

UNIB20005

Language and Computation

Introduction to Language Processing

Steven Bird, Department of Computing and Information Systems

Lesley Stirling, Department of Linguistics

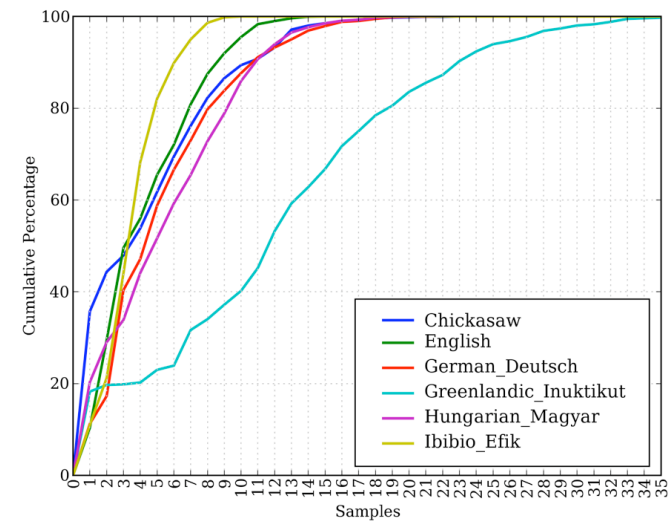
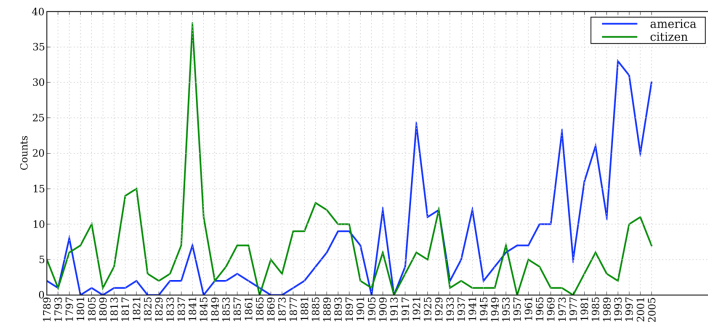
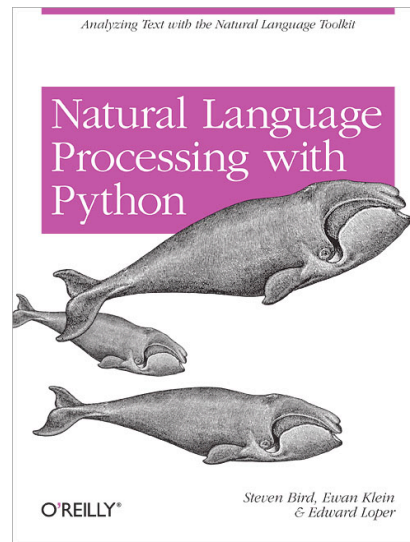
Frequency Distribution

Word Tally

the	
been	
message	
persevere	
nation	

General Information

1. LMS
2. What is “Language and Computation”?
3. Staff, Students
4. Lectures
5. Workshops
6. Textbook
7. Assessment
8. Expectations
9. Getting Help



The Richness of Language

- Overhead the day drives level and grey, hiding the sun by a flight of grey spears. (*William Faulkner, As I Lay Dying, 1935*)
- When using the toaster please ensure that the exhaust fan is turned on. (*sign in dormitory kitchen*)
- Amiodarone weakly inhibited CYP2C9, CYP2D6, and CYP3A4-mediated activities with Ki values of 45.1-271.6 μM (*Medline, PMID: 10718780*)
- Iraqi Head Seeks Arms (*spoof news headline*)
- The earnest prayer of a righteous man has great power and wonderful results. (*James 5:16b*)
- Twas brillig, and the slithy toves did gyre and gimble in the wabe (*Lewis Carroll, Jabberwocky, 1872*)
- There are two ways to do this, AFAIK :smile: (*internet discussion archive*)

Natural Language Understanding

- natural vs artificial languages
- shortcomings of web search
- language technologies
- examples: answerbus, newsblaster, wordnet
- lower-level components

Word Sense Disambiguation

- “he served the dish”
 - a. serve: help with food or drink; hold an office; put ball into play
 - b. dish: plate; course of a meal; communications device
- “by”
 - a. The lost children were found **by** the searchers (agentive)
 - b. The lost children were found **by** the mountain (locative)
 - c. The lost children were found **by** the afternoon (temporal)

Other tasks

- pronoun resolution

- a. The thieves stole the paintings. **They** were subsequently sold.
- b. The thieves stole the paintings. **They** were subsequently caught.
- c. The thieves stole the paintings. **They** were subsequently found.

