Effects of Trump's Campaign Rhetoric on Local Language

Keis Bejgo, Dylan Lewis, Luis Solis

Massachusetts Institute of Technology

Main Puzzle

There have been concerns that nationalist, right-wing sentiments have gained momentum over the past three and a half years of the Trump presidency. We were interested in understanding how Trump's rhetoric may influence public sentiment on a regional level. To this end, we analyzed tweets from 4 cities and Florida where Trump gave speeches during his campaign. We focused particularly on how words' frequencies in tweets by locals are affected by Trump's speeches.

Methods

We gathered Trump's campaign speeches from the Trump Campaign Corpus github repository. To isolate Trump specific rhetoric in his campaign speeches from general GOP rhetoric, we chose to normalize Trump's campaign speeches on the campaign speeches of the 2012 GOP candidate, Mitt Romney. When determining important words in documents, we also normalized by the number of total words in the document, lemmatized according to parts of speech, and removed stopwords. We utilized the SentimentAnalysis R package when determining the sentiments of words in Trump's speeches. We decided to use a radius of about 8 km for cities to account for more conservative users that live outside of city limits.

Descriptive Analysis

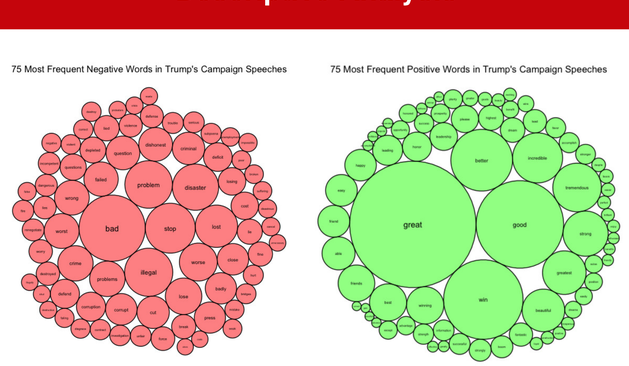

Fig. 1: Most Frequent Negative and Positive Words in Trump's Campaign Speeches

Trump's positive sentiment words tend to be adjectives with "great" far exceeding the rest. Among words with negative sentiment, there are more meaningful words related to his speech topics such as "investigation", "defense", "deficit", & "press".

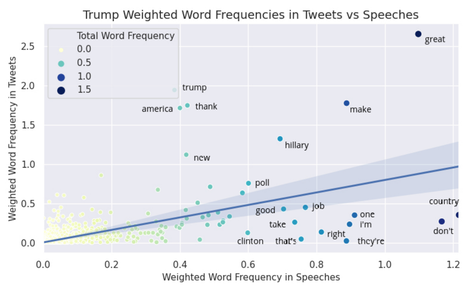

Fig. 2: Word Frequencies in Trump's Tweets vs Speeches

To ensure that Trump's twitter and speech language did not diverge significantly from each other we found the correlation between the weighted word frequencies of words in his tweets vs those in his speeches. We found a correlation of about 0.62, indicating the two do not diverge significantly from each other.

Results

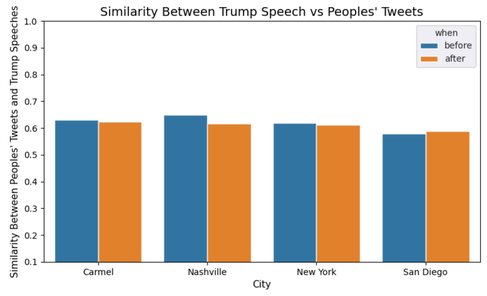

Fig. 3: Twitter Users' Language Similarity to Trump's Speech Rhetoric 2 Weeks Before vs After the Speech

We hypothesized that twitter users' rhetoric in red-state cities would become more similar to Trump's speech rhetoric in the 2 weeks following his speech in the city. We analyzed tweets from 2 red state and 2 blue state cities before and after Trump's speech in the city in order to account for differences in political leaning. We see no significant difference between the two, however due to the small sample size of tweets that we were allowed to access via the Twitter API, our results are not statistically significant and further sampling and testing will have to be done.

