Social Identity and Punctuation Variation in the #BlueLivesMatter and #BlackLivesMatter Twitter Communities

Rachael Tatman & Amandalynne Paullada

Department of Linguistics, University of Washington

The 33rd Northwest Linguistics Conference
University of British Columbia, May 5-7 2017

Introduction

- ► The social factor of political affiliation is reflected in speech [1]
- ► Sociolinguistic variation has been shown to occur in text, including on Twitter [2, 3]
- ► Does this extend to sublexical features, like punctuation?
  - ▷ Punctuation as a feature useful for authorship identification, personality traits [4, 5, 6, 7]

> **Do people with different social identities use punctuation differently on Twitter?**

Two communities: #blacklivesmatter & #bluelivesmatter

#blacklivesmatter

- ► Activist movement begun in response to killing of Trayvon Martin
- ► One goal of the movement: ending police brutality, especially towards the Black community

#bluelivesmatter

- ► Activist movement begun in response to killings of two NYPD officers
- ► Focus on supporting police; apparent opposition to #blacklivesmatter

Both are large, geographically diffuse communities that participate in conversations on Twitter using these hashtags to identify themselves.

Example Tweets:

#blacklivesmatter

- ► "being a sleepy girl with a busy life is hard"
- ► "i got in the wrong car the other day and my dad won't stop tormenting me"
- ► "Here's my #ShortStoryChallenge it's a Lil personal because my dad was a victim of police brutality but #BlackLivesMatter"

#bluelivesmatter

- ► "OOPS! I nearly followed a #BlackLivesMatter supporter. Quick! Pour bleach on me so I can get clean!! #AllLivesMatter #BlueLivesMatter"
- ► "@user I don't care if you believe me or not. If you don't believe #BlueLivesMatter, your opinion has no value, anyway."

Which community does Trump tweet like?
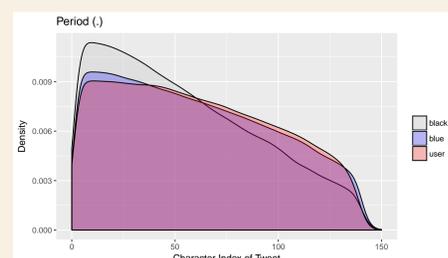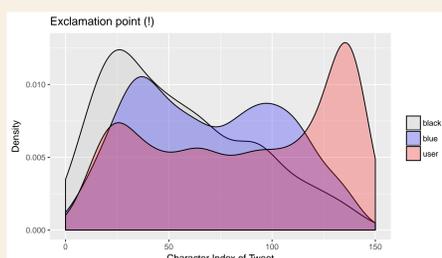
Data collection

- ► Account identification:
  - ▷ Collect usernames using Twitter's web-based interface to search for accounts associated with #blacklivesmatter and #bluelivesmatter
  - ▷ Hashtags occured in either users' names or bios
  - ▷ Identified 98 #blacklivesmatter, 90 #bluelivesmatter accounts
- ► Sampling from accounts:
  - ▷ Up to 50 tweets from each account sampled using public Twitter API (8 Dec., 2016) (some accounts had fewer tweets available through the public API)
- ► Final data set:
  - ▷ 5170 total tweets:
    - ► 2031 from 76 #blacklivesmatter accounts
    - ► 3139 from 68 #bluelivesmatter accounts

Punctuation patterns over Tweet length
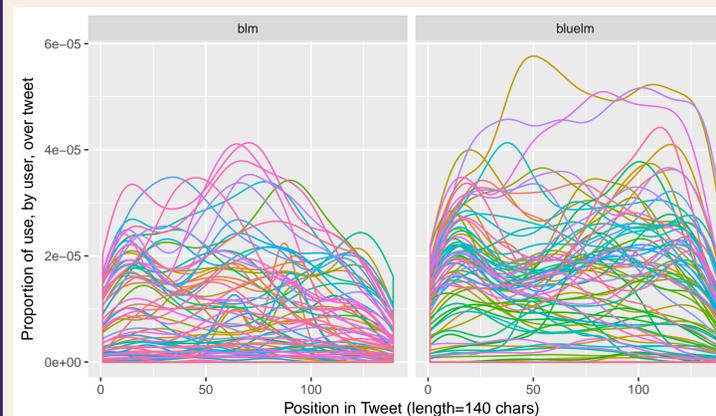
- ► #Blacklivesmatter:
  - ▷ Users use less punctuation
  - ▷ Relatively uniform distribution across tweets
- ► #Bluelivesmatter:
  - ▷ Users use more punctuation overall
 - ▷ More punctuation at the beginning and end of tweets

Distinguishing lexical features

- ► 'Fist' emoji unique to #blacklivesmatter Tweets
- ► Unicode 'white star' symbol unique to #bluelivesmatter Tweets
- ► Tweets collected 8 Dec. 2016 – only #bluelivesmatter Tweets mention Pearl Harbor

#blacklivesmatter	#bluelivesmatter
standingrock	pjnet
girlposts	bluelivesmatter
nodapl	maga
[userName]	foxnews
[userName]	ussarizona
justin	thinblueline
whitney	harbor

Top 7 most frequent words unique to each community.

Some individual punctuation marks
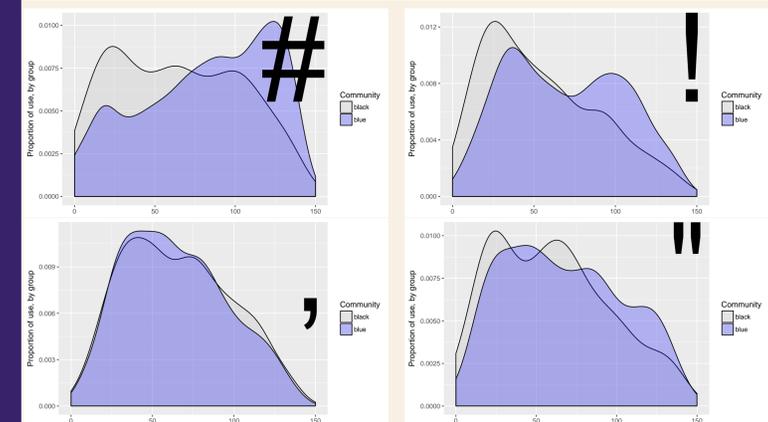
References

[1] William Labov. *Principles of linguistic change, cognitive and cultural factors*, volume 3. John Wiley & Sons, 2011.

[2] Jacob Eisenstein. Phonological factors in social media writing. In *Proc. of the Workshop on Language Analysis in Social Media*, volume 3, pages 11–19, 2013.

[3] Tyler Schnoebelen. Do you smile with your nose? stylistic variation in twitter emoticons. *University of Pennsylvania Working Papers in Linguistics*, 18(2):14, 2012.

[4] Carole E Chaski. Who's at the keyboard? authorship attribution in digital evidence investigations. *International journal of digital evidence*, 4(1):1–13, 2005.

[5] Ahmed Abbasi and Hsinchun Chen. Applying authorship analysis to extremist-group web forum messages. *IEEE Intelligent Systems*, 20(5):67–75, 2005.

[6] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. The development and psychometric properties of liwc2015. Technical report, 2015.

[7] Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. Predicting personality from twitter. In *Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on*, pages 149–156. IEEE, 2011.

Email:rctatman@uw.edu, paullada@uw.edu | Twitter:@rctatman, @amandalynneP