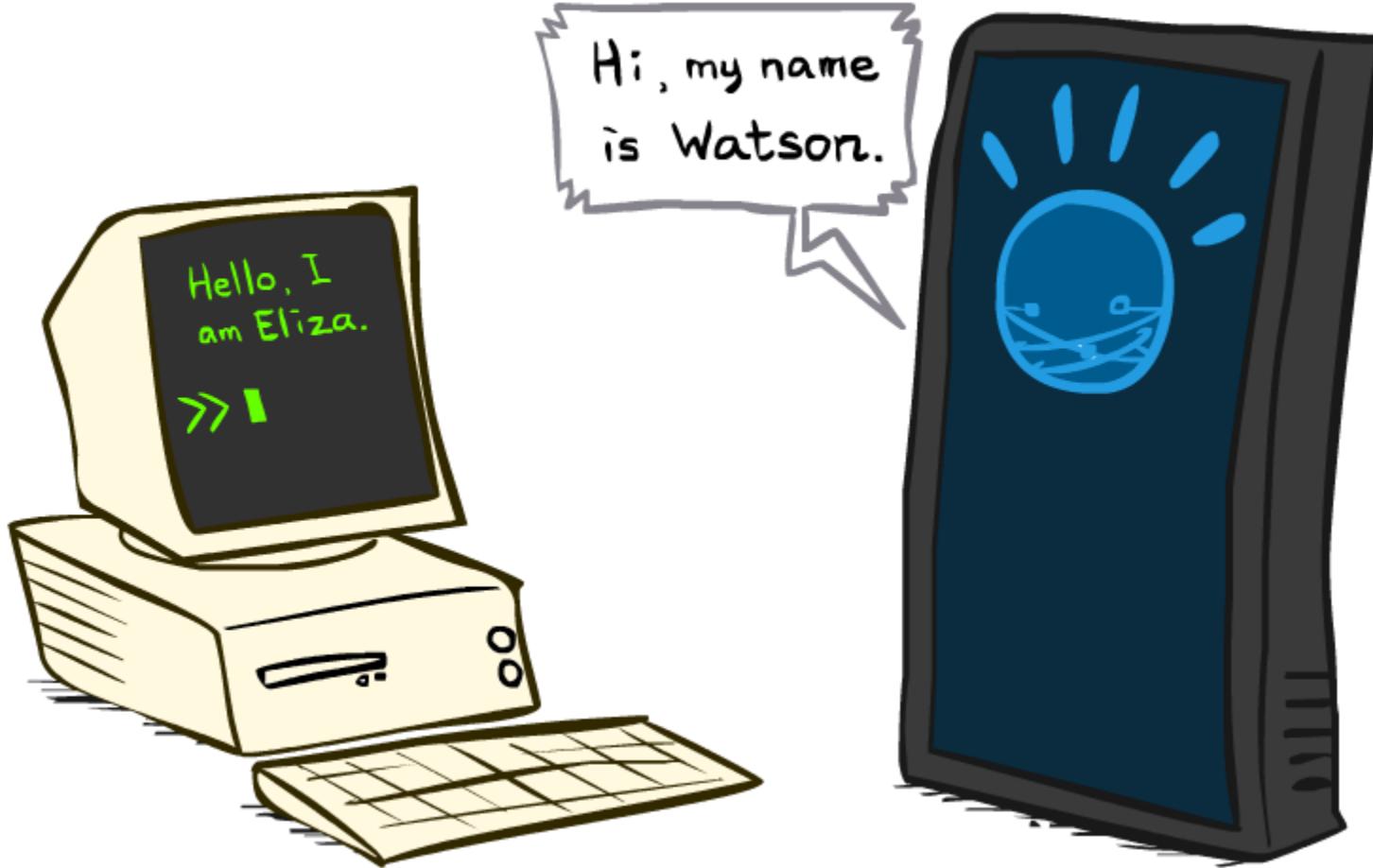


Natural Language Processing



What is NLP?



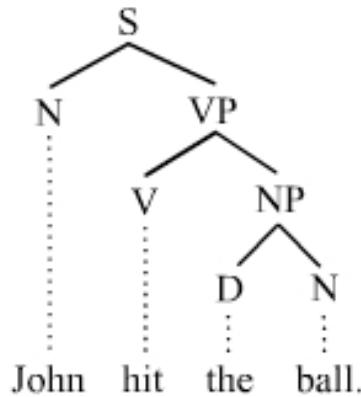
- Fundamental goal: analyze and process human language, broadly, robustly, accurately...
- End systems that we want to build:

Speech recognition

Machine translation

Question answering /dialogue

NLP at Berkeley



Core linguistic analysis



Grounded language learning

Starcraft and pianos?

Problem: Ambiguities

- **Ambiguous grammar**

- Enraged Cow Injures Farmer With Ax
- Hospitals Are Sued by 7 Foot Doctors
- Ban on Nude Dancing on Governor's Desk

- **Ambiguous sounds**

- “lonely Starbucks lovers”

- **Ambiguous words**

- Iraqi Head Seeks Arms
- Local HS Dropouts Cut in Half
- Kids Make Nutritious Snacks



Machine Translation



Machine Translation

"Il est impossible aux journalistes de rentrer dans les régions tibétaines"

Bruno Philip, correspondant du "Monde" en Chine, estime que les journalistes de l'AFP qui ont été expulsés de la province tibétaine du Qinghai "n'étaient pas dans l'illégalité".

Les faits Le dalaï-lama dénonce l'"enfer" imposé au Tibet depuis sa fuite, en 1959

Vidéo Anniversaire de la rébellion tibétaine : la Chine sur ses gardes



"It is impossible for journalists to enter Tibetan areas"

Philip Bruno, correspondent for "World" in China, said that journalists of the AFP who have been deported from the Tibetan province of Qinghai "were not illegal."

Facts The Dalai Lama denounces the "hell" imposed since he fled Tibet in 1959

Video Anniversary of the Tibetan rebellion: China on guard



- Translate text from one language to another
- Recombines fragments of example translations
- Challenges:
 - What fragments? [learning to translate]
 - How to make efficient? [fast translation search]

The Problem with Dictionary Lookups

顶部 /**top**/roof/

顶端 /summit/peak/**top**/apex/

顶头 /coming directly towards one/**top**/end/

盖 /lid/**top**/cover/canopy/build/Gai/

盖帽 /surpass/**top**/

极 /extremely/pole/utmost/**top**/collect/receive/

尖峰 /peak/**top**/

面 /fade/side/surface/aspect/**top**/face/flour/

摘心 /**top**/topping/

MT: 60 Years in 60 Seconds



Warren Weaver

When I look at an article in Russian, I say: "This is really written in English, but it has been coded in some strange symbols. I will now proceed to decode."



John Pierce

"Machine Translation" presumably means going by algorithm from machine-readable source text to useful target text... In this context, there has been no machine translation...

