CS 6474 Social Computing: Analyzing Language II

Munmun De Choudhury

munmund@gatech.edu

Week 9 | October 15, 2018

Midterm Presentations

- On Oct 29 and Oct 31
- 10 teams in all
- Each team gets 15 minutes
  - 10 minutes for actual presentation
  - 5 minutes for Q&A
  - All team members need to attend both days
  - Any combination can present
- Schedule at the following link (also on class website): https://docs.google.com/spreadsheets/d/1DYt53JeQhR_F6RTHi-ciBDoW5lyiVmVh547h2CkE8zY/edit#gid=0

Midterm Presentations

- Introduction and motivation
- Background
- Project Goals
- Expected outcomes
- Prior Work
- Work accomplished so far
- Projected plan

Gender and Power: How Gender and Gender Environment Affect Manifestations of Power

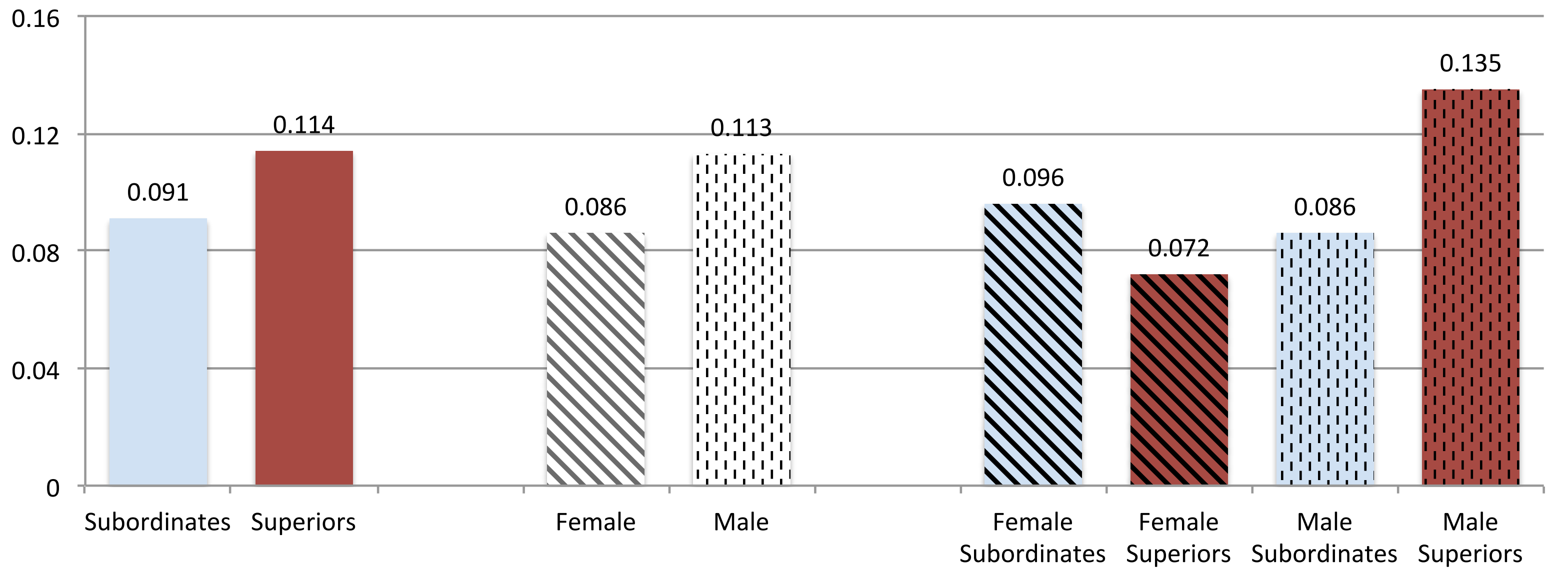
Summary

- Interplay between gender, gender environment of online conversations and power
- Contributions:
  - Automatic gender assignment of 87% of the Enron corpus – US Social Security Administration list of names matching the approx. age range of Enron employees; first names gathered from email headers
  - Test the sociolinguistic hypotheses: face-saving use of language, and to the use of language to strengthen social relations
  - Gender-based features boosts the accuracy of predicting the direction of power between pairs of email interactants

