Natural Language Processing

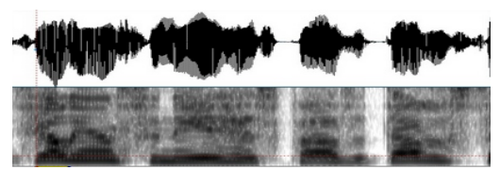
Midterm Review and Practice Problems

10/11/2018

COMS W4705
Daniel Bauer

Levels of Linguistic Representation

phonetics phonology	sounds and sound patterns of language



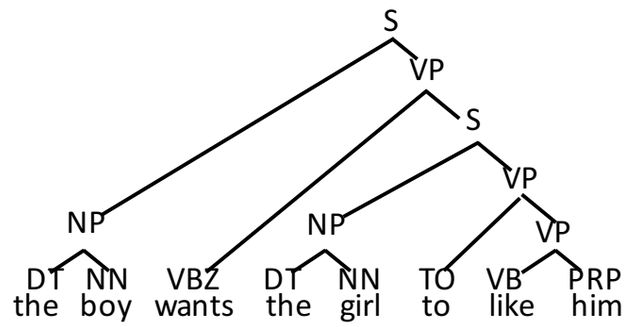
/bɔɪ/

morphology	formation of words

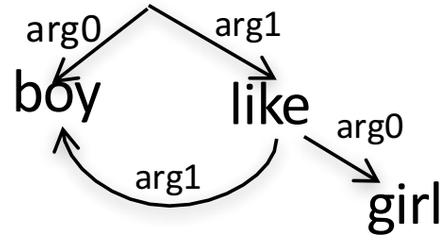
in- + validate + -ed

syntax	word order

```
DT  | NN  |   VBZ     |  DT  | NN   | TO | VB   | PRP |.
the | boy | want+s    | the  | girl | to | like | him |.
```



semantics	word and sentence meaning



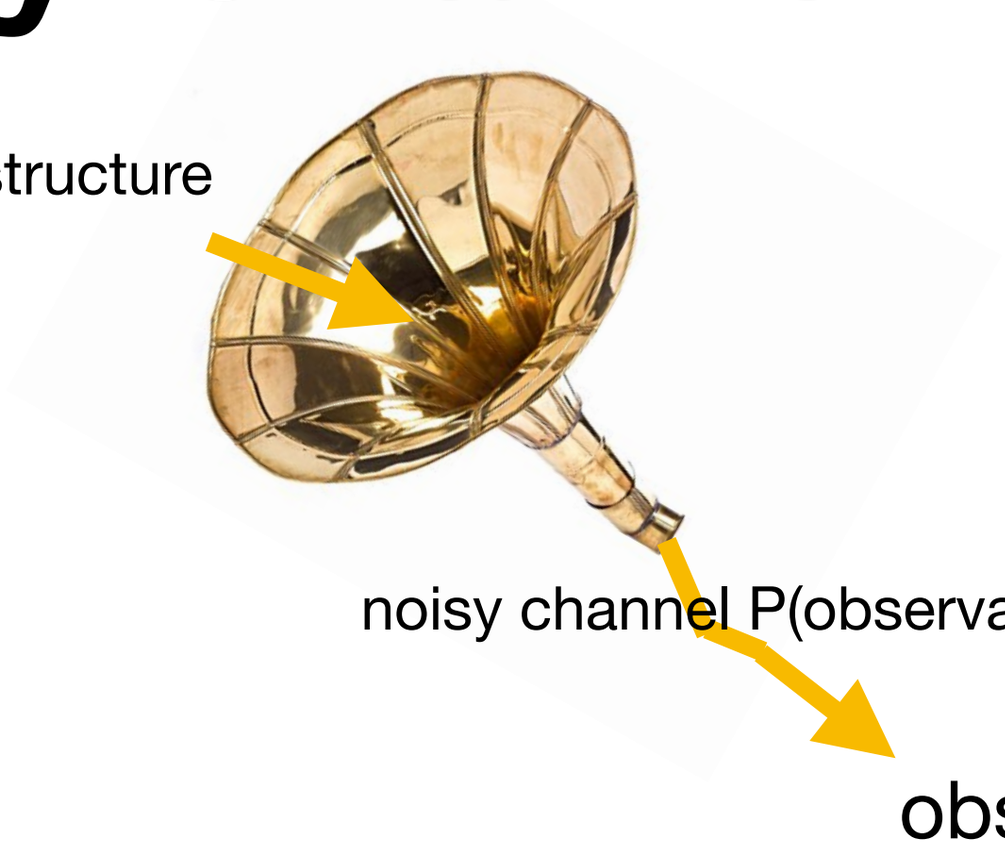
pragmatics	influence of context and situation

Natural Language Processing as Translation

- Most NLP techniques can be understood as translation tasks from one structure into another.

- For each translation step:
    - Construct search space of possible translations.
    - Find best paths through this space (decoding) according to some performance measure.
- Modern NLP relies on Machine Learning to figure out these translation steps.
    - Generative vs. discriminative models.

Noisy Channel Model



- Decoding: figure out what the original input to the the channel was. Use Bayes' rule:

Naive Bayes

a **great game**	sports
the **election** was **over**	politics
very clean match	sports
a **clean** but **close game**	sports
it was a **close election**	politics

what is the category for "*the close election was over*"

HMM Exercise

given the following POS annotated text corpus:

1. *(D,the) (N,cheese) (V,ages) (A,alone)*
2. *(D,the) (N,cheese) (N,waits) (V,last) (N,ages)*
3. *(D,some) (N,flies) (V,dove) (P,for) (D,the) (N,cheese)*
4. *(D,the) (N,dove) (V,flies) (P,for) (D,some) (N,flies)*
5. *(D,the) (A,last) (N,dove) (V,waits) (A,alone)*

- Estimate the emission probabilities.
- Estimate the transition probabilities.

- Using the Viterbi algorithm, what is the most likely tag sequence for

some dove waits for the flies

Context Free Grammar

S → NP VP
NP → NP PP
NP → N
NP → PRP
PP → P NP
VP → V NP
VP → VP PP

PRP → *I*
N → *sushi*
N → *tuna*
N → *chopsticks*
N → *friends*
P → *with*
V → *eat*

- Is this grammar in CNF?
- Parse the sentence using the Earley algorithm.
- Convert grammar to CNF, then parse using CKY.

I eat sushi with friends

Dependency Parsing

I eat sushi with friends

- What are the dependency structures corresponding to the different interpretations?
- How would an arc-standard transition-based dependency parser produce each tree?