Berkeley's first MT grant

MT is the "first" non-numeral compute task

ALPAC report deems MT bad

Statistical MT thrives

Statistical data-driven approach introduced



'47

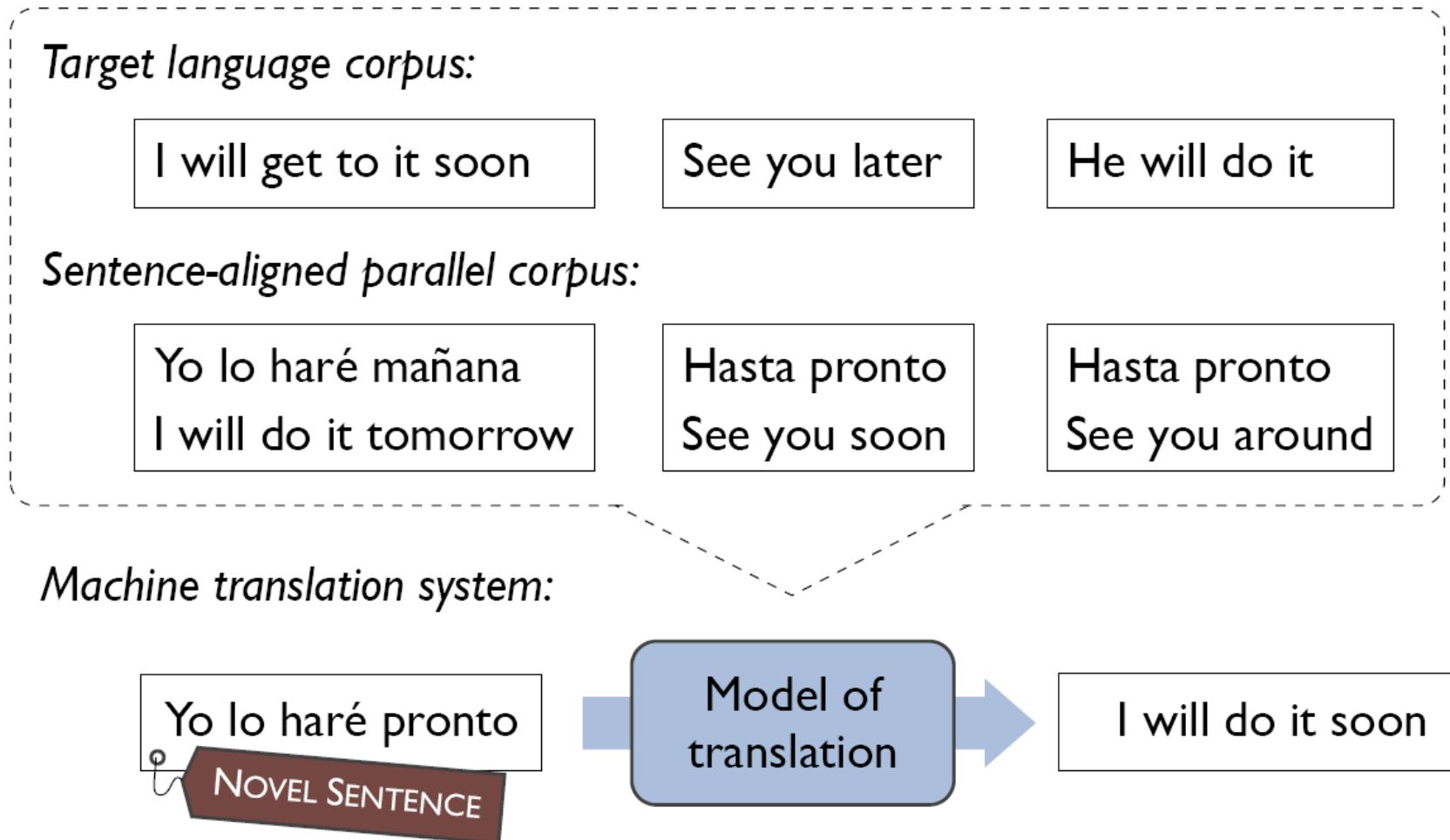
'58

'66

'90's

'00's

Data-Driven Machine Translation

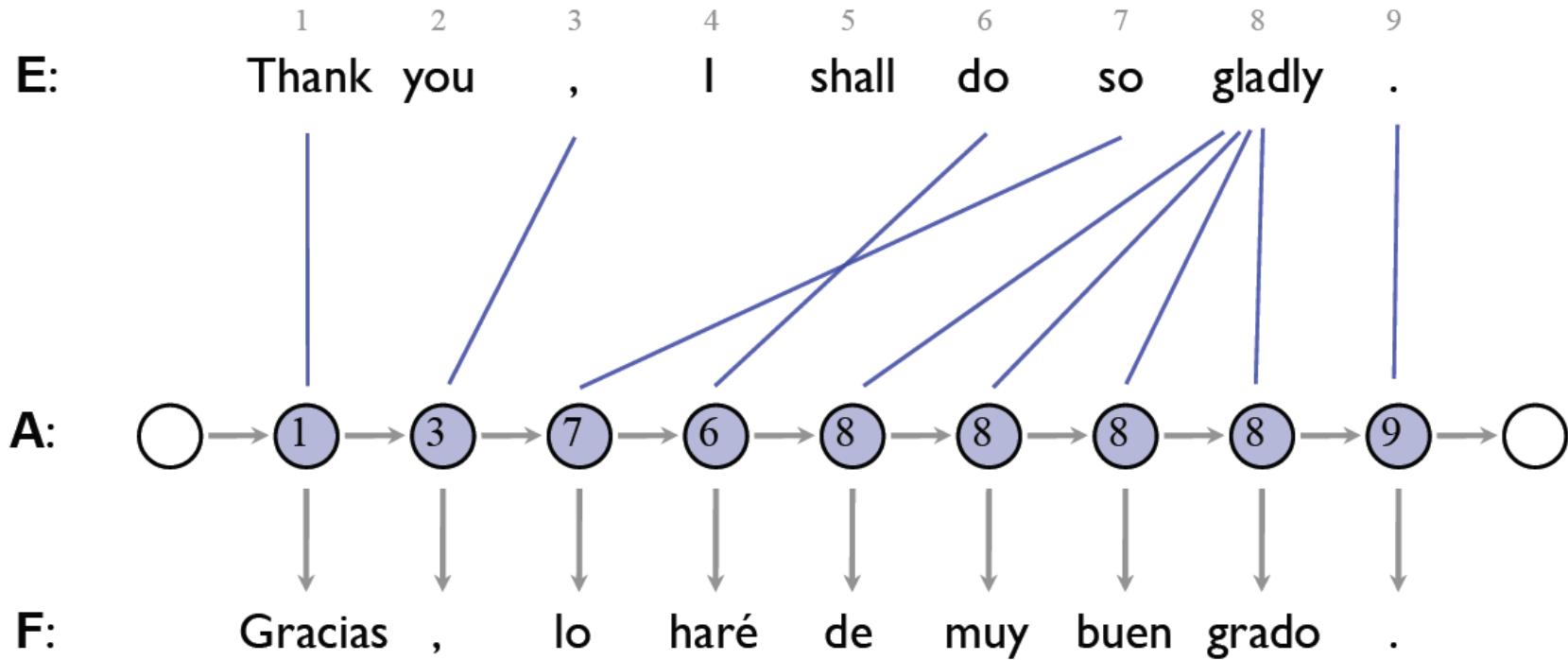


Learning to Translate

		CLASSIC SOUPS		Sm.	Lg.
清 燉 雞 湯	57.	House Chicken Soup (Chicken, Celery, Potato, Onion, Carrot)		1.50	2.75
雞 飯 湯	58.	Chicken Rice Soup		1.85	3.25
雞 麵 湯	59.	Chicken Noodle Soup		1.85	3.25
廣 東 雲 吞	60.	Cantonese Wonton Soup.....		1.50	2.75
蕃 茄 蛋 湯	61.	Tomato Clear Egg Drop Soup		1.65	2.95
雲 吞 湯	62.	Regular Wonton Soup		1.10	2.10
酸 辣 湯	63.	Hot & Sour Soup		1.10	2.10
蛋 花 湯	64.	Egg Drop Soup.....		1.10	2.10
雲 蛋 湯	65.	Egg Drop Wonton Mix.....		1.10	2.10
豆 腐 菜 湯	66.	Tofu Vegetable Soup		NA	3.50
雞 玉 米 湯	67.	Chicken Corn Cream Soup		NA	3.50
蟹 肉 玉 米 湯	68.	Crab Meat Corn Cream Soup.....		NA	3.50
海 鮮 湯	69.	Seafood Soup.....		NA	3.50