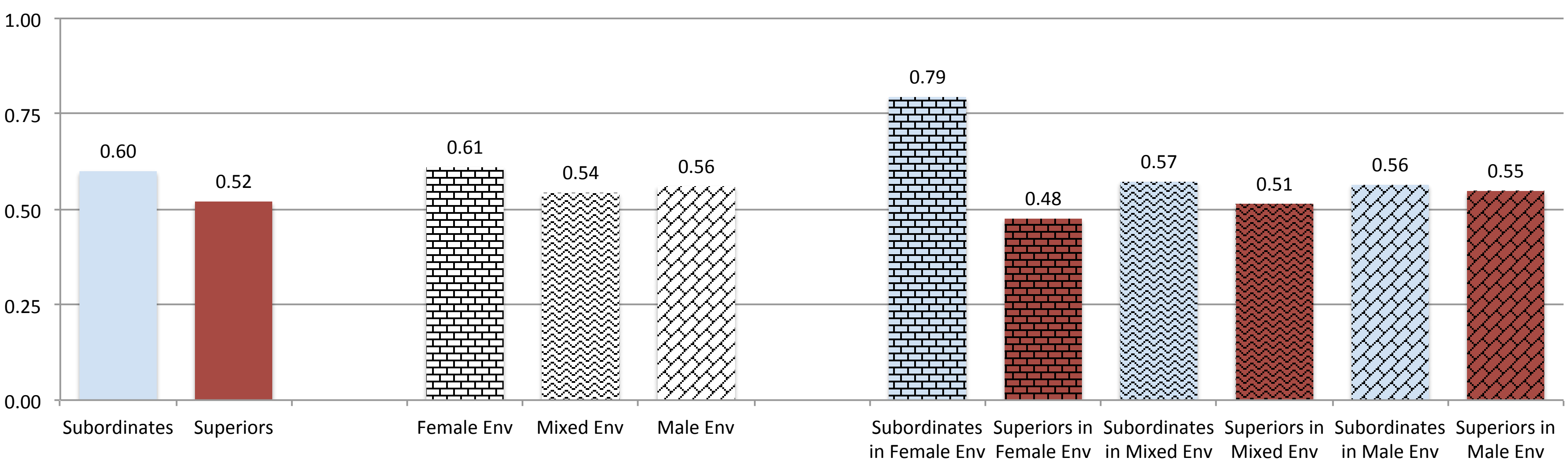
Summary

- Hypothesis 1: Female superiors tend to use “face-saving” strategies at work that include conventionally polite requests and impersonalized directives, and that avoid imperatives (Herring, 2008).

- Hypothesis 2: Women use language to create and maintain social relations, for example, they use more small talk (based on a reported “stereotype” in (Holmes and Stubbe, 2003)).



Overt display of power



Conventional Counts

Would these results hold in modern enterprise social media contexts?

As people in corporations increasing adopt platforms like Facebook and Twitter, how do you expect these findings to generalize/change?

How do the preexisting social/corporate structures and biases at a large corporation impact the data, especially given that less than 35% of the unique discourse participants were women?

The style of communication depends on the context or environment

How can the sociolinguistic findings on gender, gender environment, and power be useful for social computing research?

Ex: The power framework provided an interpretable and actionable set of hypotheses that could apply productively to other social situations, such as the difference in moderator vs. user behavior in an online forum.

Ex: Danescu-Niculescu-Mizil et al. identified connections between linguistic coordination and social power relations using discussions among Wikipedians and arguments before the U. S. Supreme Court

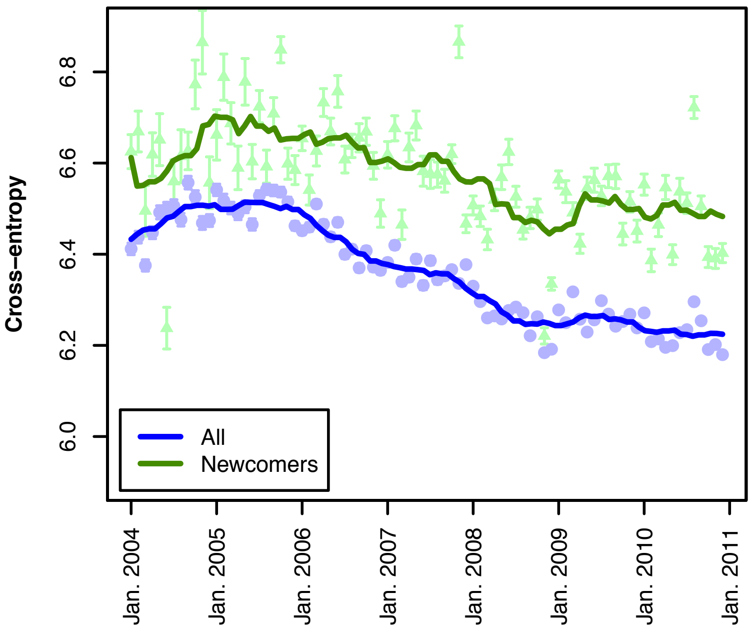
Class Exercise I

What kind of design considerations could incorporate the sociolinguistic findings on gender, gender environment, and power? What would they enable/what are the advantages?

