Amber Boydstun, University of California at Davis: Issue Framing as a Generalizable Phenomenon

Framing—portraying an issue from one perspective to the necessary exclusion of alternative perspectives—is a central concept in political communication. It is also a powerful political tool, as evidenced through experiments and single-issue studies beyond the lab. Yet compared to its significance, we know very little about framing as a generalizable phenomenon. Do framing dynamics, such as the evolution of one frame into another, play out the same way for all issues? Under what conditions does framing influence public opinion and policy? Understanding the general patterns of framing dynamics and effects is thus hugely important. It is also a serious challenge, thanks to the volume of text data, the dynamic nature of language, and variance in applicable frames across issues (e.g., the 'innocence' frame of the death penalty debate is irrelevant for discussing smoking bans). To address this challenge, I describe a collaborative project with Justin Gross, Philip Resnik, and Noah Smith. We advance a unified policy frames codebook, in which issue-specific frames (e.g., innocence) are nested within high-level categories of frames (e.g., fairness) that cross cut issues. Through manual annotation bolstered by supervised learning, we can track the relative use of different frame cues within a given issue over time and in an apples-to-apples way across issues. Preliminary findings suggest our work may help unlock the black box of framing, pointing to generalizable conditions under which we should expect to see different types of framing dynamics and framing effects.

Ed Chi, Google: Location and language use in social media

We now know that social interactions are critical in many knowledge and information processes. In this talk, I plan to illustrate a model-driven approach to understanding social behavior around user location and different languages in social media.

First, in 2010, we performed the first in-depth study of user location field in Twitter user profiles. We found that 34% of users did not provide real location information, frequently incorporating fake locations or sarcastic comments that can fool traditional geographic information tools. We then performed a simple machine learning experiment to determine whether we can identify a users location by only looking at contents of a user's tweets. We found that a users country and state can in fact be determined easily with decent accuracy, indicating that users implicitly reveal location information, with or without realizing it.

Second, despite the widespread adoption of Twitter in different locales, little research has investigated the differences among users of different languages. In prior research, the natural tendency has been to assume that the behaviors of English users generalize to other language users. We studied 62 million tweets collected over a four-week period. We discovered cross-language differences in adoption of features such as URLs, hashtags, mentions, replies, and retweets. We also found interesting patterns of how multi-lingual Twitter users broker information across these language boundaries. We discuss our works implications for research on large-scale social systems and design of cross-cultural communication tools.

Justin Grimmer, Stanford University: Creating and destroying party brands

Party brands are central to theories of Congressional action. While previous work assumes that a party's brand—its long run reputation—is a direct consequence of the content of legislation, in this presentation I show how partisans from both parties use public statements to craft their own party's reputation and to undermine their opponents party. The incentive to craft and destroy brands varies across legislators, creating systematic distortions in who contributes to partisan branding efforts,

what is said about a party's brand, and when partisan criticism becomes salient. To demonstrate the construction of party brands I use new collections of newsletters from Congressional offices, along with press releases, floor speeches, and media broadcasts. Across the diverse sources, I show that ideologically extreme legislators are the most likely to explain their party's work in Washington and the most likely to criticize the opposing party—particularly when their is an opposing party president. Extreme legislators also engage in more vitriolic criticism of the opposing party, particularly when opposing presidents are unpopular. The result is that parties in rhetoric appear even more combative and polarized in public debate outside Congress than inside Congress.

Lillian Lee, Cornell University: Is it all in the phrasing? Computational explorations in how we say what we say, and why it matters

Louis Armstrong (is said to have) said, "I don't need words — it's all in the phrasing". As someone who does natural-language processing for a living, I'm a big fan of words; but lately, my collaborators and I have been studying aspects of phrasing (in the linguistic, rather than musical sense) that go beyond just the selection of one particular word over another. I'll describe some of these projects in this talk. The issues we'll consider include: Does the way in which something is worded in and of itself have an effect on whether it is remembered or attracts attention, beyond its content or context? Can we characterize how different sides in a debate frame their arguments, in a way that goes beyond specific lexical choice (e.g., "pro-choice" vs. "pro-life")? The settings we'll explore range from movie quotes that achieve cultural prominence; to posts on Facebook, Wikipedia, Twitter, and the arXiv; to framing in public discourse on the inclusion of genetically-modified organisms in food.