Example from Adam Lopez

An HMM Translation Model

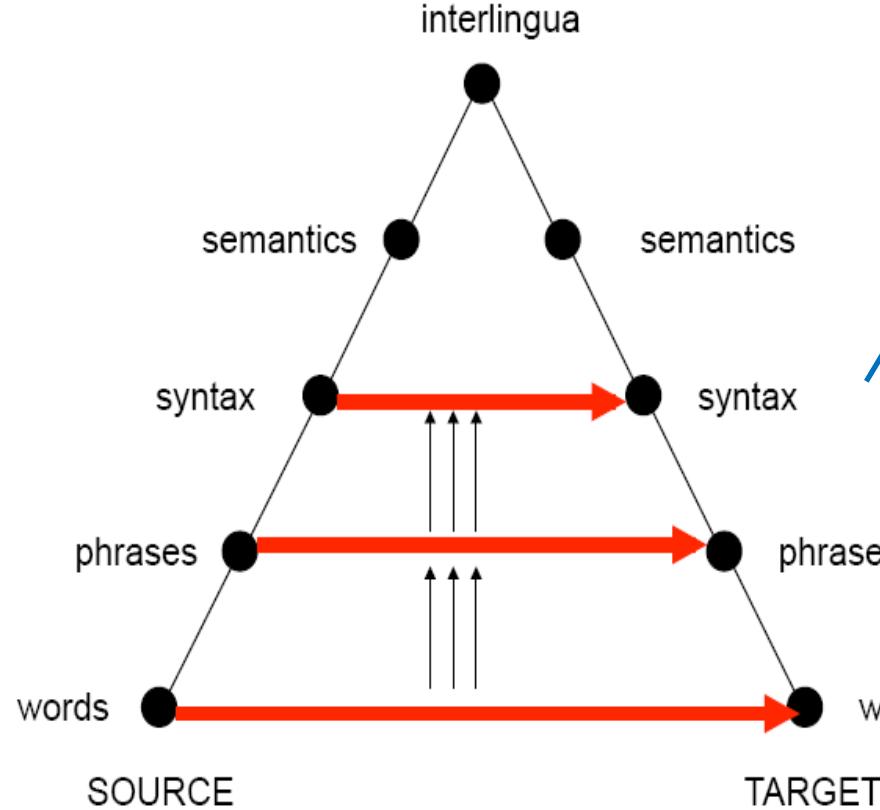


Model Parameters

Emissions: $P(F_1 = \text{Gracias} | E_{A1} = \text{Thank})$

Transitions: $P(A_2 = 3 | A_1 = 1)$

Levels of Transfer



Yo lo haré mañana
I will do it tomorrow

Yo lo haré mañana
I will do it, tomorrow

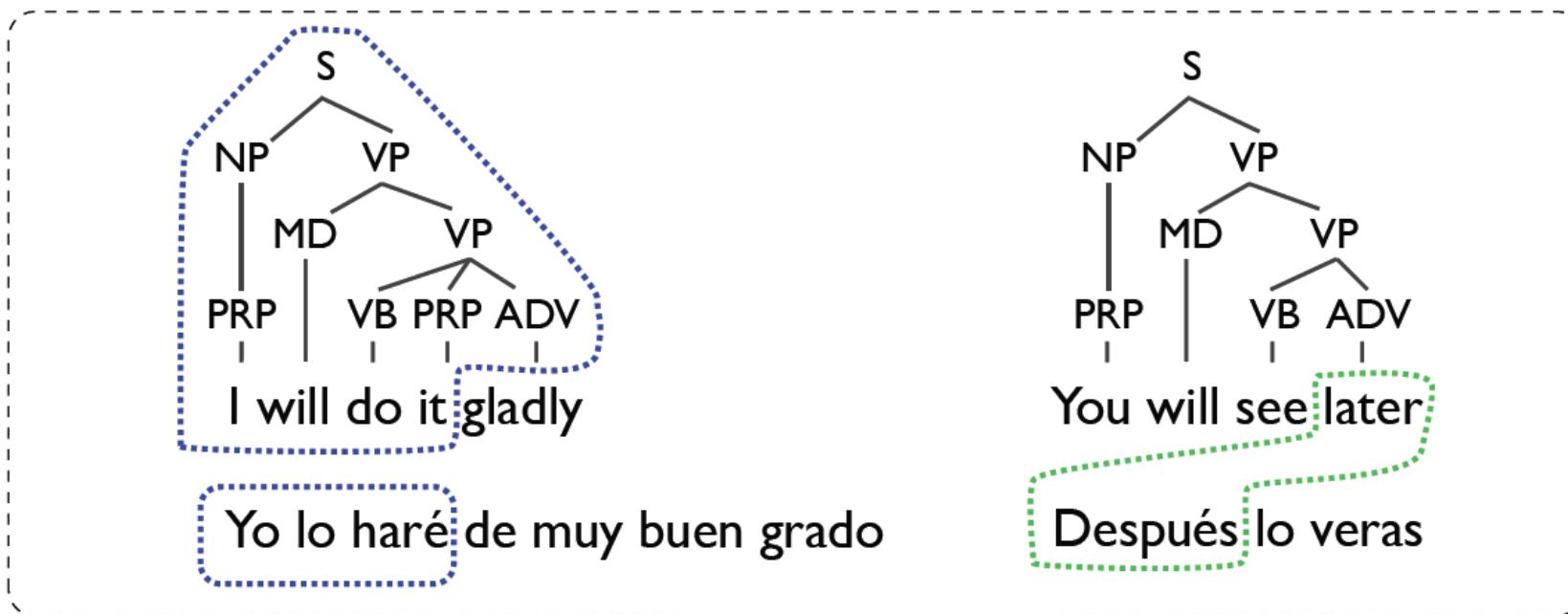
Yo lo haré mañana
I will do it tomorrow

$$P(\text{MD VP} \mid \text{will do it}) = 0.8$$

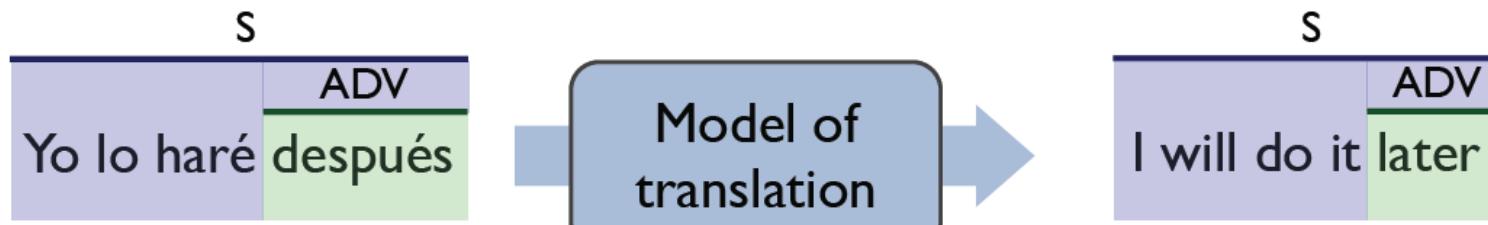
English (E)	$P(E \mid \text{lo haré})$
will do it	0.8
will do so	0.2

English (E)	$P(E \mid \text{mañana})$
tomorrow	0.7
morning	0.3

Syntactic Translation



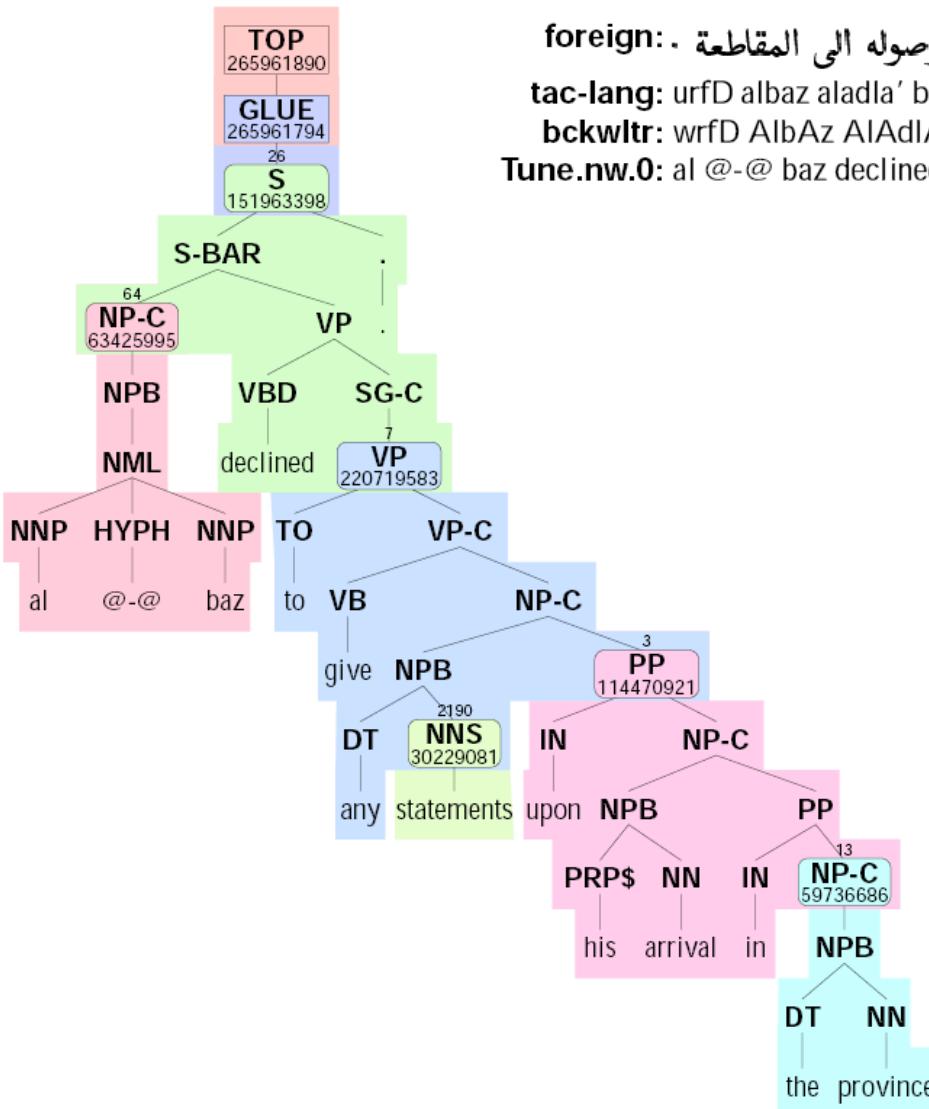
Machine translation system:



Example: Syntactic MT Output

ورفض الباز الأدلاء باى تصريحات فور وصوله الى المقاطعة .
foreign:

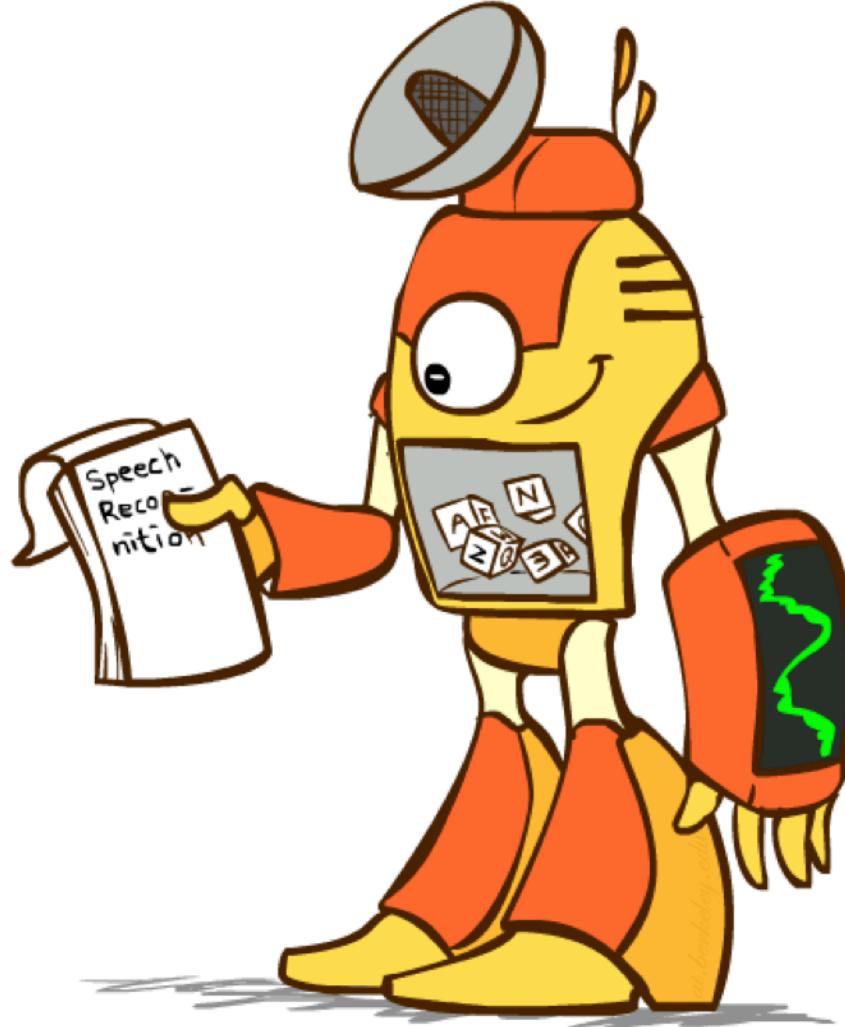
Example: Syntactic MT Output



foreign: .: ورفض البار الادلاء باى تصريحات فور وصوله الى المقاطعة .
tac-lang: urfD albaz aladla' baá tSryHat fur uSulh alá almqaT'e .
bckwltr: wrfD AlbAz AlAdIA' bAY tSryHAt fwr wSwlh AIY AlmqATEp .
Tune.nw.0: al @-@ baz declined to make any statements upon his arrival in the province .

[ISI MT system output]

