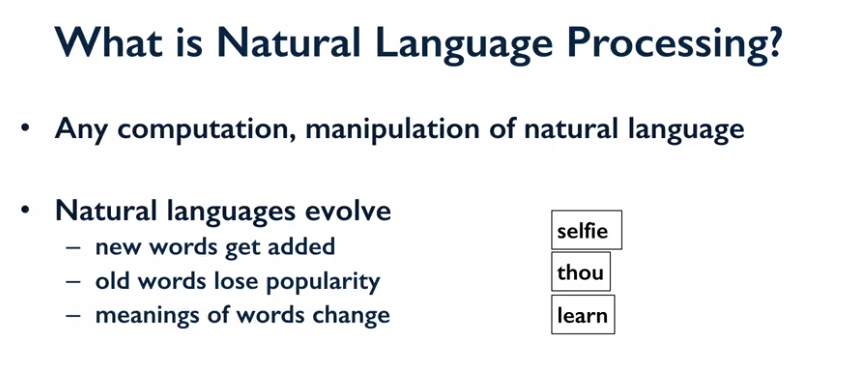
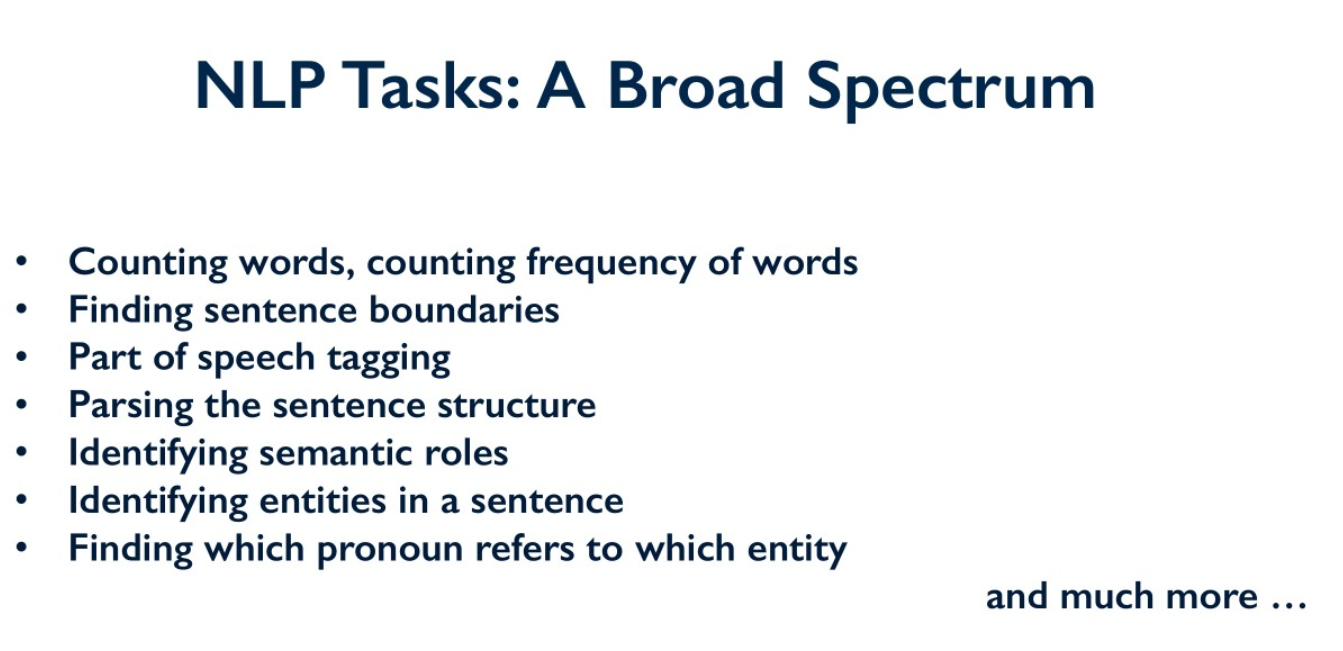
Course1

* Regular expression to find pattern in the sentence

Course 2

* NLP
* NLTK: natural language toolkit
* Normalisation and stemming
  + stemming -> ex: count listed, listing, list as the same
  + porter = nltk.PorterStemmer() -> [porter.stem(t) for t in words1]
  + Lemmatization
* Tokenize
  + nltk.word\_tokenize()





Course 3

* text classification
* feature identification
* tfidf
* ngrams

