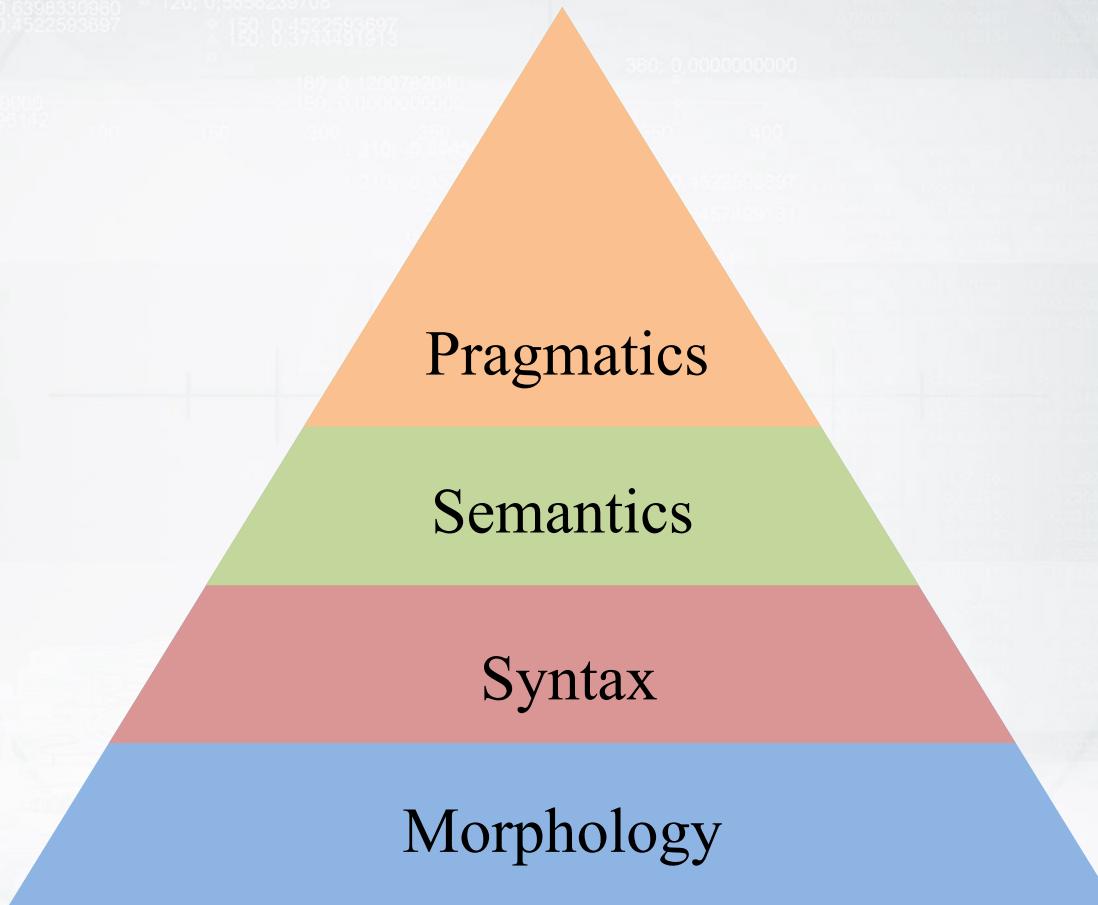




## Linguistic knowledge in NLP

# NLP Pyramid



Natural Language Processing Pyramid

# Libraries and tools

- **NLTK**

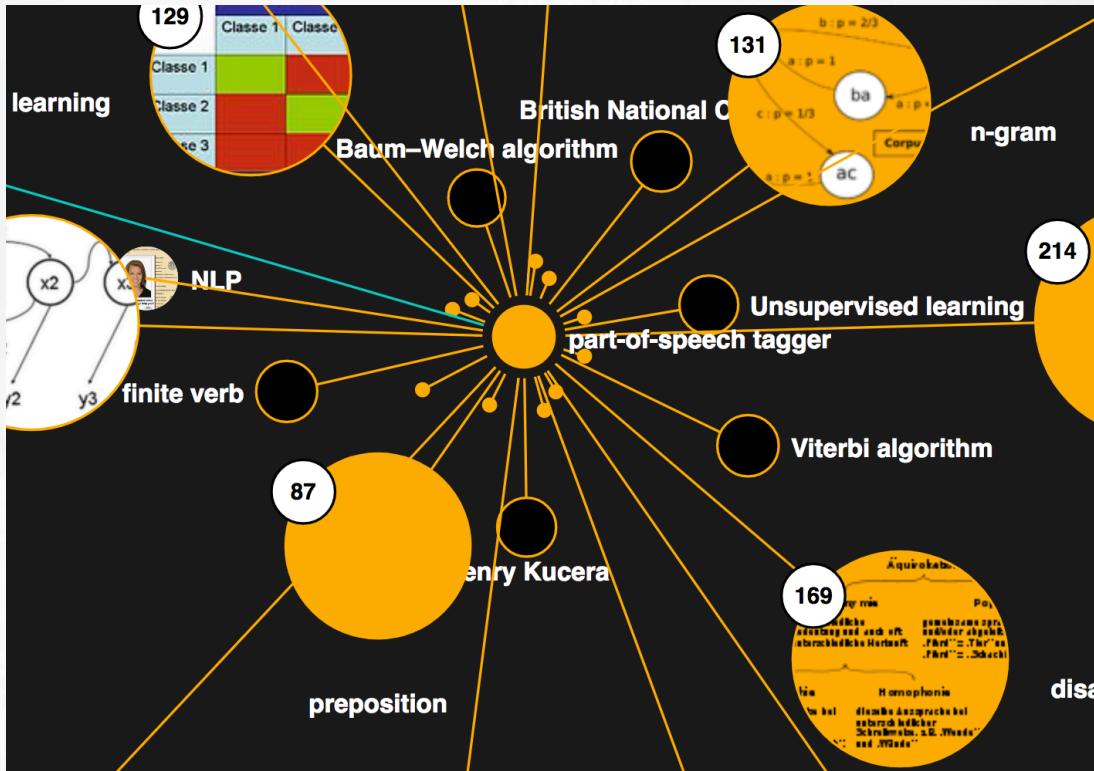
- Small but useful datasets with markup
- Preprocessing tools: tokenization, normalization...
- Pre-trained models for POS-tagging, parsing...

- **Stanford parser**

- **spaCy**: python and cython library for NLP