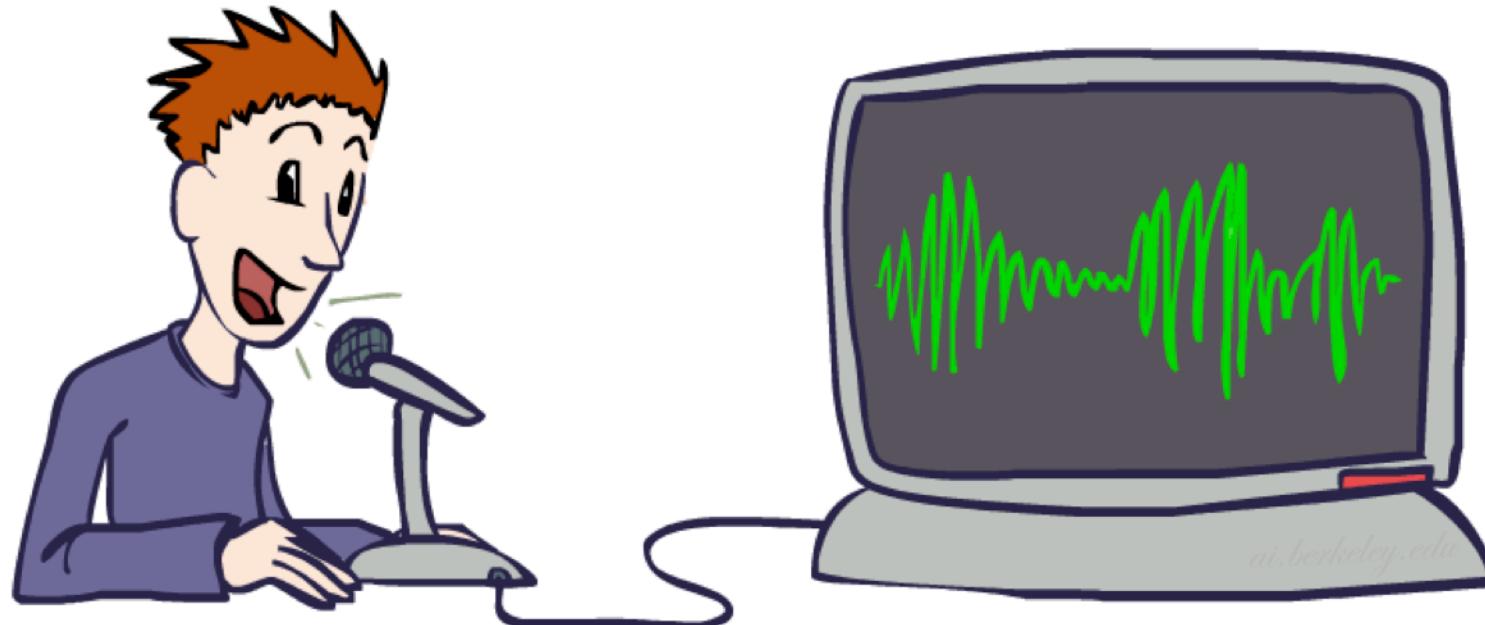
Speech Recognition



Speech Recognition in Action

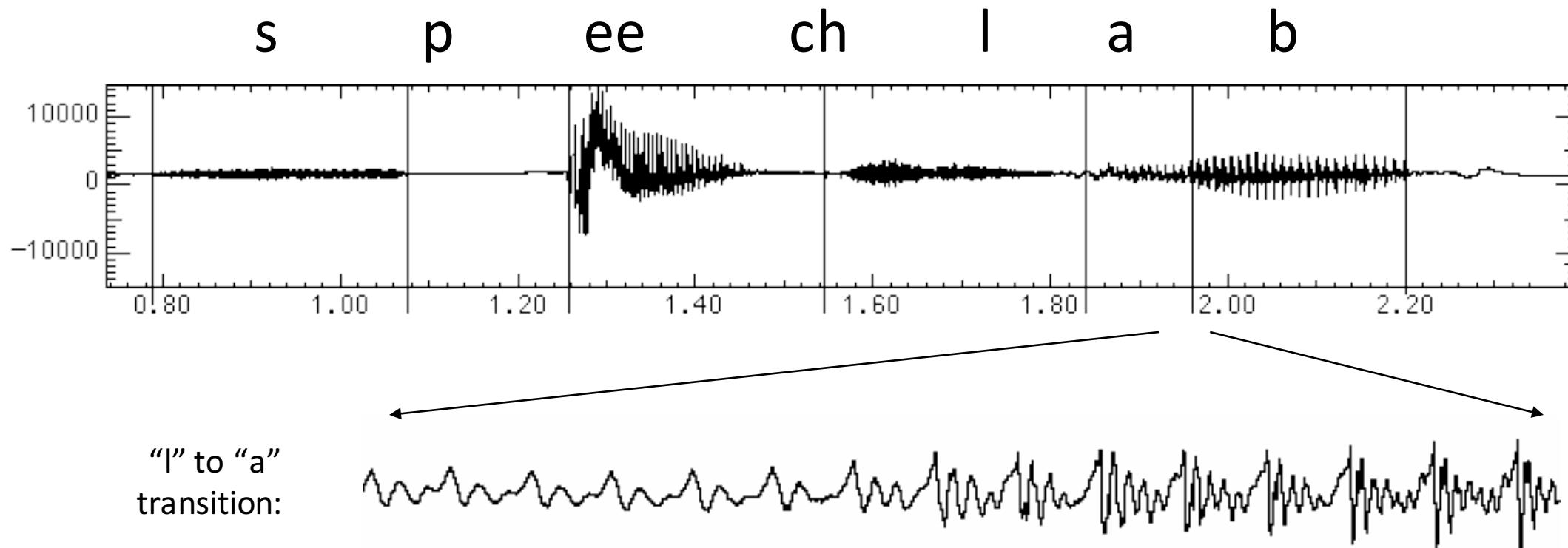
[Video: NLP – ASR tvsample.avi (from Lecture 1)]

Digitizing Speech



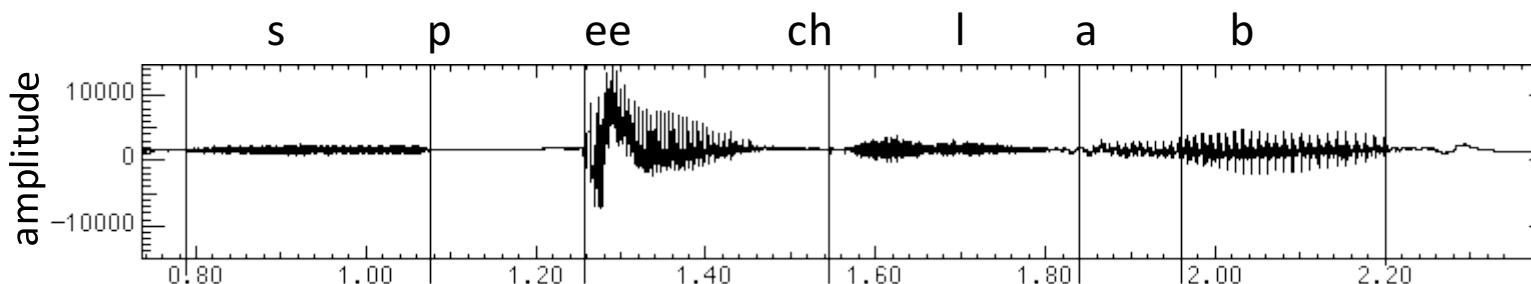
Speech in an Hour

- Speech input is an acoustic waveform

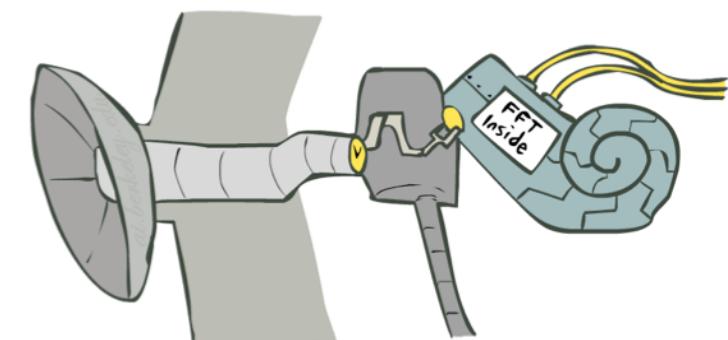
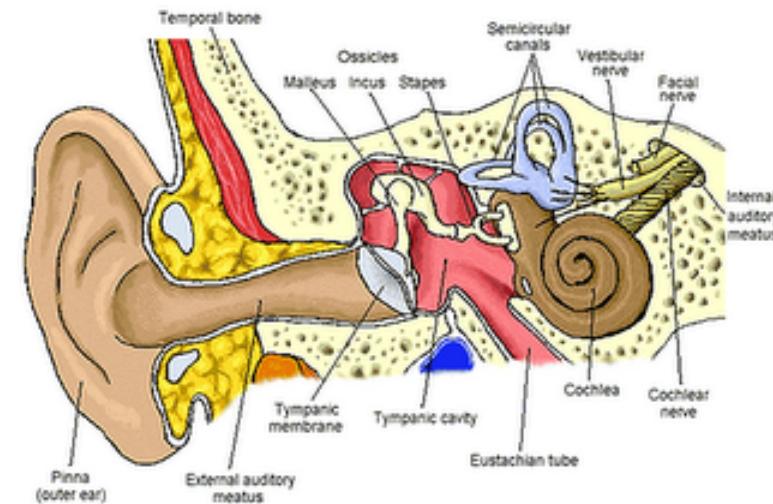
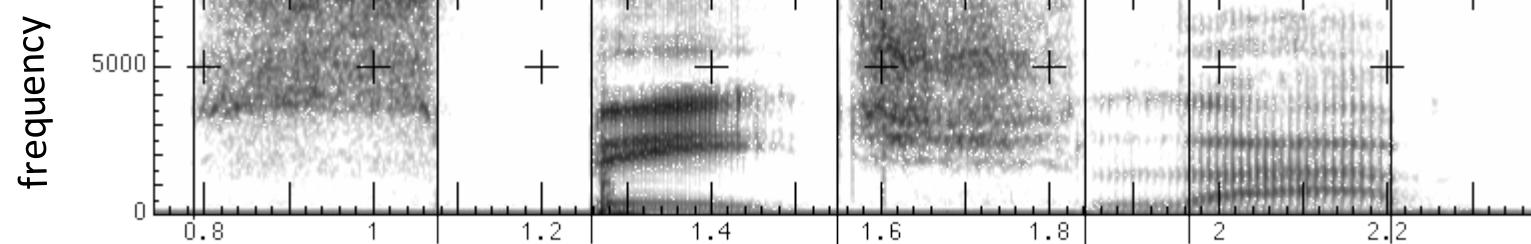


Spectral Analysis

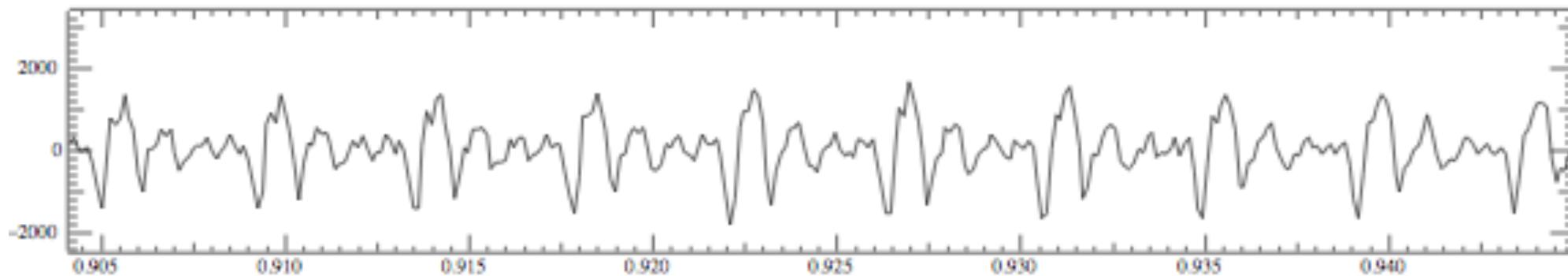
- Frequency gives pitch; amplitude gives volume
 - Sampling at ~8 kHz (phone), ~16 kHz (mic) (kHz=1000 cycles/sec)