- translation

- a. The thieves stole the paintings. **They** were subsequently found.
- b. Les voleurs ont volé les peintures. **Ils** ont été trouvés plus tard.
- c. Les voleurs ont volé les peintures. **Elles** ont été trouvées plus tard.

Dialogue Systems

S: How may I help you?

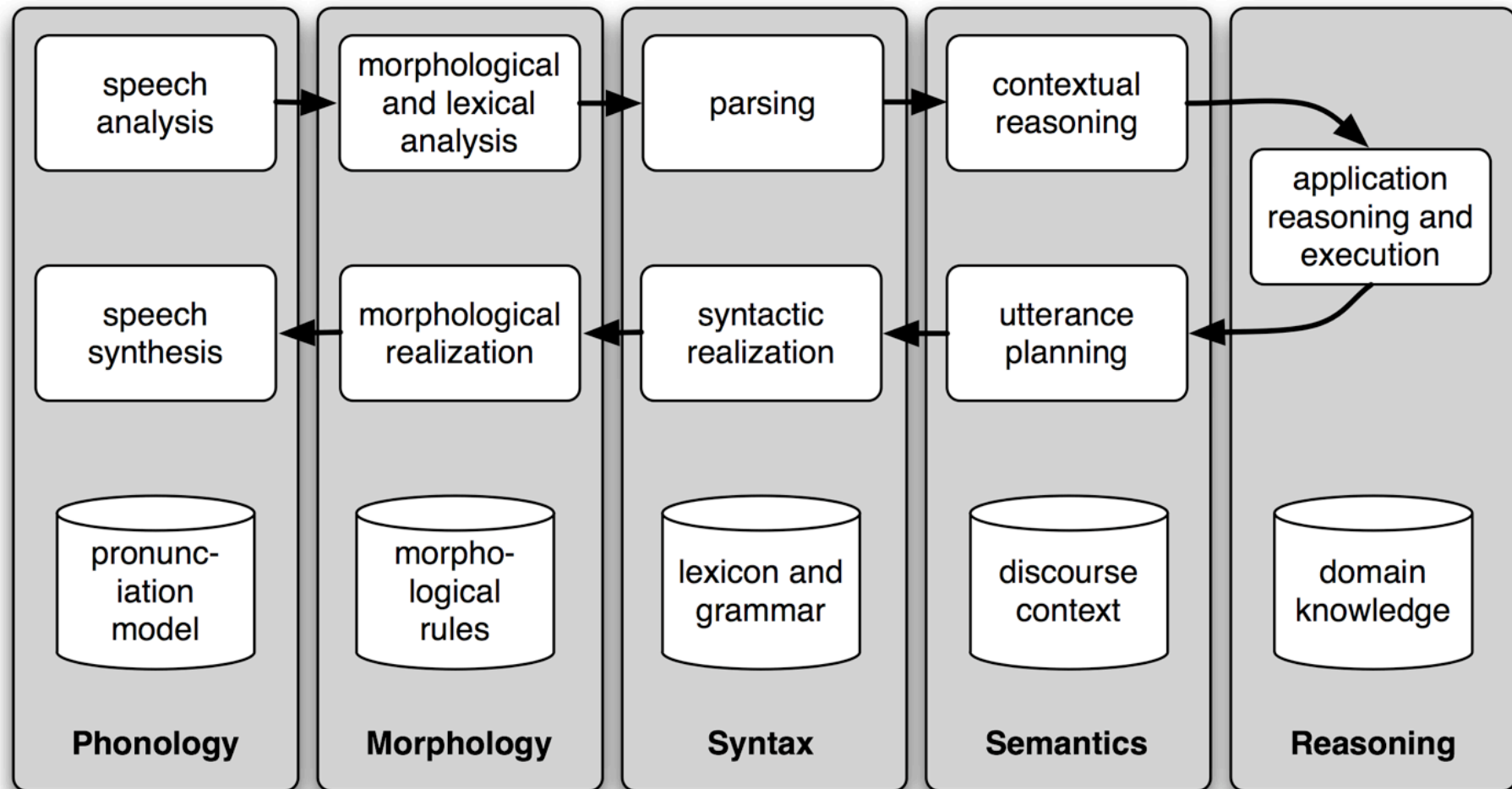
U: When is Saving Private Ryan playing?

S: For what theater?

U: The Paramount theater.

S: Saving Private Ryan is not playing at the Paramount theater, but it's playing at the Madison theater at 3:00, 5:30, 8:00, and 10:30.

Dialogue Systems



For Wednesday

- Start reading chapter 1 of the textbook
<http://www.nltk.org/book>
- Check out language.log.org
- How many words do you know?