Why isn’t linguistics more useful in NLP?

• Be clear about your goals
  ▶ science ≠ engineering

• More linguistic theory is not necessarily better
  ▶ features should be *relevant* to task
  ▶ true ≠ useful for some specific task

• Commonplace linguistic insights are often most useful
  ▶ especially if we want good *average-case* accuracy

• No good models of “world knowledge” or “common-sense reasoning”
  ▶ but they are necessary for understanding language
  ▶ formalizing world knowledge is not a goal of linguistics (nor should it be)
  ▶ lexical statistics (e.g., head-to-head dependencies) are a crude approximation
Lessons from parsing and related applications

- Banal linguistic insights are sometimes very useful
  - words group together to form phrases
  - phrases have head words
  - relative locations within a phrase matter
- Linguists’ insights are sometimes useful, but their representations and formalisms are not
  - machine-learning based approaches map parses to feature vectors
    ⇒ details of parse representations don’t matter (as long as the features can be read off somehow)
- Corpus annotation seems a more economical way of getting linguistic information into a system than writing grammars
  - linguistic grammars are *closed world* (everything not explicitly permitted is ungrammatical)
  - stochastic models are typically *open world* (everything is possible)
Accuracy reduction removing a feature class
Doing feature selection well is hard!

Averaged perceptron feature selection

f-score on sections 20-21

f-score on section 24

0.908 0.906 0.904 0.902 0.900 0.898 0.896 0.894 0.892 0.901 0.902 0.903 0.904 0.905 0.906 0.907 0.908 0.909 0.91 0.911