Spoken Language Processing

Julia Hirschberg

CS 4706

Speech Processing

- How do you produce sounds that other people interpret as language?
- How does your hearer decode what you are trying to convey?
- In a conversation, how do you know when it's your turn to talk?
- Once you decide ***what*** you want to say, how do you decide ***how*** to say it?:
  - How do you decide where to pause in the sentence? 
  - How do you decide what words to emphasize?  
  - How do you decide what intonational contour to use? 
  - How do you convey your own feelings and emotions?

  

Applications for Speech Technologies

- Speech synthesis (TTS): AT&T, IBM (Jeopardy 2/14-16), SitePal
- Speech recognition (ASR): Nuance
- Speech to Speech Translation
- Speech Search: Google Voice Search
- Homeland Security: Deception Detection, Dialect and Language ID, and Speaker ID, trust
- Spoken Dialogue Systems:
  - Over-the-phone services: Voice Actions for Android
  - Tutoring systems: KTH's Ville
  - Amtrak Julie (or here)

What will we do in this course?

- Learn about fundamental aspects of speech signals and how to analyze them
  - Phonemes and phones: sounds of a language
  - Acoustic/prosodic information: pitch, energy
  - Intonational contours
- Study two basic speech technologies, TTS and ASR and their application to Spoken Dialogue Systems (SDS)
  - Build your own SDS using the Festival and HTK Toolkits

Course information

- Course syllabus and readings (Jurafsky & Martin, second edition; Keith Johnson, second edition)
- Speech tools and speech lab
- Courseworks: discussion, course files, gradebook
- TAs: Bob Coyne, Erica Cooper

Projects and Exams

- Build and demo a Spoken Dialogue System
  - Teams of 3
    - Organize your own or
    - Advertise for team members on Courseworks discussion category “Find a team” or
    - Send mail to coyne@cs.columbia.edu
  - 4 Deadlines:
    - Project Description
    - TTS component (using Festival tools)
    - ASR component (using HTK tools)
    - SDS demos (during final exam period)
- Midterm (March 9) and final (May 2) in class

- Honor policy on syllabus
- Late policy on syllabus (5 ‘free’ late days per project for the semester)
- My office hours: M 4-6, CEPSR 705
- Bob Coyne: TBD, CEPSR 7LW1 (Speech Lab)
- Erica Cooper: TBD, CEPSR 7LW1 (Speech Lab)

Any Questions?