- **Gensim**: python library for text analysis, e.g. for word embeddings and topic modeling
- **MALLET**: Java-based library, e.g. for classification, sequence tagging, and topic modeling
- ...

# Linguistic knowledge

- Ideas and evaluation
- External resources: WordNet, BabelNet, etc.



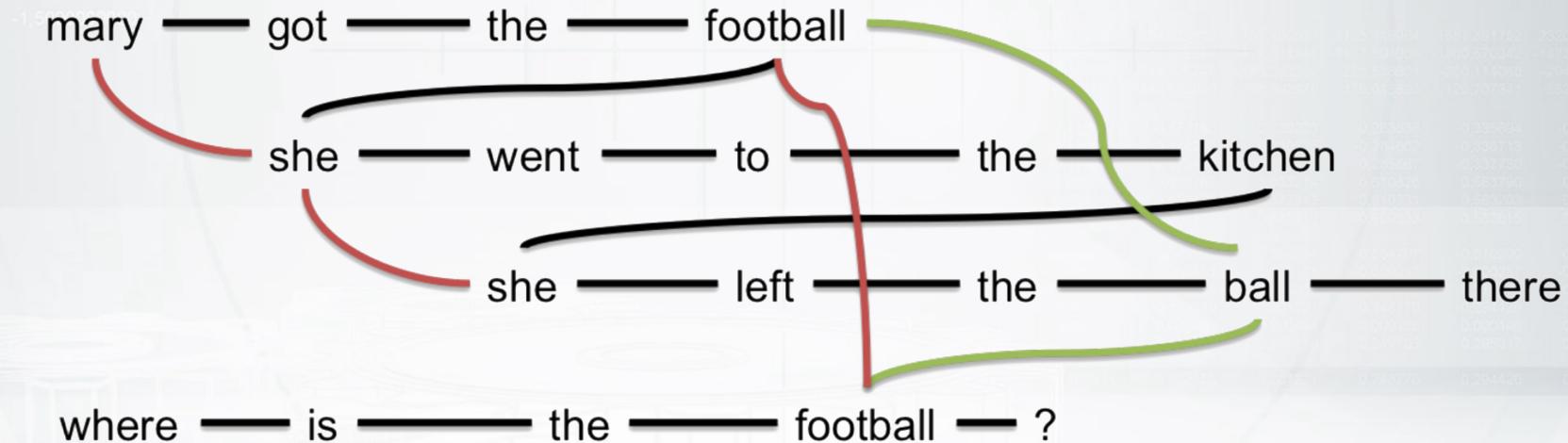
<http://babelnet.org/synset?word=NLP&lang=EN&details=1&orig=NLP>

