

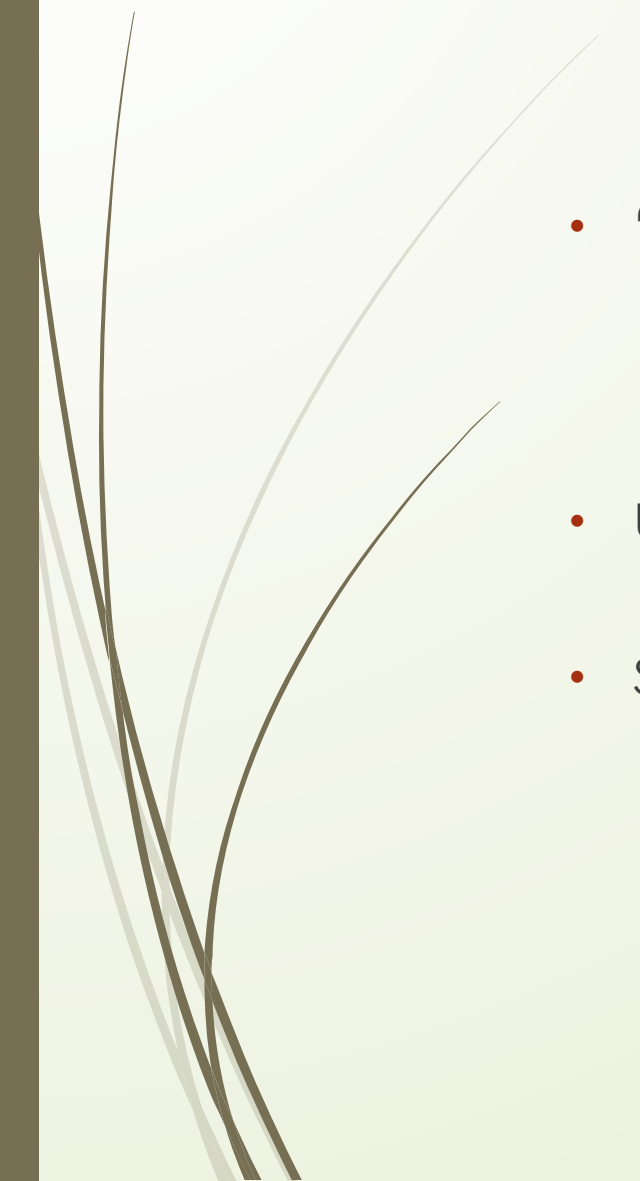


NATURAL LANGUAGE PROCESSING

Presented by: Aseem Upadhyay (Grad no. 7)



What is NLP?

- “Natural” languages
 - English, Hindi, Mandarin, French, Swahili, Arabic, Nahuatl,
 - NOT Java, C++, Perl, ...
 - Ultimate goal: Natural human-to-computer communication
 - Sub-field of Artificial Intelligence, but very interdisciplinary
 - Computer science, human-computer interaction (HCI), linguistics, cognitive psychology, speech signal processing (EE), ...
- 

SHALL WE PLAY A GAME?

