LING 300 - Topics in Linguistics:
Introduction to Programming and Text Processing for Linguists

Week 9

Python for Text (and Beyond)
Roadmap for This Week

Monday

- Content:
  - Dependency Parsing
  - WordNet
- Final Assignment
- Some Applications of What We’ve Learned
  (from my research)

Wednesday

- Content:
  - Word Vectors Classification
- Assignment 6 Notes
- Final Self-Evaluation
- Where To Go From Here
Dependency Parsing gives a syntax representation

- Words are connected to other words with a tag representing their relationship
- Main verb is the sentence root
- Directed: head → dependent
- Tag is role the played by the dependent

https://spacy.io/usage/visualizers
Dependency Parsing gives a syntax representation

- Most common formalism for syntax in Comp Ling / NLP
  - Interesting contrast with formal syntax!
- Partially because of computational feasibility
- Very exciting project: Universal Dependencies
  - https://universaldependencies.org/
  - (you can contribute!)
Dependency Parsing gives a syntax representation

- spaCy does dependency parsing inherently (if you don’t disable "parser")
- Access dependency tag with token.dep_
  List of children with token.children
- More info: https://spacy.io/usage/linguistic-features
WordNet is a lexical resource for semantic relations

- Represents semantic relationships in a large network
- Allows to calculate e.g. “path similarity”
- Play with directly:

http://wordnetweb.princeton.edu/perl/webwn
WordNet is a lexical resource for semantic relations

- NLTK has an interface for working with WordNet
- ... but it’s not the most intuitive thing in the world
- More info here: https://www.nltk.org/howto/wordnet.html
Final Assignment

● More and more on your own! Get creative!

● Key point!
  ○ If you want to use LDC or BYU data, let me know by Wednesday

● Please turn in on time! Next Tuesday EOD, where you’re at
  ○ For grading purposes, but I’m always available to talk more later if you keep working on it!

● Brief walkthrough