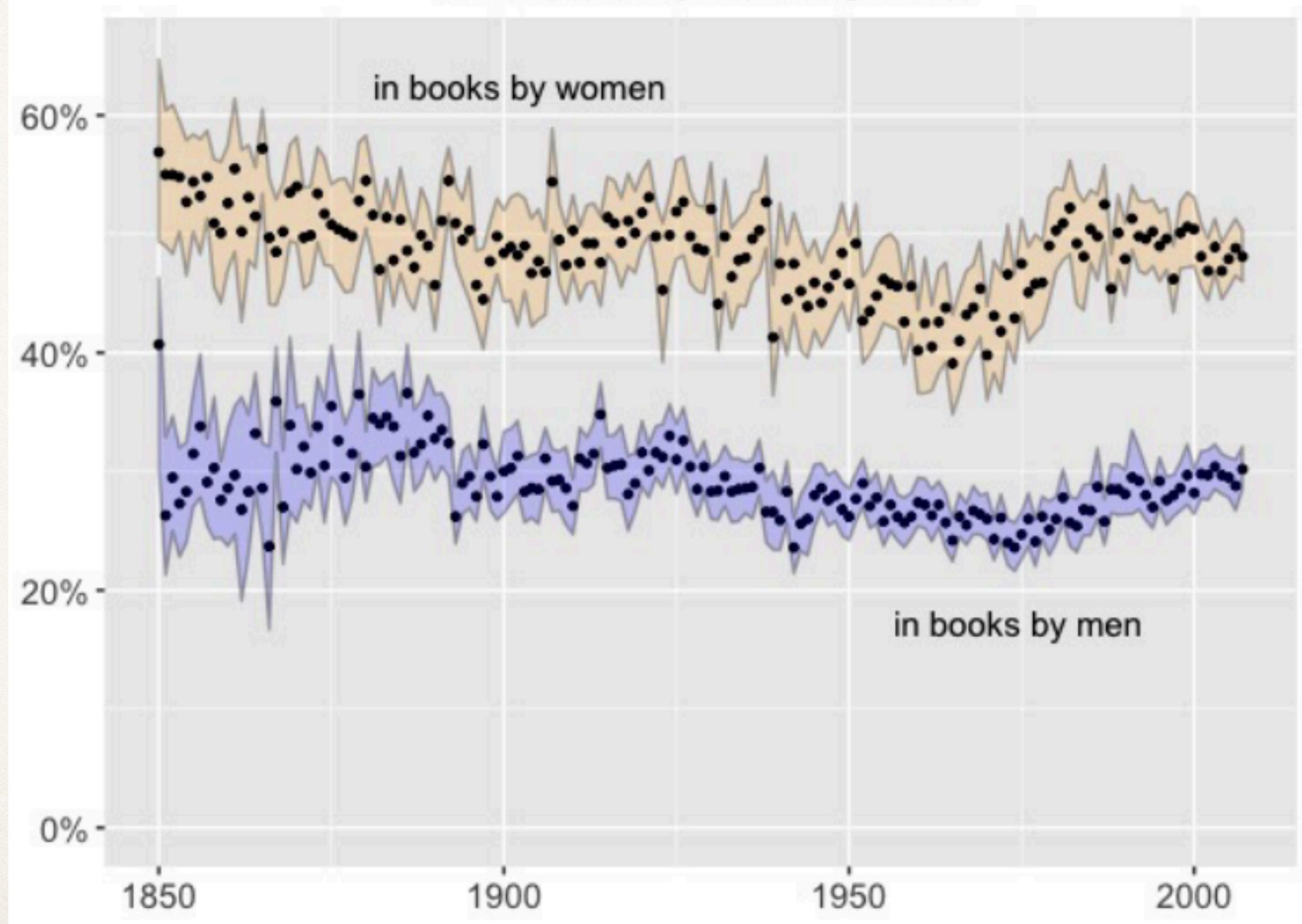


Figure taken from Underwood et. al. 2018

IN CLASS

- Sequence labeling demo