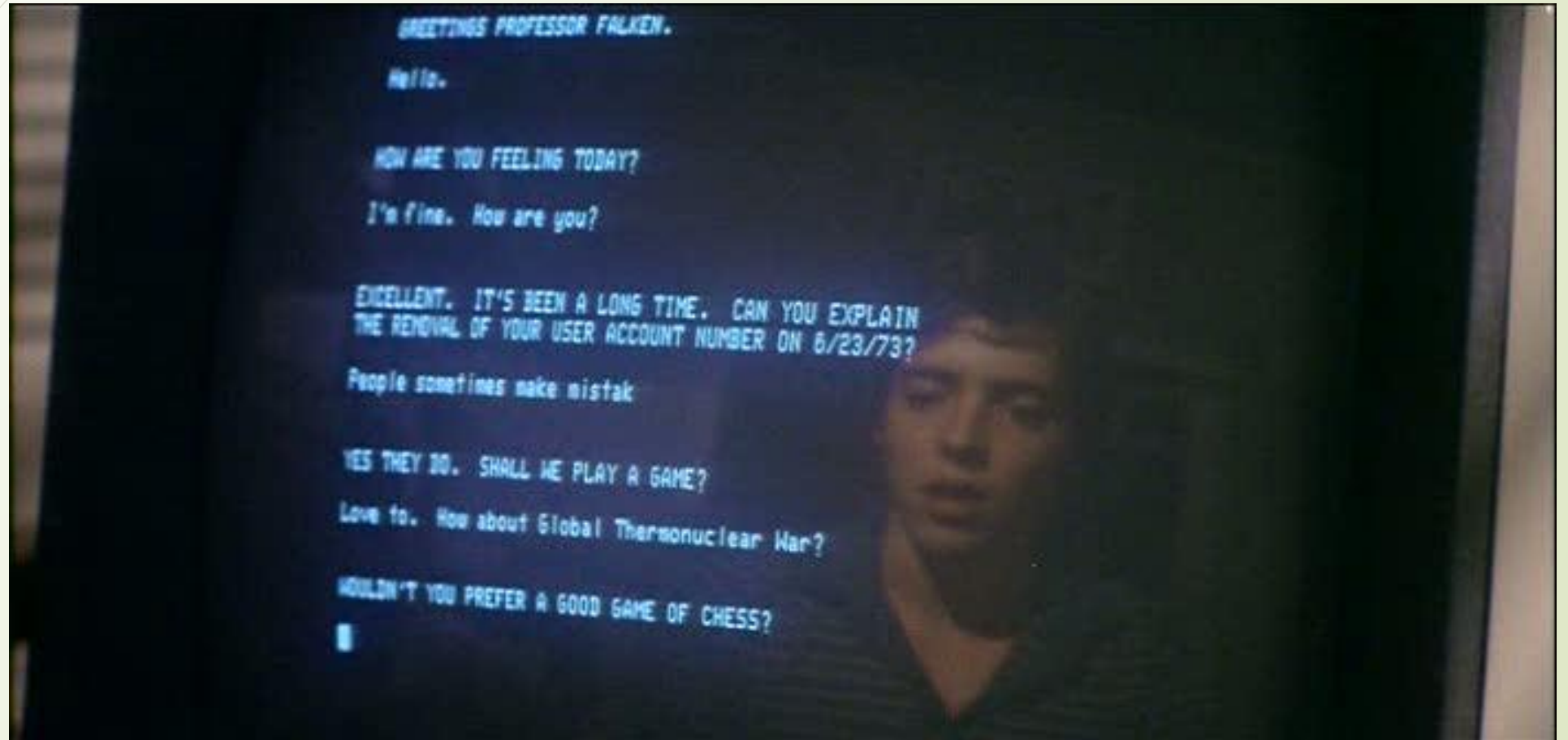


Image from WARGAMES (1983)

REAL WORLD NLP

Ask how tall is the statue of liberty? **Search**

Web Images News Videos Maps Q&A Beta More ▾

Related searches for how tall is the statue of liberty?

Statue of Liberty Facts	Where Is the Statue of Liberty Located	History of the Statue of Liberty
Why Is the Statue of Liberty Green	When Was the Statue of Liberty Built	Statue of Liberty Original Color

Answer [The Statue of Liberty](#) is 305 feet tall (93 meters) from the ground to the tip of the flame

adios amigo in english

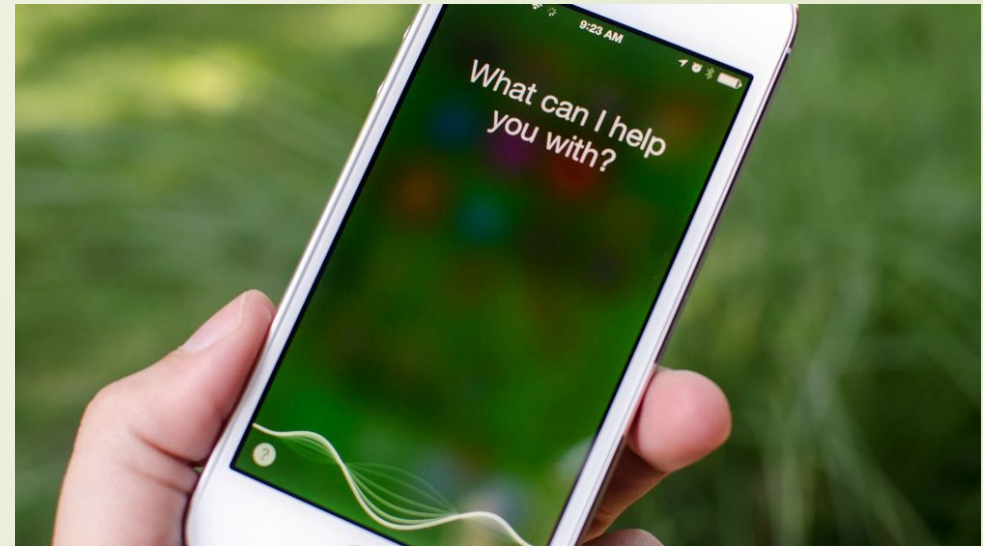
All News Videos Images Shopping More Settings Tools