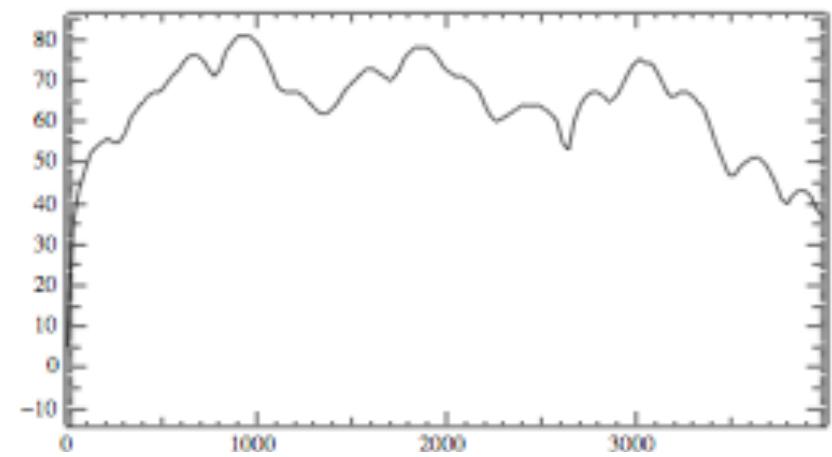
- Fourier transform of wave displayed as a spectrogram
 - Darkness indicates energy at each frequency



Part of [ae] from “lab”



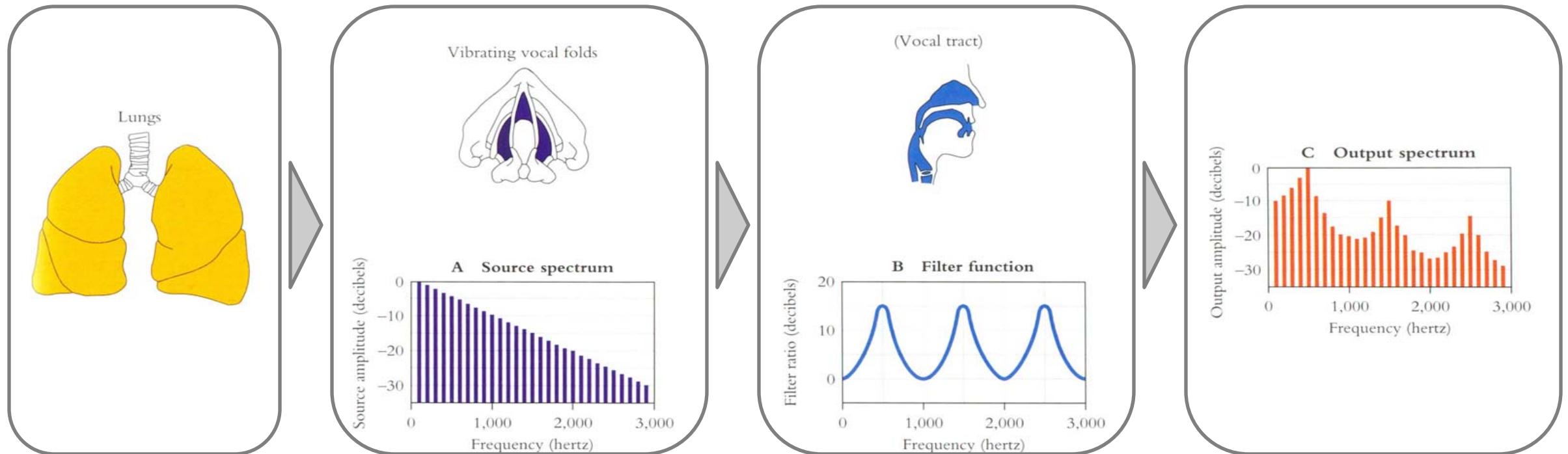
- Complex wave repeating nine times
 - Plus smaller wave that repeats 4x for every large cycle
 - Large wave: freq of 250 Hz (9 times in .036 seconds)
 - Small wave roughly 4 times this, or roughly 1000 Hz



Why These Peaks?

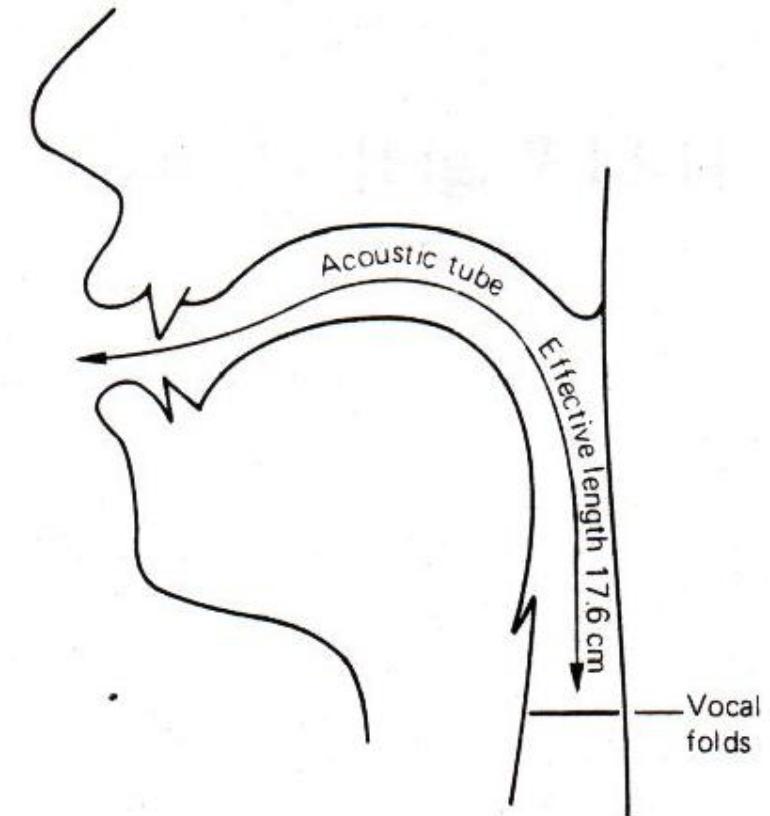
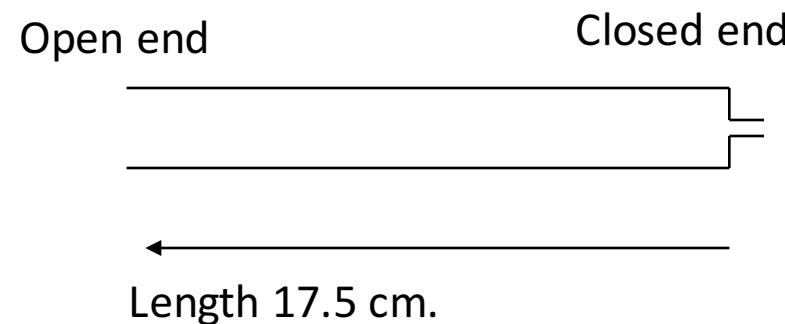
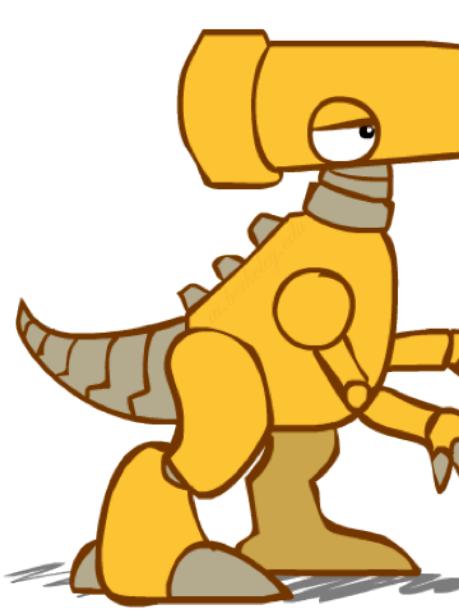
- Articulator process:

- Vocal cord vibrations create harmonics
- The mouth is an amplifier
- Depending on shape of mouth, some harmonics are amplified more than others



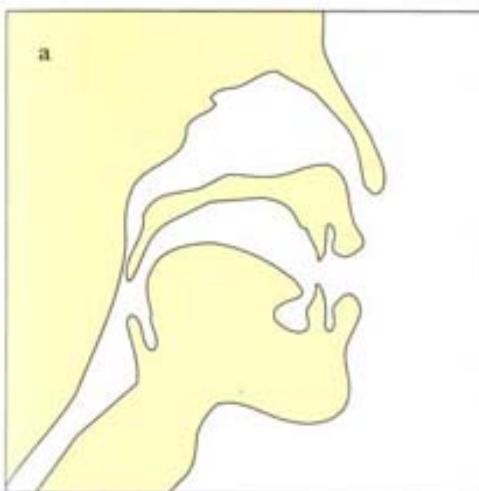
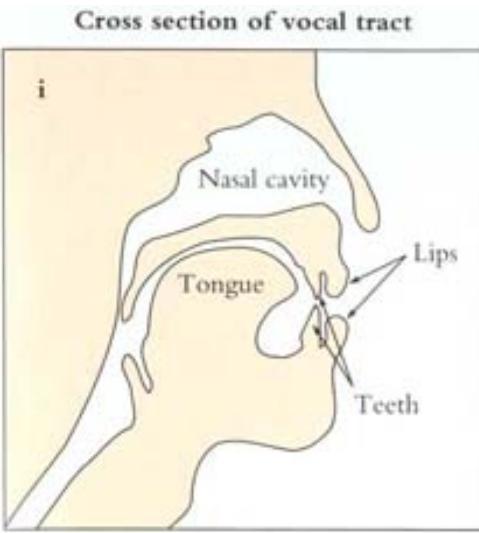
Resonances of the Vocal Tract

- The human vocal tract as an open tube

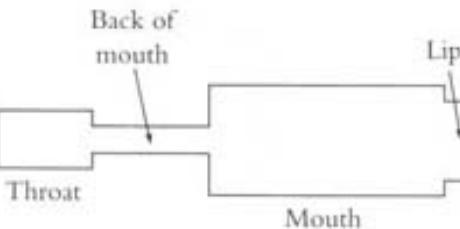
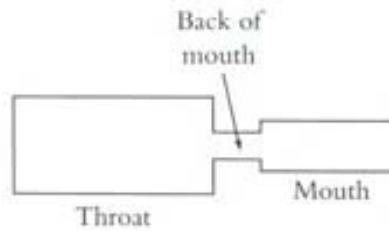


