Week 14 Readings
C. Hidey
April 22, 2015

[PIP2]Daniel Jurafsky, Rajesh Ranganath, Daniel A. McFarland (2009)
Jurafsky et al. attempt to identify aspects of personality in the context of speed dating. They used a speed dating corpus created in 2005 and segmented acoustic information and transcribed speech with marked turns. They used a number of prosodic features for amplitude and pitch such as the range and standard deviation. They added lexical features based on the LIWC categories and added 2 additional features for story-telling and content about the date itself. They also included dialog features and disfluency features such as questions, discourse, or pauses. They trained a logistic regression classifier and found that they were able to do better than random chance. I thought this paper was very interesting and I was surprised they were able to obtain fairly high F-score using only speech and linguistic features. I would expect that they could obtain better performance by including other social cues like facial expressions, gestures, or physical contact. One negative about the paper was the assumption that the reader is familiar with acoustic and prosodic features and that could have benefited from some explanation. Also, it would have been interesting to compare human performance on this task.

[PIP3] Andrew Rosenberg and Julia Hirschberg. (2008)
Rosenberg and Hirschberg examine charisma in speech. They created a corpus by annotating public speeches from 9 Democratic Presidential candidates with adjectives such as accusatory, passionate, intense or trustworthy, reasonable, believable. They extracted features and modeled correlation with various attributes of charisma. This experiment was repeated using different subjects using text only. They found that there was some consistency in which speakers were rated highly for charisma. For both experiments, they used lexico-syntactic features such as number of words, content words, repeated words, and number of pronouns. This paper was very detailed, and the experiment should be reproducible. One way that I think the paper could have been improved would be to discretize the human judgments and attempt to build a predictive model to classify the charisma of a subject.