About 2,570,000 results (0.78 seconds)

Spanish - detected English

adios amigo Edit goodbye friend

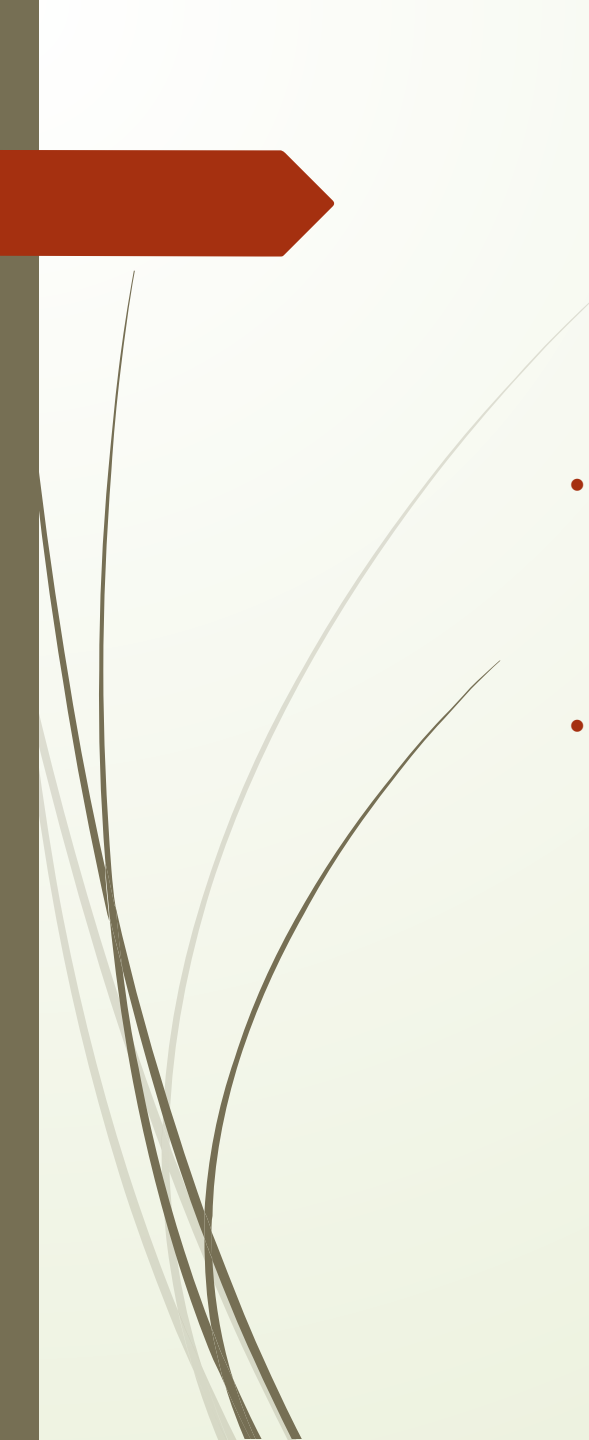
Open in Google Translate Feedback



How Does NLP work?