- Air in a tube of a given length will tend to vibrate at resonance frequency of tube
- Constraint: Pressure differential should be maximal at (closed) glottal end and minimal at (open) lip end

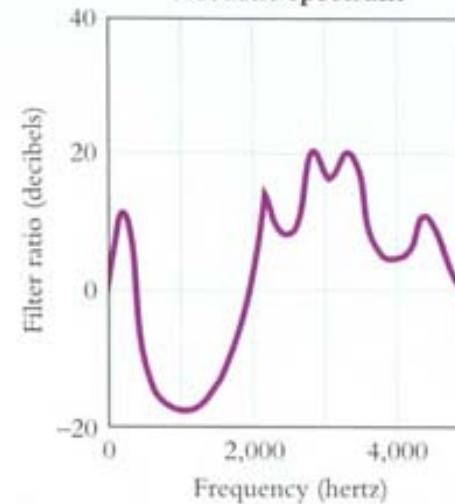
Spectrum Shapes



Model of vocal tract



Acoustic spectrum



Acoustic spectrum

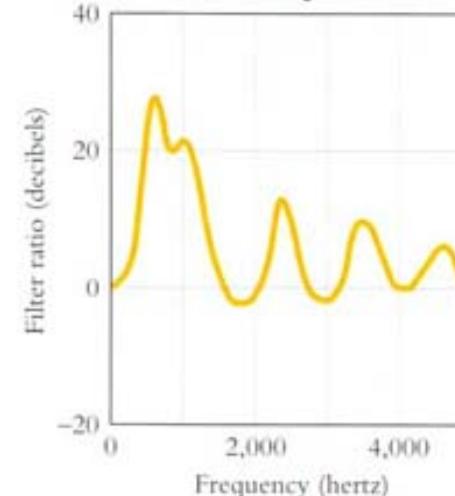


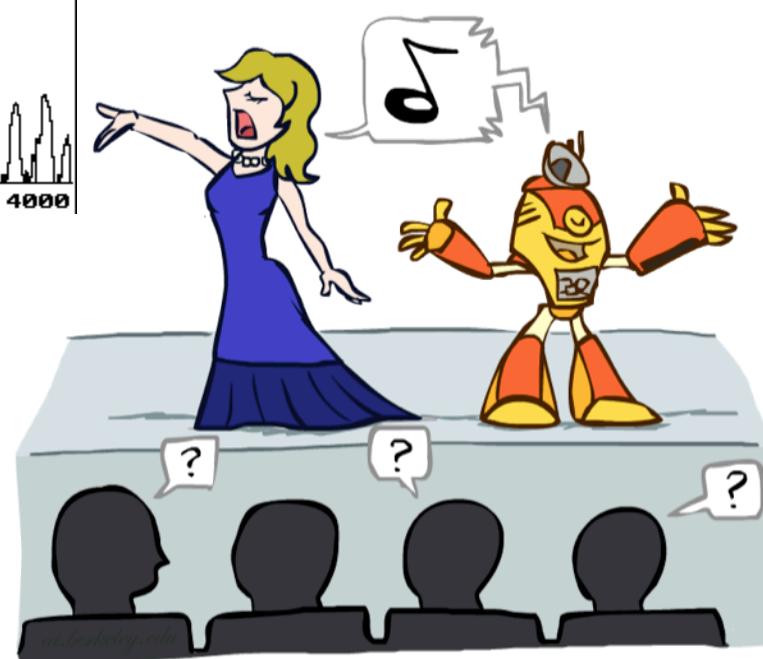
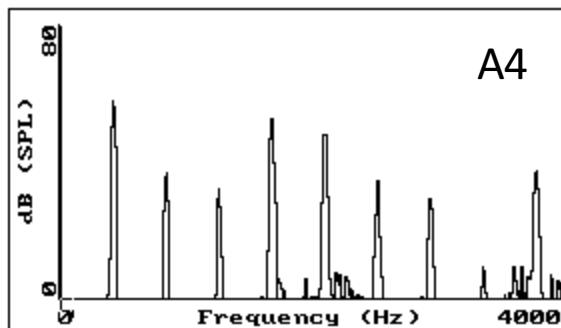
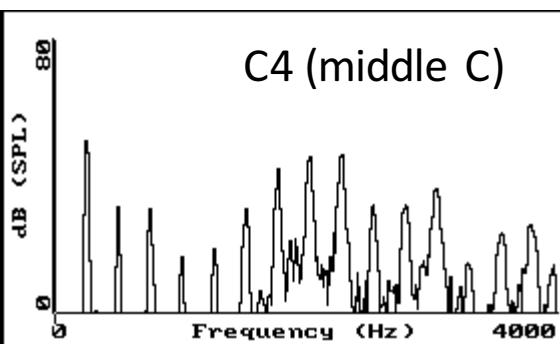
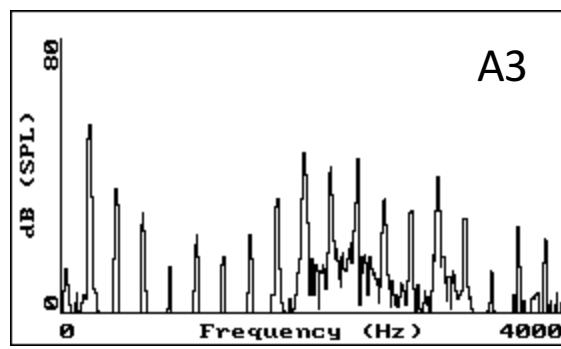
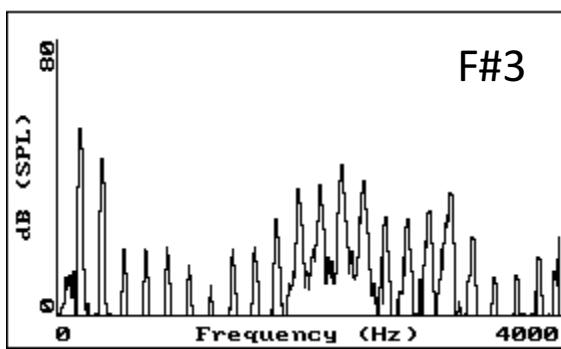
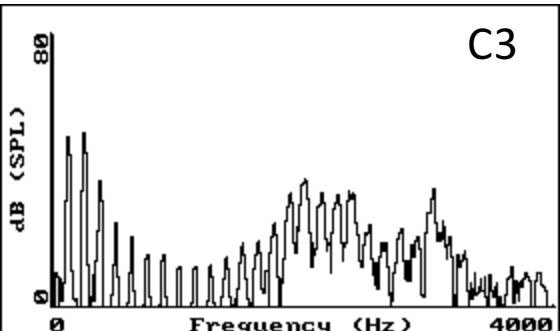
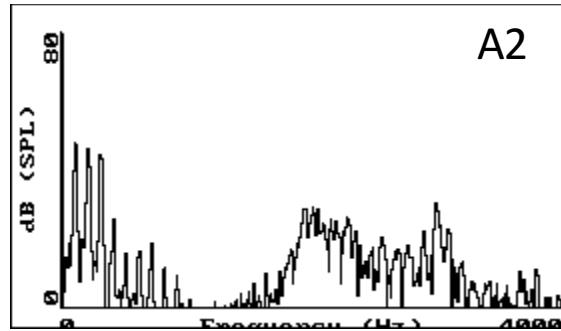
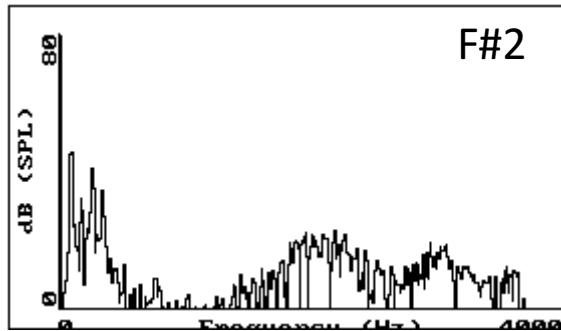
Figure: Mark Liberman

[Demo: speech synthesis]

Video of Demo Speech Synthesis

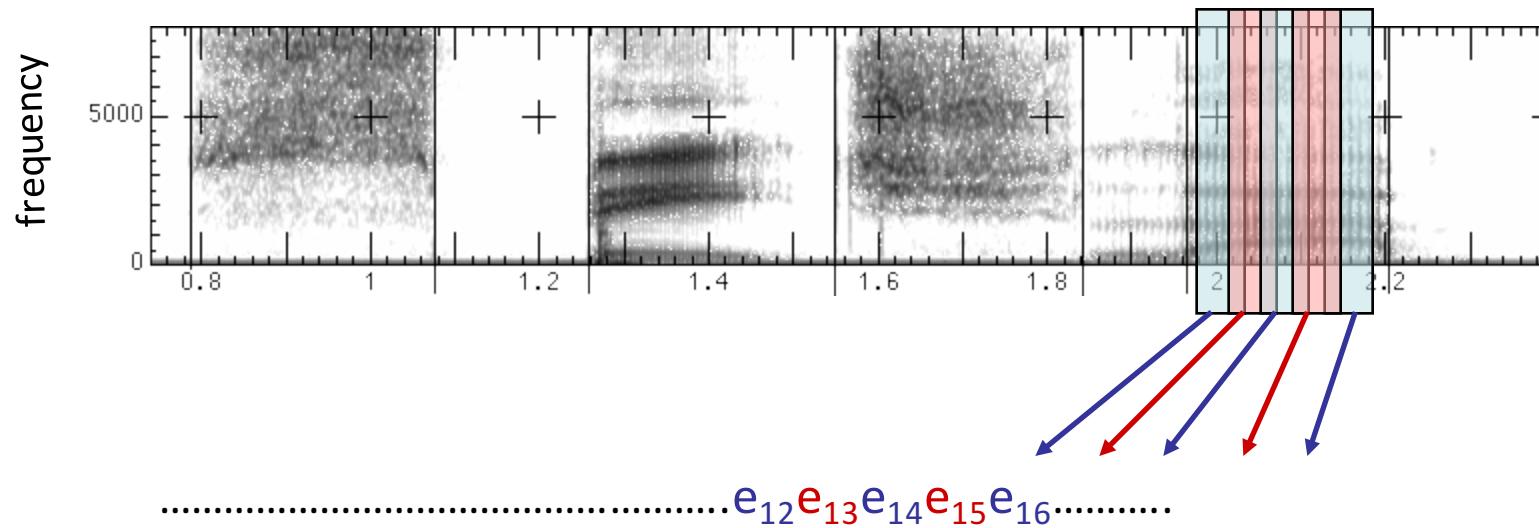


Vowel [i] sung at successively higher pitches



Acoustic Feature Sequence

- Time slices are translated into acoustic feature vectors (~39 real numbers per slice)