# Linguistic knowledge + Deep Learning

*Task:* Question Answering / Reasoning

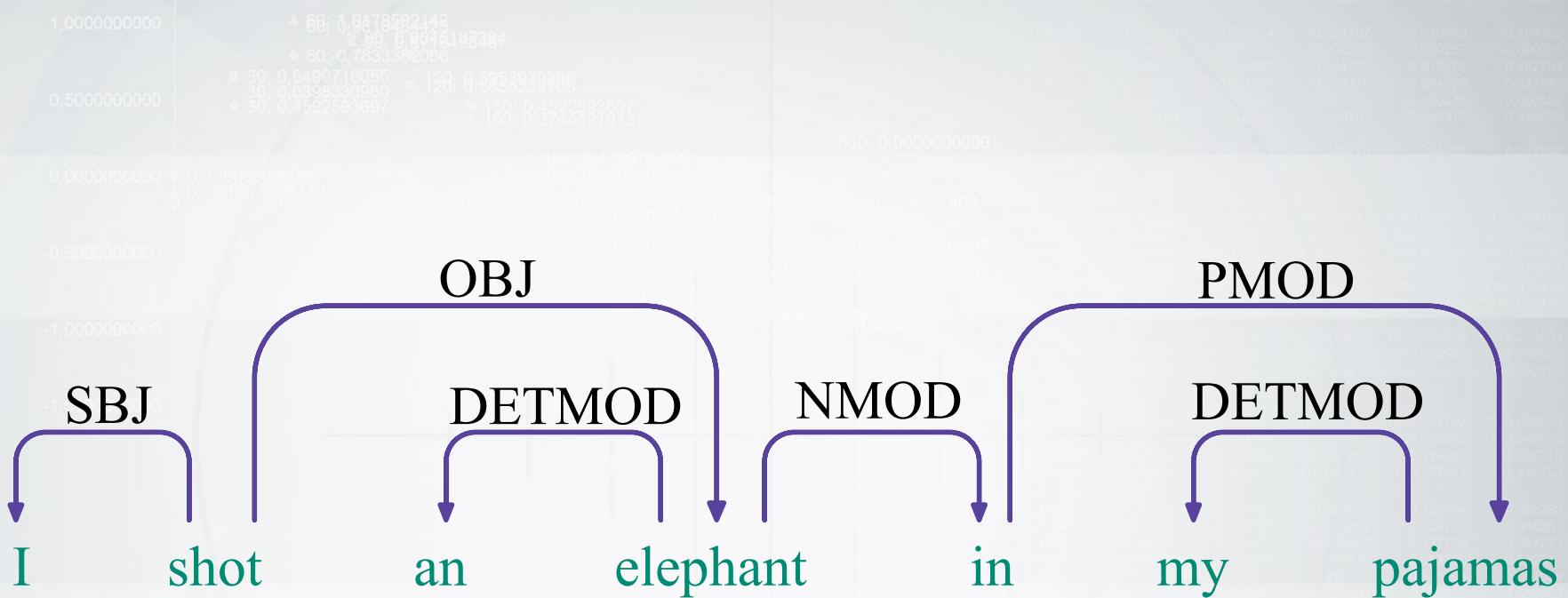
*Linguistic links:* co-reference (red), hypernyms (green)