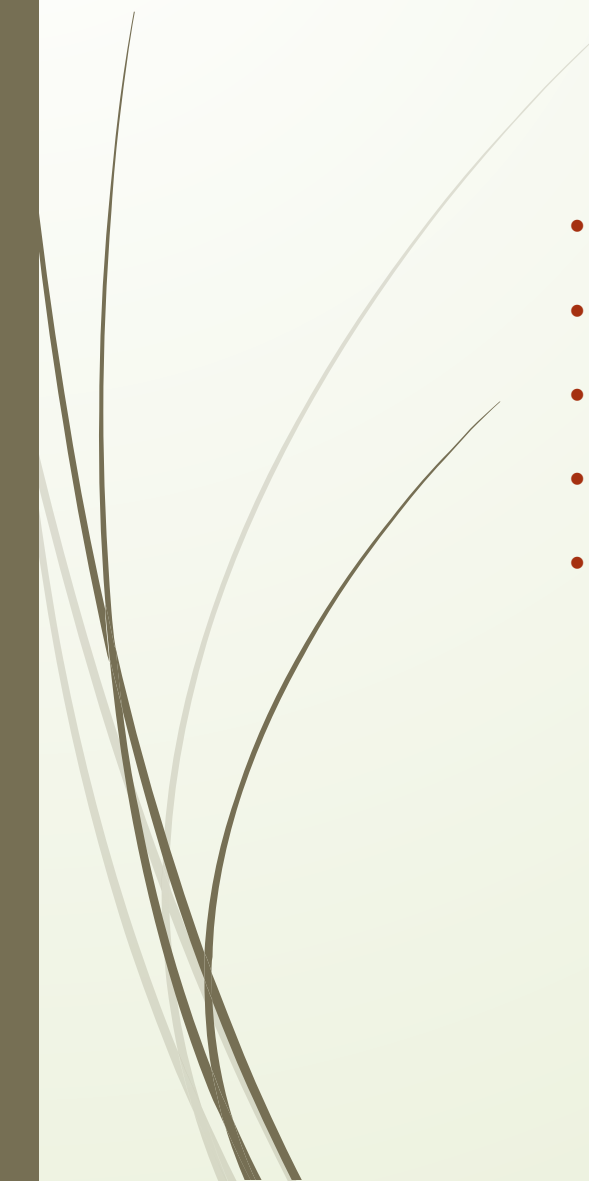
- Always two parts : Understanding and Generation
- Morphology : Identification of the structure of a word, such as the root word, suffixes, prefixes etc.
- Lexicography: What does each word mean?
 - He plays bass guitar.
 - That bass was delicious!
- Syntax: How do the words relate to each other?
 - The dog bit the man. ≠ The man bit the dog.
 - But in Russian: человек собаку съел = человек съел собаку



- 
- Semantics: How can we infer meaning from sentences?
 - I saw the man on the hill with the telescope.
 - Discourse: How about across many sentences?
 - President Bush met with President-Elect Obama today at the White House. He welcomed him, and showed him around.
 - Who is “he”? Who is “him”? How would a computer figure that out?



Why is NLP hard?

- Highly ambiguous at all levels
 - Complex and subtle use of context to convey meaning
 - Fuzzy
 - Involves knowledge about the world
 - Understanding how people interact with each other (persuasion, sarcasm, insulting etc.)
- 



Why NLP is difficult: Newspaper headlines

- Ban on Nude Dancing on Governor's Desk
- Iraqi Head Seeks Arms
- Juvenile Court to Try Shooting Defendant
- Teacher Strikes Idle Kids
- Stolen Painting Found by Tree
- Local High School Dropouts Cut in Half
- Red Tape Holds Up New Bridges
- Clinton Wins on Budget, but More Lies Ahead
- Hospitals Are Sued by 7 Foot Doctors
- Kids Make Nutritious Snacks
- Minister Accused Of Having 8 Wives In Jail

Image taken from one of Dr. Chris Manning's presentations



Question answering

A: And, what day in May did you want to travel?

C: OK, uh, I need to be there for a meeting that's from the 12th to the 15th.

Note that client did not answer question.

- Meaning of client's sentence:
 - Meeting
 - Start-of-meeting: 12th
 - End-of-meeting: 15th
 - Doesn't say anything about flying!!!!
- How does agent infer client is informing him/her of travel dates?

what is 1 meter in feet?



All News Images Videos Shopping More Settings Tools

About 16,100,000 results (0.48 seconds)

Length =
Meter

[More info](#)

[Feedback](#)

How old is stephen hawking?



All Images News Videos Shopping More Settings Tools

About 3,890,000 results (0.85 seconds)

Stephen Hawking / Age

75 years

January 8, 1942



what is the temperature outside?



All Maps News Shopping Books More Settings Tools

About 751,000,000 results (0.85 seconds)

Houghton, MI

Monday 9:00 PM

Clear

 39 °F | °C

Precipitation: 15%

Humidity: 37%

Wind: 8 mph

Temperature

Precipitation

Wind


- May want to ask questions about non-English, non-text documents... and get responses back in English text.

why did the us enter ww2?