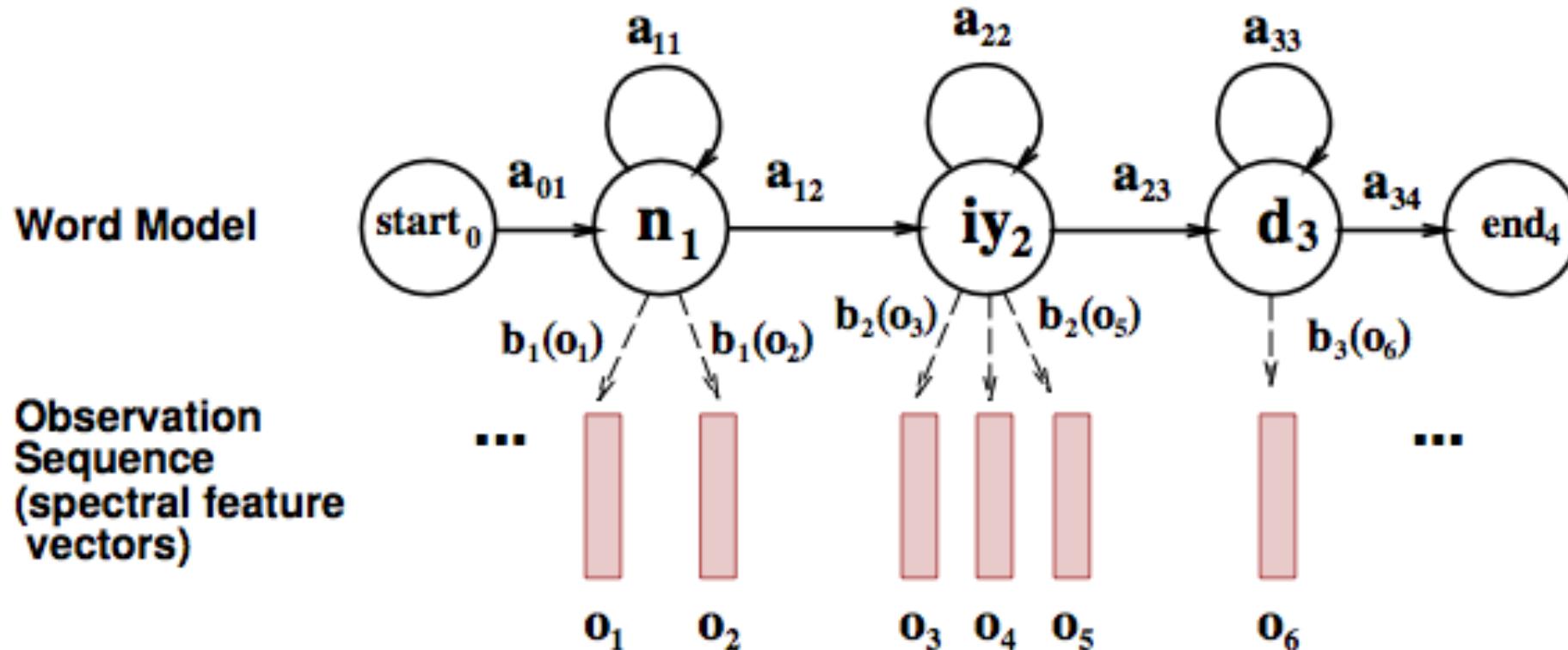


- These are the observations E , now we need the hidden states X

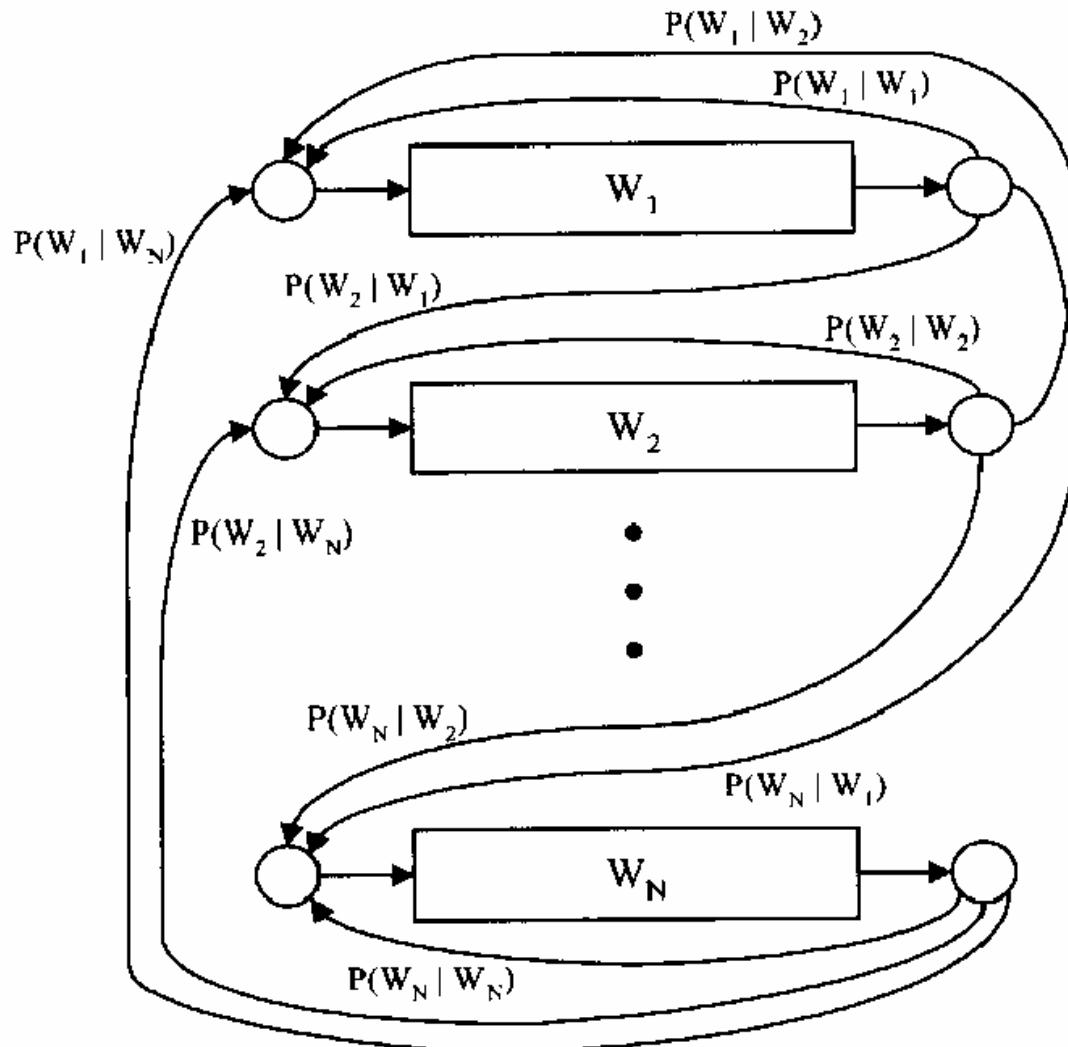
Speech State Space

- HMM Specification
 - $P(E|X)$ encodes which acoustic vectors are appropriate for each phoneme (each kind of sound)
 - $P(X|X')$ encodes how sounds can be strung together
- State Space
 - We will have one state for each sound in each word
 - Mostly, states advance sound by sound
 - Build a little state graph for each word and chain them together to form the state space X

States in a Word



Transitions with a Bigram Model



Training Counts

198015222	the first
194623024	the same
168504105	the following
158562063	the world
...	
14112454	the door

23135851162	the *

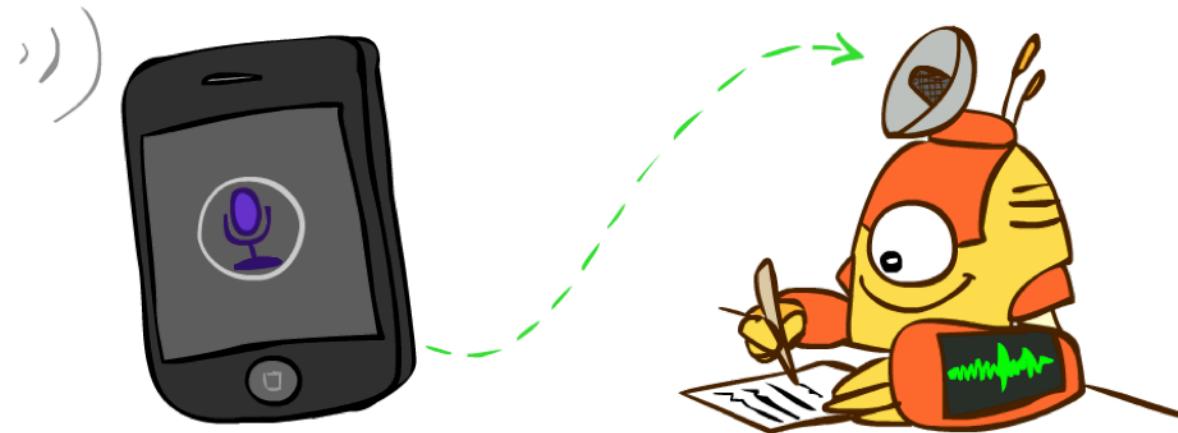
$$\hat{P}(\text{door}|\text{the}) = \frac{14112454}{23135851162}$$
$$= 0.0006$$

