

# Computational Linguistics (INF2820 - Morphology) 

\{ eat, eats, eating, ate, eaten \}

Stephan Oepen
Universitetet i Oslo \& CSLI Stanford

```
oe@ifi.uio.no
```


## Some Areas of Descriptive Grammar

$$
\begin{array}{ll}
\text { Phonetics } & \text { The study of speech signals. } \\
\hline \text { Phonology } & \text { The study of sound systems. } \\
\hline \text { Morphology } & \text { The study of word structure. }
\end{array}
$$

Syntax $\quad$ The study of sentence structure.

Semantics The study of language meaning.

$$
\text { Pragmatics } \quad \text { The study of language use. }
$$

$\qquad$
Computational Linguistics at Work (2)

## Some Areas of Descriptive Grammar

## Phonetics The study of speech signals.

Phonology The study of sound systems.

Morphology $\quad$ The study of word structure.

Syntax $\quad$ The study of sentence structure.

Semantics The study of language meaning.

## Pragmaties <br> The study of language use.

## Tokenization: Finding the Basic Building Blocks

## Morphological Categories (1 of 3)

Parts of Speech (PoS)

| cat, dog, neighbours, ... | noun (N) |
| :--- | ---: |
| barks, chased, was, ... | verb (V) |
| fierce, angry, black, young, ... | adjective (Adj) |
| quickly, probably, not, ... | adverb (Adv) |
| a, the, my, that, ... | determiner (D) |
| of, by, on, at, under, ... | preposition (P) |
| she, mine, those, what, ... | pronoun (Pro) |
| and, neither ... nor, because, ... | conjunction (C) |

How to discover the inventory of categories?

## A Quick Tour of English Morphology

## Morphological Categories (2 of 3)

Parts of Speech (PoS)
cat, dog, neighbours, ...
barks, chased, was, ...
fierce, angry, black, young, ...
quickly, probably, not, ...
a, the, my, that, ...
of, by, on, at, under, ...
she, mine, those, what, ...
and, neither ... nor, because, ...
noun (N) verb (V) adjective (Adj) adverb (Adv) determiner (D) preposition (P) pronoun (Pro) conjunction (C)

- Paradigm set of word forms, e.g. \{bark, barks, barking, barked \};
- Unit Categories dimensions structuring a paradigm internally;
- Paradigm Categories properties common to all paradigm units.


## Morphological Categories (3 of 3)

## Number - Person - Case - Gender

That dog barks. - Those dogs bark.
I bark. — You bark. - They bark. - Sam shaved himself.
We bark. - You bark. - Those dogs bark.
I saw her. - She saw me. - My dog barked.

How many distinct verb forms according to number and person?

## Tense - Aspect - Mood

The dog barks. - The dog barked - The dog will bark. The dog has barked. - The dog is barking. If I were a carpenter, ...
$\qquad$