All Videos News Images Maps More Settings Tools

About 2,620,000 results (0.72 seconds)

Although the war began with Nazi Germany's attack on Poland in September 1939, the United States **did not enter** the war until after the Japanese bombed the **American** fleet in Pearl Harbor, Hawaii, on December 7, 1941.



U.S. soldiers take cover under fire in Germany, WWII

[World War II \(1939-1945\) - The George Washington University](https://www2.gwu.edu/~erpapers/teaching/glossary/world-war-2.cfm)
<https://www2.gwu.edu/~erpapers/teaching/glossary/world-war-2.cfm>

About this result • Feedback

People also ask

- Why did the US get involved in World War 2? ▾
- What brought us into World War 2? ▾
- When did the Japanese enter the war? ▾
- What events led to the United States entry into World War 2? ▾

Feedback



Machine Translation

- **About \$10 billion spent annually on human translation**
- Hotels in Beijing, China

In Chinese: 昨天我打电话订的时候艺龙信誓旦旦的保证说是四星级的酒店,住进去以后一看没,我靠,这在80年代可能算得上是四星的,我要的是368的大床房,房间只有一个0.5米*1米的小窗户,打开一看,我靠, ...

In English: Yesterday, I called out when Art Long vowed to ensure that the four-star hotel, to live in. I see no future, I rely on it in the 80s may be regarded as a four-star, and I want the big 368-bed Room, the room is only one 0.5 m * 1-meter small windows, what we can see, I rely on, ...



Why is machine translation hard?

- Requires both understanding the “from” language and generating the “to” language.

Que hambre tengo yo

What hunger have I

I've got that hunger

I am so hungry

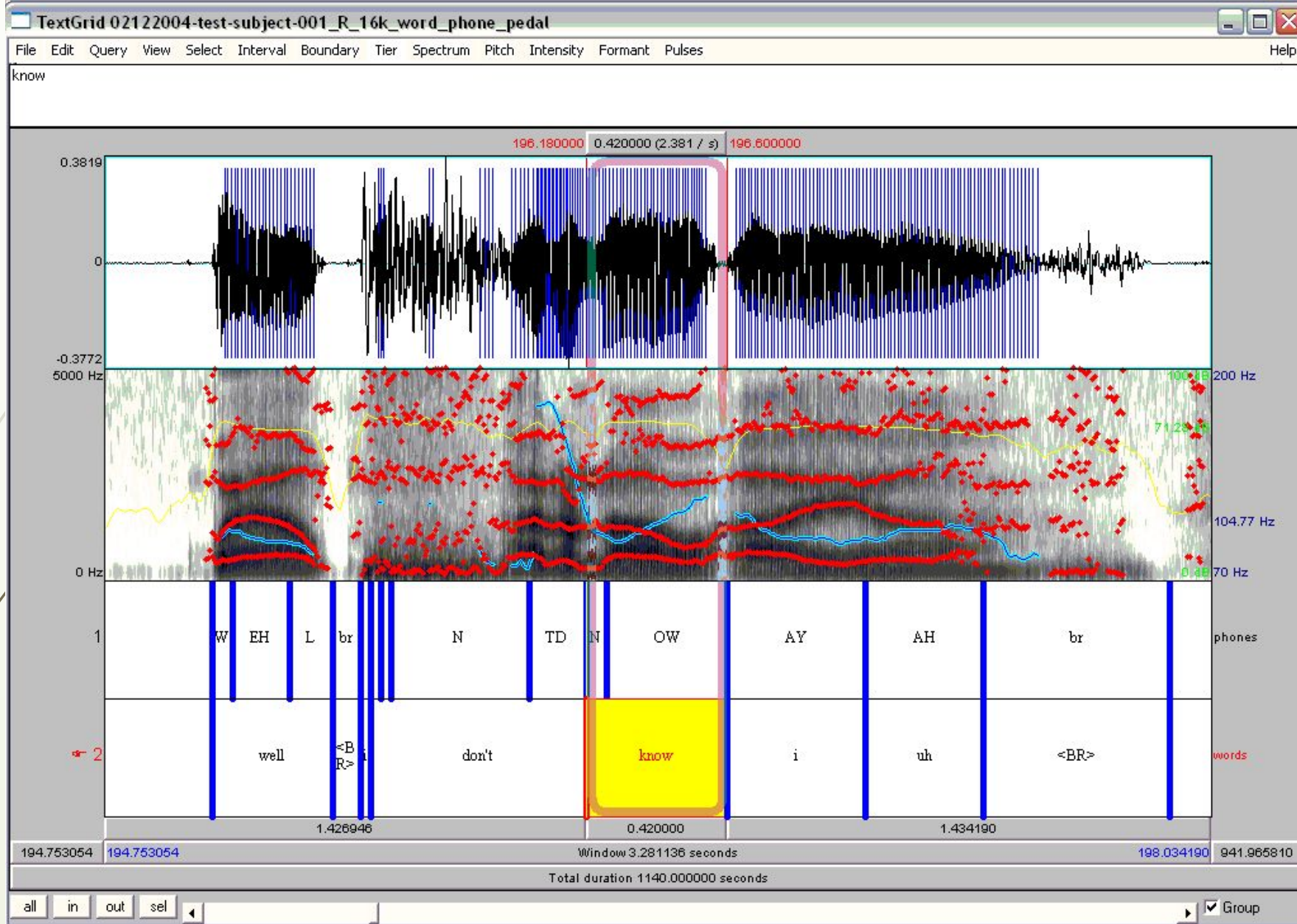
- How can we teach a computer a “second language” when it doesn't even really have a first language?



