

# Experiences with Undergraduate Research (Computationally Analyzing Audio, Video, and Transcriptions of Team Conversations)

Professor Diane Litman

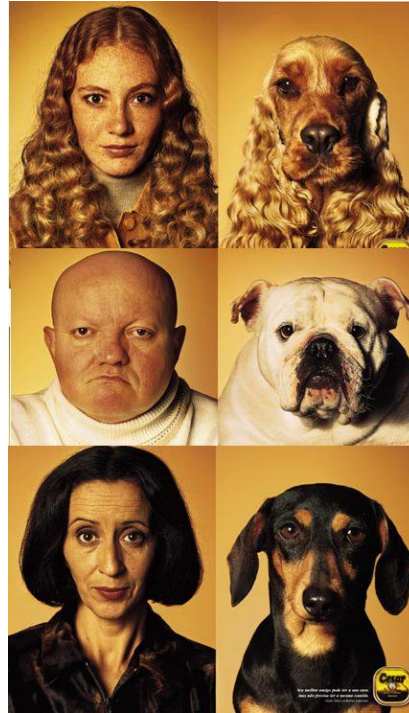
# Natural Language Processing (NLP)

- Getting computers to perform useful and interesting tasks involving **human languages**
  - languages such as English, Spanish, Chinese, etc.
  - as opposed to computer languages such as Python

# Why is NLP needed?

- An enormous amount of machine readable text, audio, and video is now available
- Conversational agents such as Siri and Alexa are becoming an important form of human-computer communication

# *Teams Project:* Entrainment and Task Success in Team Conversations



# *Teams Project: Entrainment and Task Success in Team Conversations*

- *Multi-party* entrainment measures that are *computable using NLP*
- Applications
  - Conversational agents
  - Browsers for (un)successful teams

# Experimentally Collected Data

- Experimental Design
  - Team Training or Not
  - First vs. Second Games
- Audio-Video
  - 47 hours
  - 63 teams
- Questionnaires
  - 216 individuals



# What teams say (transcriptions)

M: And then [I'm here.]

E: [Oh.]

P: [Yeah] probably wanna save [Whispering Garden.]

E: [Whispering- Yeah.]

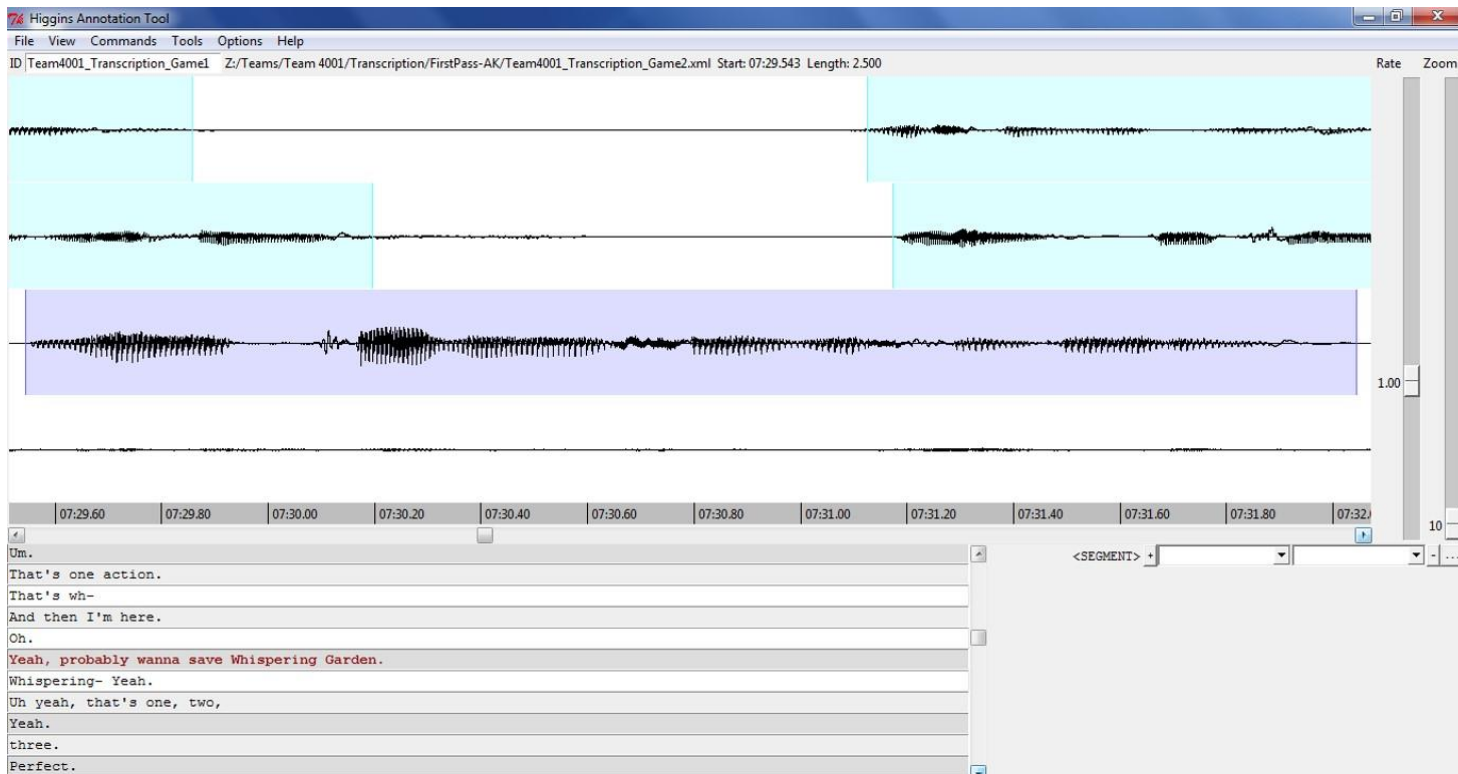
M: [Uh yeah, that's one,] [two,]