Joint work with Lars Backstrom, Justin Cheng, Eunsol Choi, Cristian Danescu-Niculescu-Mizil, Jon Kleinberg, Bo Pang, Jennifer Spindel, and Chenhao Tan.
References:

- ACL 2012: `http://www.mpi-sws.org/~cristian/memorability.html`

- EXPROM 2012: `https://confluence.cornell.edu/display/llresearch/HedgingFramingGMOs`

- WSDM 2013: `http://www.cs.cornell.edu/home/llee/papers/convcuration.home.html`

- ACL 2014: `http://chenhaot.com/pages/wording-for-propagation.html`

- ACL 2014 short paper: `http://chenhaot.com/pages/statement-strength.html`

Philip Resnik, University of Maryland: "I Want to Talk About, Again, My Record On Energy ...": Modeling Agendas and Framing in Political Debates and Other Conversations

Computational social science has been emerging over the last several years as a hotbed of interesting work, taking advantage of, to quote Lazer et al. (Science, v.323), "digital traces that can be compiled into comprehensive pictures of both individual and group behavior, with the potential to transform our understanding of our lives, organizations, and societies." Within that larger setting, I'm interested in how language is used to influence people, with an emphasis on computational modeling of agendas (who is most effectively directing attention, and toward what topics?), framing or "spin" (what underlying perspective does this language seek to encourage?), and sentiment (how does someone feel, as evidenced in the language they use)? These questions are particularly salient in political discourse. In this talk, I'll present recent work looking at political debates and other conversations using Bayesian models to capture relevant aspects of the conversational dynamics,

as well as new methods for collecting people's reactions to speeches, debates, and other public conversations on a large scale.

This talk includes work done in collaboration with Jordan Boyd-Graber, Viet-An Nguyen, Deborah Cai, Amber Boydstun, Rebecca Glazier, Matthew Pietryka, Tim Jurka, and Kris Miler.

Sali Tagliamonte, University of Toronto: Sociolinguistics for computational social science

In recent years, a major growth area in applied natural language processing has been the application of automated techniques to massive datasets in order to answer questions about society, and by extension people. Sociolinguistics, which combines anthropology, statistics and linguistics (e.g. Labov 1994, 2001), studies linguistic data in order to answer key questions about the relationship of language and society. Sociolinguists focus on frequency and patterns in linguistic usage, correlations, strength of factors and significance, which together reveal information about the sex, age, education and occupation of speakers/writers but also their history, culture, place of residence, social relationships and affiliations. The findings arising from this type research offer important insights into the nature of human organizations at the global, national or community level. They also reveal connections and interactions, the convergence and divergence of groups, historical associations and developing trends.

In this paper, I will introduce Sociolinguistic research and the nature of sociolinguistic field techniques and sample design. I will argue that socially embedded data is critical for analyzing and discovering social meaning. Then, I will summarize the findings of several case studies. What does the use of a 3rd singular morpheme -s, as in (1), tell us about the history and culture of a community (Tagliamonte 2012, 2013)? How is quotative be like, (2), spreading in geographic space (Tagliamonte to appear)? What is the mechanism that underlies linguistic change (Tagliamonte & D'Arcy 2009) and by extension cultural trends and projections?

1. The English people speaks with grammar.

2. I was *like*, "Hey how are you going?" And hes *like*, "Im fine."

Using sociolinguistic datasets, the answers to these questions have successfully addressed prevailing puzzles and offered solutions to real world problems. However this type of research is only be as good as the quality of the data, the capability of the technologies for extracting and analyzing what is important, and the relevance of the socially cogent and statistically sound interpretations. I will argue that Sociolinguists and Computational Scientists could be powerful allies in uncovering the complex structure of language data and in so doing, offer unsurpassed insight into varying human states and conditions.

References:

- Labov, William (1994). *Principles of linguistic change: Volume 1: Internal factors.* Cambridge and Oxford: Blackwell Publishers.

- Labov, William (2001). *Principles of linguistic change: Volume 2: Social factors.* Malden and Oxford: Blackwell Publishers.

- Tagliamonte, Sali A. (2012). *Variationist Sociolinguistics: Change, observation, interpretation.* Malden and Oxford: Wiley-Blackwell.

- Tagliamonte, Sali A. (2013). *Roots of English: Exploring the history of dialects.* Cambridge: Cambridge University Press.

- Tagliamonte, Sali A. (to appear). System and society in the evolution of change: the view from Canada. In Green, E. & Meyer, C. (Eds.), *Faces of English*. Berlin, Germany: De Gruyter-Mouton.

- Tagliamonte, Sali A. & D'Arcy, Alexandra (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language* 85(1): 58–108.