Speech Processing



- Speech Recognition
 - Automatic dictation, assistance for blind people, text-to-speech, speech-to-text ...
- Factors that affect speech recognition ...
 - How does intonation affect semantic meaning?
 - Detecting uncertainty and emotions
 - Detecting deception!
- Why is this hard?
 - Each speaker has a different voice (male vs female, child versus older person)
 - Many different accents (Scottish, American, non-native speakers) and ways of speaking
 - Conversation: turn taking, interruptions, ...



Example from one of Dr. Julia Hirschberg's presentations

Summarization

- Two approaches : Extraction and Abstraction
- Due to the problem of information overload i.e. availability of excess information, which hides the desired part of the information, the need for summarization is also increasing.
- The challenge here is that the summary should not miss out on any of the important elements or lose the actual meaning of the original document.



CS4760 & CS5760: Human-Computer Interactions & Usability
Skip to content

Home Announcements Assignments Lectures Resources Policies

The usability test sessions must be scheduled three weeks before the actual testing because many individuals must be coordinated and space made available.

Clear

Summarize

The usability test sessions must be scheduled three weeks before the actual testing because many individuals must be coordinated and space made available.

Each team member should attend at least 5 usability test sessions.

I will email you access to google form that the participates will use to reserve a test session.

While you make your schedule, you need to also consider the location of your testing.

Make sure you have access to these classrooms during your scheduled hour for your usability tests.

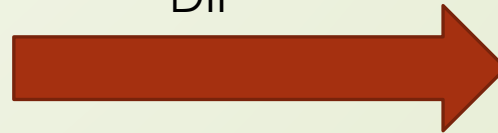
Before fill out your schedule in the Google Form, you should check the sheets for other grad students to see if they have already reserved the location.

Assisted Text Input

- Various approaches provide for detecting and recognizing text to enable a user to perform various functions or tasks.
- For example, a user could point a camera at an object with text, in order to capture an image of that object.
 - This image can be digitally processed, and it's meaning extracted



DIP



TEXT or VOICE generation





References



- Christopher D. Manning. 1991. Lexical Conceptual Structure and Marathi. ms. Stanford University, Stanford CA.
- Christopher D. Manning. 1995. Ergativity: Argument Structure and Grammatical Relations. Paper presented at the 69th annual meeting of the Linguistic Society of America, New Orleans.
- Joan Bresnan, Shipra Dingare, and Christopher D. Manning. 2001. Soft Constraints Mirror Hard Constraints: Voice and Person in English and Lummi. *Proceedings of the LFG01 Conference*, pp. 13-32, Hong Kong
- Roger Levy and Christopher D. Manning. 2003. Is it harder to parse Chinese, or the Chinese Treebank? *ACL 2003*, pp. 439-446.
- Julia Hirschberg and Christopher D. Manning. 2015. Advances in natural language processing. *Science* 349(6):261-266.
- Christopher D. Manning. 2016. Texting and Talking ... with Language-Understanding Computers? *Boao Review*.
- R. Mihalcea. 2004. "Graph-based ranking algorithms for sentence extraction, applied to text summarization." In Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics (ACL 2004) (companion volume), Barcelona, Spain.
- www.cs.columbia.edu/~julia/talks/afosr14.pdf
- cse.unl.edu/~choueiry/S02-976/Davis-NLP-Overview_of.ppt