P: [Yeah.]

M: [three.]

P: [Perfect.]

# How teams say it (audio)



The screenshot displays the Higgins Annotation Tool interface. The window title is "Higgins Annotation Tool". The menu bar includes "File", "View", "Commands", "Tools", "Options", and "Help". The status bar shows "ID Team4001\_Transcription\_Game1 Z:/Teams/Team 4001/Transcription/FirstPass-AK/Team4001\_Transcription\_Game2.xml Start: 07:29.543 Length: 2.500".

The main area shows four audio waveforms. The top two waveforms are partially highlighted with light blue rectangular boxes. The third waveform is highlighted with a light purple rectangular box. The bottom waveform is not highlighted. A vertical zoom slider on the right is set to 1.00, with a maximum value of 10.00.

The timeline at the bottom shows time markers from 07:29.60 to 07:32.00 in increments of 0.20. Below the timeline is a transcript window with the following text:

Um.  
That's one action.  
That's wh-  
And then I'm here.  
Oh.  
Yeah, probably wanna save Whispering Garden.  
Whispering- Yeah.  
Uh yeah, that's one, two,  
Yeah.  
three.  
Perfect.

On the right side of the transcript window, there is a dropdown menu showing "<SEGMENT> +".



# How teams say it (video)

- Non-verbal communication
  - Gaze
  - Gesture
  - Facial expressions
  - Etc.

# Litman Lab Undergraduate Research

## Connecting

- I recruit (class performance, email to colleagues)
- I advertise (Pitt's First Experiences in Research Program)
- Students initiate contact
  - Ideal prerequisites: one or more of AI, NLP, ML

## Implementing

- *Employment*: REU funds from the National Science Foundation
- *Credit*: CS Capstone project
  - lab or office space
  - Individual and group meetings

# Selected Prior Projects / Outcomes

- ***Exploiting Word-level Features for Emotion Prediction***
  - First-authored publications
  - CS Day Award for Best Undergraduate Student Poster
  - Honorable Mention in the Computing Research Association's Outstanding Undergraduate Award
  - MS from Brown University
- ***In the Zone: Towards Detecting Student Zoning Out using Supervised Machine Learning***
- ***Examining the Impacts of Dialogue Content and System Automation on Affect Models in a Spoken Tutorial Dialogue System***
  - First-authored publications
  - CS Day Award for Best Undergraduate Student Poster
  - National Science Foundation Graduate Research Fellowship
  - now at U of Toronto
- ***Differences in User Responses to a Wizard-of-Oz versus Automated System***
  - First-authored publications
  - Honorable Mention in Computing Research Association's Outstanding Undergraduate Researcher Award
  - now at U of Texas