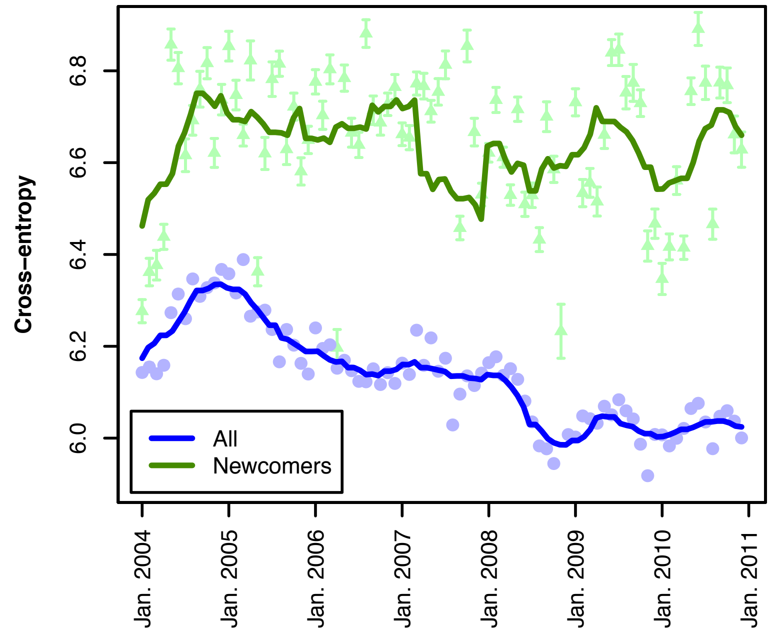
No Country for Old Members: User lifecycle and linguistic change in online communities

Summary

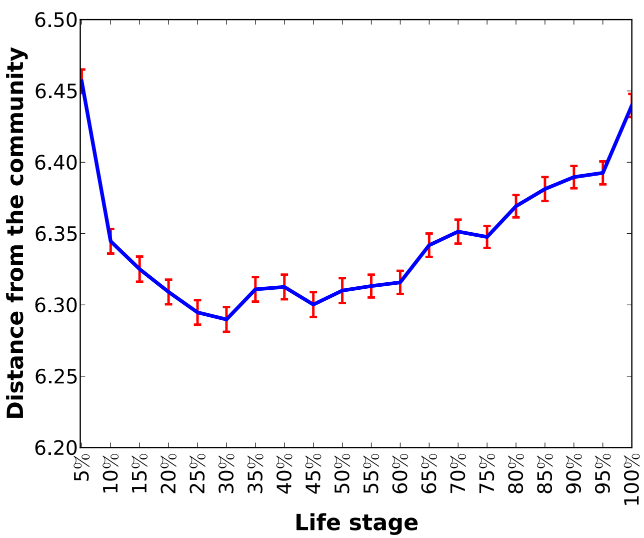
- The paper proposes a framework for tracking linguistic change as it happens in a community, to understand how specific users react to the community's evolving norms
- Results show a two-stage lifecycle of linguistic change in communities (RateBeer and BeerAdvocate):
  - a linguistically innovative learning phase in which users adopt the language of the community
  - a conservative phase in which users stop changing and the evolving community norms pass them by



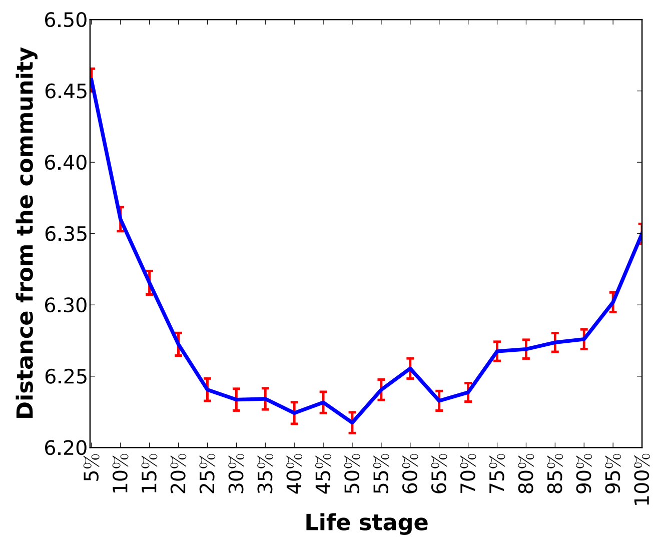
(a) BeerAdvocate



(b) RateBeer



(a) BeerAdvocate



(b) RateBeer

Class Exercise II

Danescu-Niculescu-Mizil et al. say that "[their] framework can be used to detect, early in a user's career, how long she will stay active in the community"

Describe two scenarios where this knowledge will be beneficial. Who are these stakeholders who can derive benefit?

How do evolving linguistic norms impact participation in anonymous communities?

Will the two-phase lifecycle (linguistic innovation learning and conservative phases) hold for/generalize to other online communities?

Chancellor, S., Pater, J. A., Clear, T., Gilbert, E., & De Choudhury, M. (2016, February). *# thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities*. In Proc. CSCW.