*Method:* DAG-LSTM

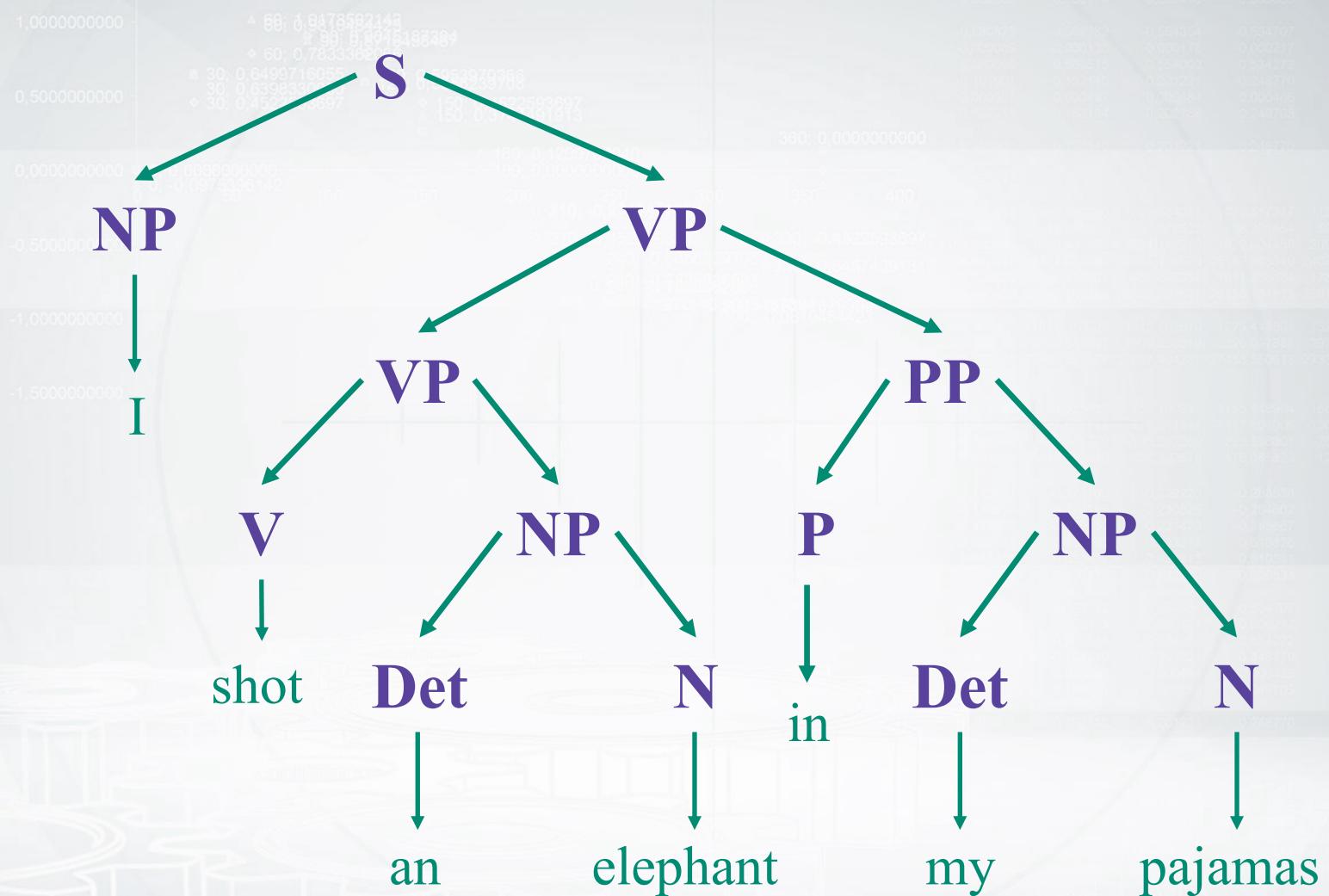


Dhingra et. al. Linguistic Knowledge as Memory for Recurrent Neural Networks, 2017.