Decoding

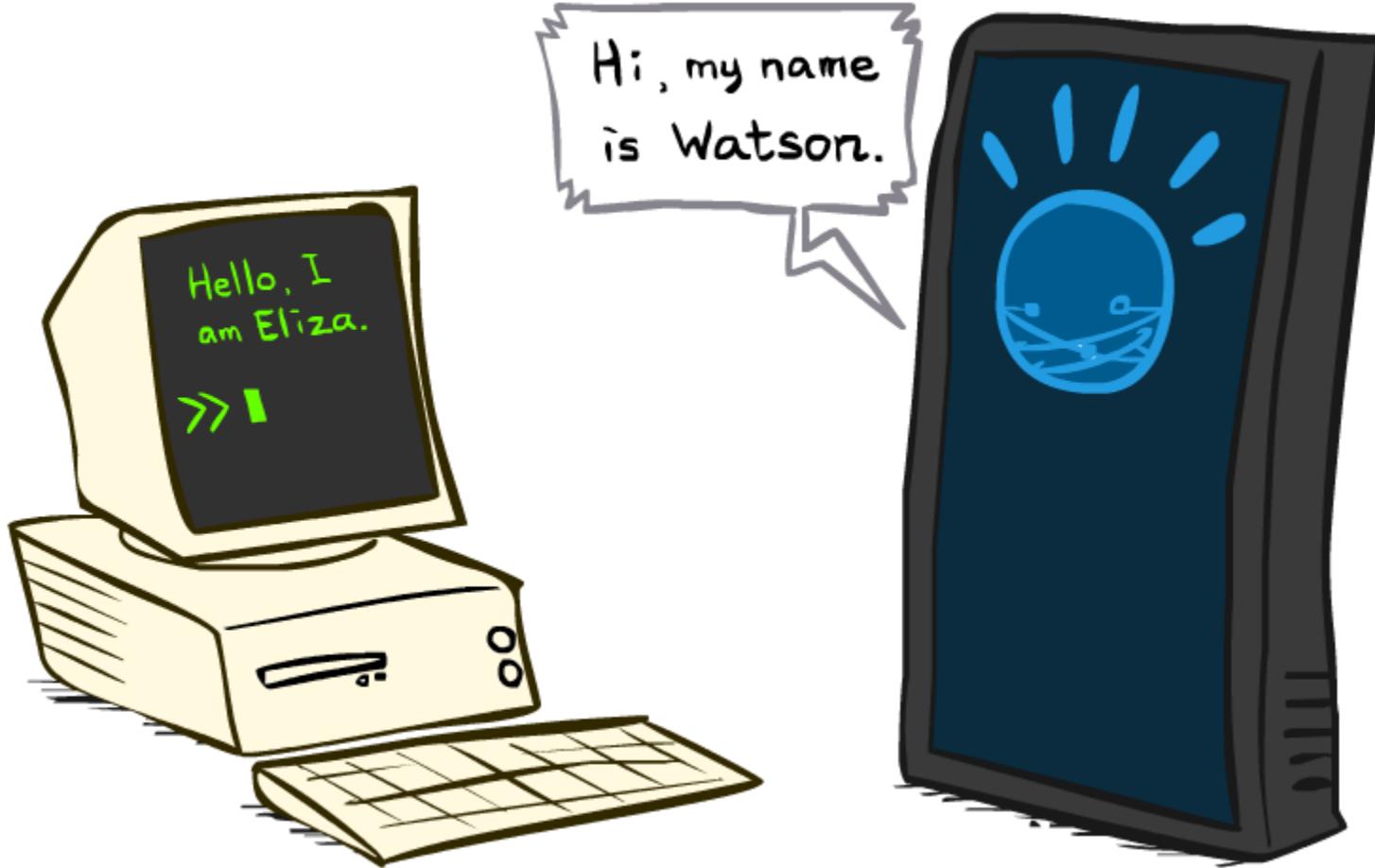
- Finding the words given the acoustics is an HMM inference problem
- Which state sequence $x_{1:T}$ is most likely given the evidence $e_{1:T}$?

$$x_{1:T}^* = \arg \max_{x_{1:T}} P(x_{1:T}|e_{1:T}) = \arg \max_{x_{1:T}} P(x_{1:T}, e_{1:T})$$

- From the sequence x , we can simply read off the words



Dialog Systems



ELIZA



- A “psychotherapist” agent (Weizenbaum, ~1964)
- Led to a long line of chatterbots
- How does it work:
 - Trivial NLP: string match and substitution
 - Trivial knowledge: tiny script / response database
 - Example: matching “I remember __” results in “Do you often think of __”?
- Can fool some people some of the time?

Dialogue systems in the news



Be careful what you wish for....

Dialogue systems in the news



Dialogue systems in the news

- Microsoft deletes ‘teen girl’ AI after it became a Hitler-loving sex robot within 24 hours
- Microsoft silences its new AI bot Tay, after Twitter users teach it racism
- Microsoft is deleting its AI chatbot’s incredibly racist tweets
- Microsoft’s racist chatbot returns with drug-smoking Twitter meltdown

It's not all bad

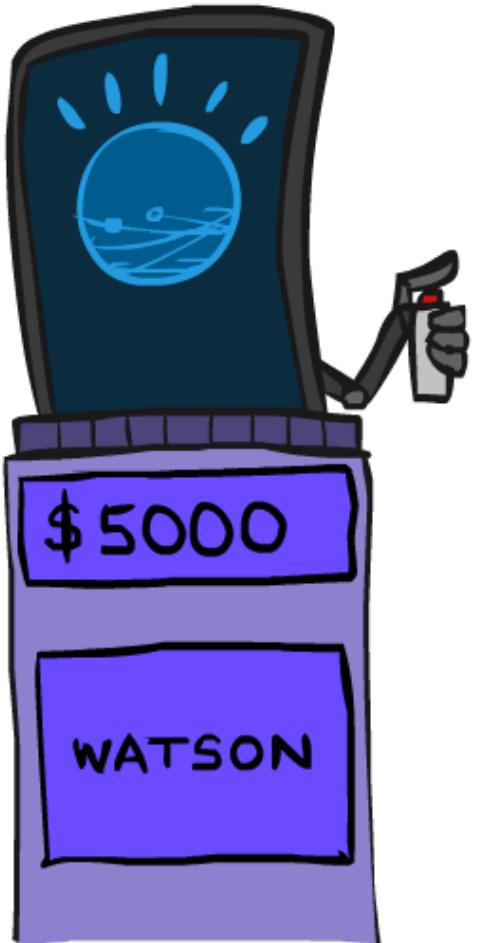


What's in Watson?

- A question-answering system (IBM, 2011)
- Designed for the game of Jeopardy
- How does it work:
 - Sophisticated NLP: deep analysis of questions, noisy matching of questions to potential answers
 - Lots of data: onboard storage contains a huge collection of documents (e.g. Wikipedia, etc.), exploits redundancy
 - Lots of computation: 90+ servers
- Can beat all of the people all of the time?



Watson



"a camel is a horse designed by"

About Wiktionary Entry Discussion Read Edit History Search Log in / create account

a camel is a horse designed by a committee

Contents [hide]
1 English
1.1 Alternative forms
1.2 Pronunciation

The Phrase Finder

> Discussion Forum

Google Custom Search Search

A camel is a horse designed by committee

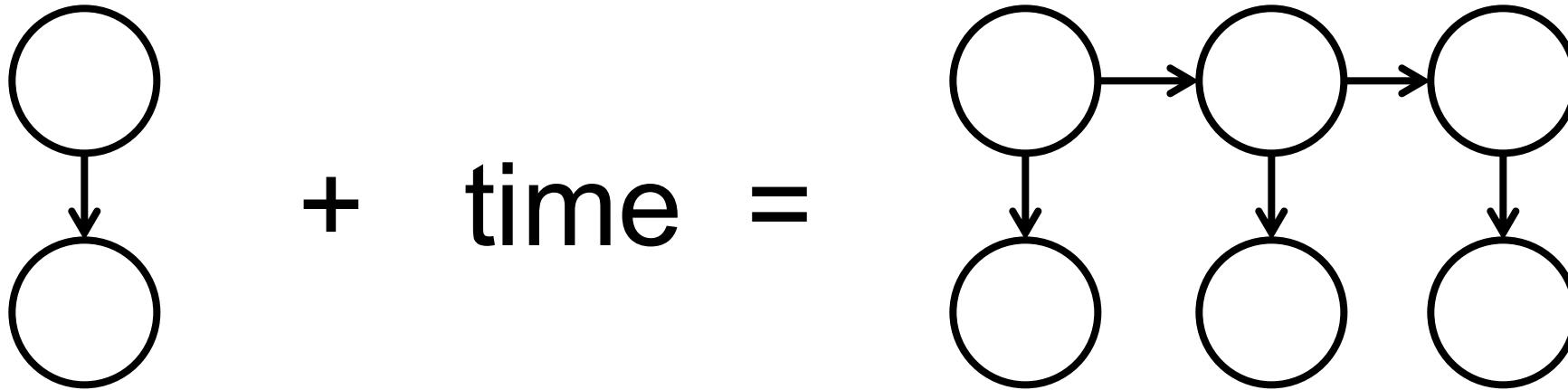
Posted by Ruben P. Mendez on April 16, 2004

Does anyone know the origin of this maxim? I heard it way back at the United Nations, which is chockfull of committees. It may have originated there, but I'd like an authoritative explanation. Thanks

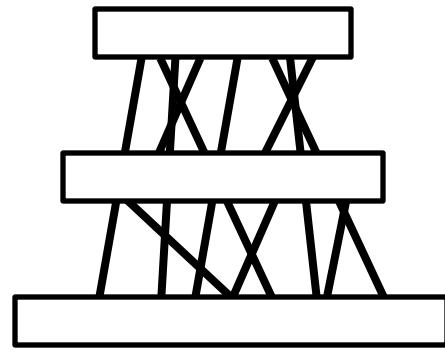
- Re: A camel is a horse designed by committee SR 16/April/04
 - Re: A camel is a horse designed by committee Henry 18/April/04

If a camel is a horse de
If a camel is a horse designed by committee then what's this contemporary Routemaster?

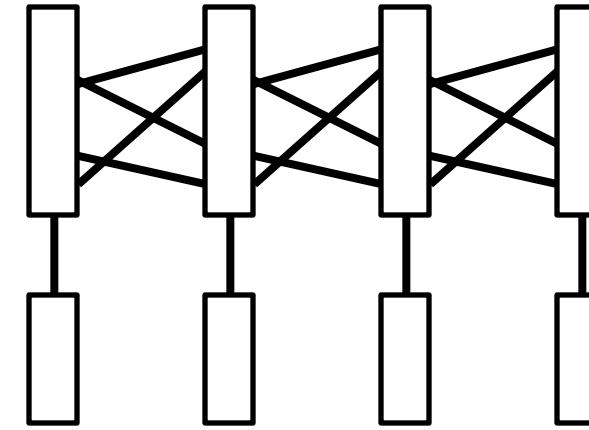
Modern techniques: recurrent neural nets



Modern techniques: recurrent neural nets



+ time =



Modern problems: Language & Vision



What is in the sheep's ear?

tag



What color is she wearing?

white