

Advanced programming in Python

Lecture 12: Python in Natural Language Processing

Arianna Masciolini

Chalmers/GU CSE (DAT515/DIT515)

- ❖ PhD student in Natural Language Processing (or Computational Linguistics, or Language Technology...) at Språkbanken Text, GU
- ❖ particularly interested in:
 - ❖ computational syntax
 - ❖ Computer-Assisted Language Learning

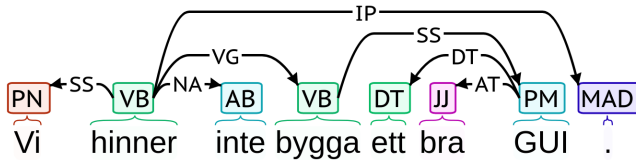
- ❖ PhD student in Natural Language Processing (or Computational Linguistics, or Language Technology. . .) at Språkbanken Text, GU
- ❖ particularly interested in:
 - ❖ computational syntax
 - ❖ Computer-Assisted Language Learning
 - ❖ Swedish

We are going to build a mini Python application to practice Swedish syntax (adverb placement), but...

Vi inte hinner
bygga ett bra GUI.

"We don't have time to build a proper GUI"

With **Part-Of-Speech** annotation and **dependency parsing**:



With **Part-Of-Speech** annotation and **dependency parsing**:



With **Part-Of-Speech annotation** and **dependency parsing**:

Vi **hinner** bygga ett bra GUI .

With **Part-Of-Speech** annotation and **dependency parsing**:



With **Part-Of-Speech** annotation and **dependency parsing**:



- ❖ **python-korp**, a wrapper for the Korp API, to retrieve linguistically pre-annotated sentences from large Swedish corpora
- ❖ **Spacy**, a widely used NLP library, to annotate our own sentences

Korp

1. get a random concordance containing an adverb
2. generate a question:
 - 2.1 set aside the adverb
 - 2.2 highlight its head
3. ask the user to input a 3+ word-long sequence *including* the adverb
4. check whether the sequence is included in the original sentence

Spacy

If the path to a custom input file is provided:

1. **read the file and split it into sentences** (we can assume that there is one sentence per line)
2. **parse all sentences with Spacy**
3. get a random concordance containing an adverb (**new query and matching function!**)
4. **transform it to a korp-like dict with to_korp_dict**

Otherwise:

1. get a random Korp concordance containing an adverb

Then, generate a question and check the user answer as usual.

2h-prototype of a simple language learning app:

- ❖ code: `github.com/harisont/advpy_nlp`
- ❖ possible extensions:
 - ❖ GUI
 - ❖ other POS/dependency relations (customizable queries)
 - ❖ other languages (with Universal Dependencies)
 - ❖ ...

- ❖ Korp user manual:
`spraakbanken.gu.se/en/tools/korp/user-manual`
- ❖ Official Spacy website: `spacy.io`
- ❖ the code: `github.com/harison/advpy_nlp`
- ❖ the exercise class tomorrow
- ❖ Aarne's Computational Syntax course