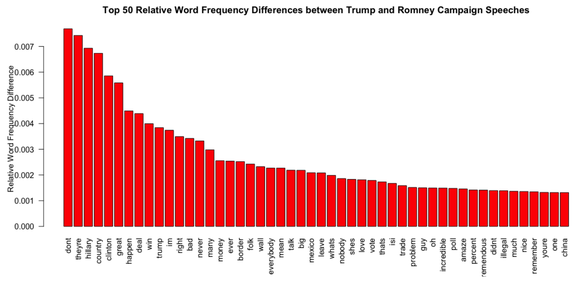

Fig. 4: Trump's Most Frequent Words across Campaign Speeches

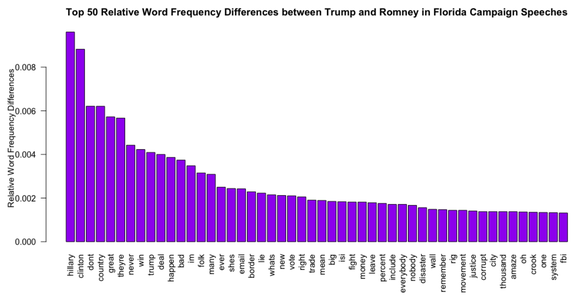

Fig. 5: Trump's Most Frequent Words across Florida Campaign Speeches

In Trump's speeches across the entire campaign, his most frequent words, normalized on Romney's campaign speeches, include, "don't", "hillary", "great", "deal", as well as words related to his election platform such as "border", "wall", "mexico", "isis", "trade", and "china". Words used to thwart Hillary Clinton's campaign such as "hillary", "email", "lies", "corrupt", "crook", and "fbi" in regards to Clinton's email scandal appear more frequently in Trump's Florida campaign speeches than across all of his campaign speeches, showing that in swing states, Trump strategizes to mention the scandal more frequently to win voters to tip the scale.

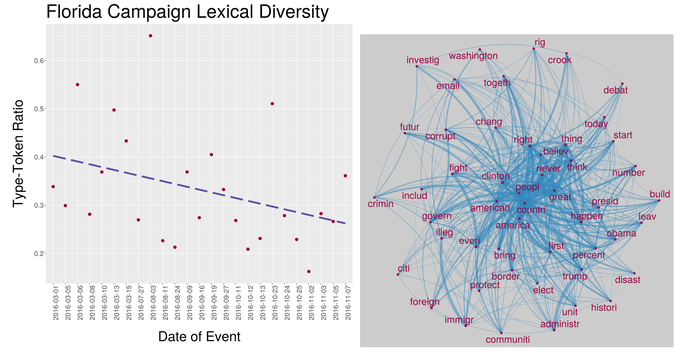

Fig. 6: Florida lexical, semantic analysis

Trump's presidential campaign focused on a small number of topics: immigration, government distrust, and the "average American." Trump's voter base applauded his 'politically incorrect' rhetoric, believing him an unconventional candidate who would not fail them as previous politicians had. Focusing on Florida over the entirety of the campaign season, Trump's lexical diversity weakened, suggesting his speeches focused on a smaller range of topics. Both Trump and Hillary campaigned heavily in Florida, but Trump ultimately won the state in the general election. The previous figure illustrates the top 50 (stemmed) words in Trump's speeches—with words closer to the center being more frequent—and how strongly connected they are with other words.

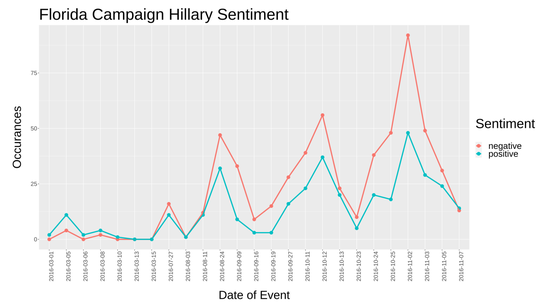

Fig. 7: Trump sentiment analysis of Hillary Clinton

The figure above illustrates a sentiment analysis of "Hillary Clinton" in Trump's Florida speeches, with occurrence and negative sentiment spiking in the days before the election. This suggests Trump's staunch criticism of Hillary—particularly her email scandal—might have dissuaded citizens from voting Democrat, landing him the election in Florida.

Future Work

We hope to use the twint package or greater access to the Twitter API to increase our sample size of tweets and analyze the effects of Trump's language on a greater number of U.S. campaign locations in order to draw more statistically meaningful results.

Acknowledgements

Thank you to Professor Kim and the course staff for providing us with the knowledge and assistance to pursue this project. Thank you to Twitter for giving us developer access to use the Twitter API.