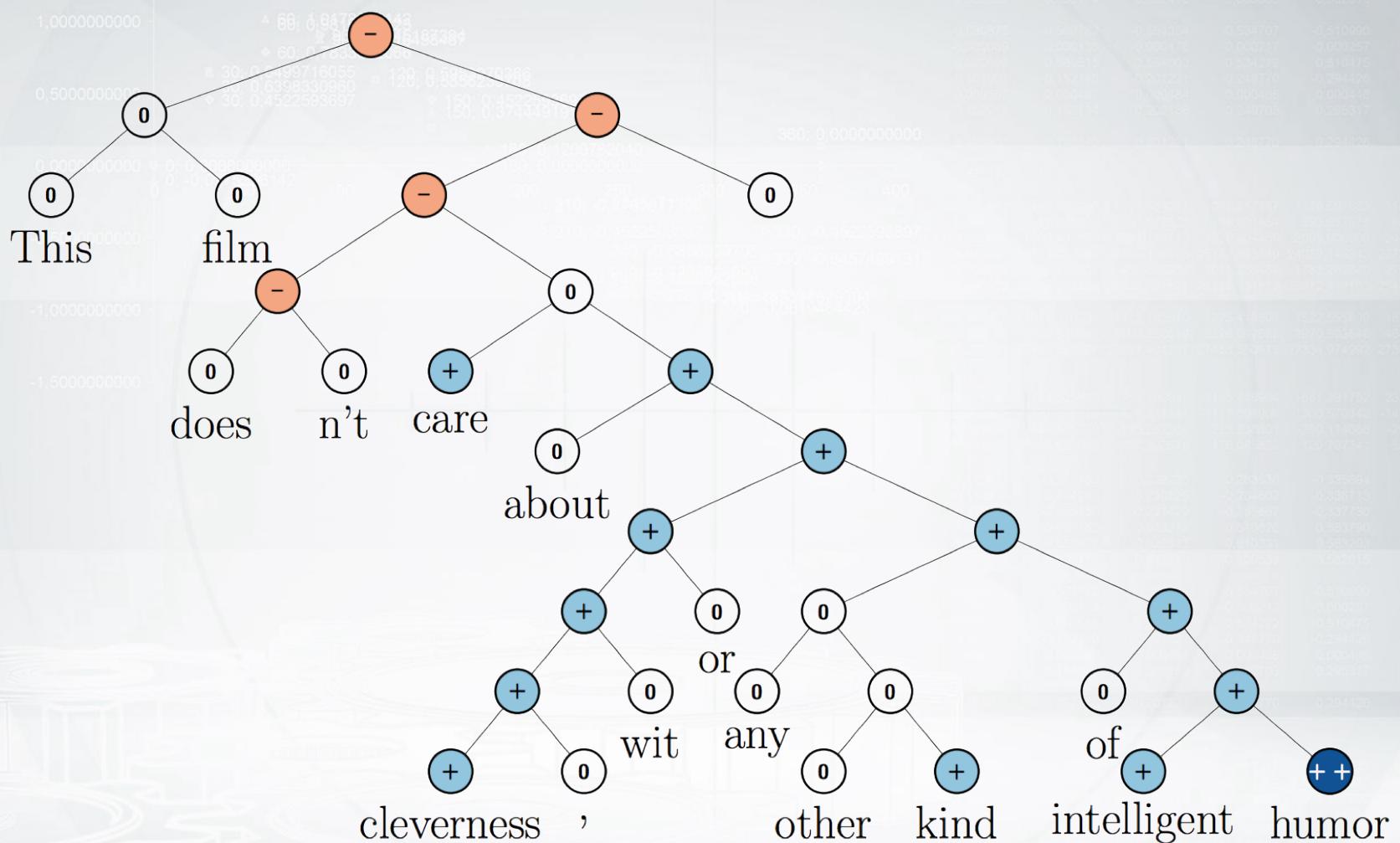
# Syntax: dependency trees



# Syntax: constituency trees



# Sentiment analysis



Socher et. al. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank, 2013.