The President on Twitter: A Characterization Study of @realDonaldTrump

Brooke Auxier(✉) and Jennifer Golbeck

University of Maryland, College Park, MD 20742, USA
{bauxier, jgolbeck}@umd.edu

Abstract. US President Donald Trump is perhaps the most powerful man on Twitter in terms of both his office and his ability to impact world events through his tweets. The way he uses the platform is unusual for someone in his position and is divisive among US citizens. Some tweets are posted by staff while others are posted by Trump himself, and in the time period of our dataset, the platform used to post distinguishes the author. We use this data to study the behavioral characteristics of the tweet sources and the public reaction to this content. Trump tweets tended to be more focused on himself or and other people, rather than the audience, and are more negative, angry, and anxious than staffers' tweets. Liberals and conservatives alike found some of the tweets inappropriate for someone in Trump's position to be posting, and the majority of inappropriate tweets came from Trump himself. The language characteristics are so distinctive that they may be used in a predictive model to correctly classify a tweet's author with 87% accuracy. Our predictive model will low for authorship determination, even when platform information is not informative, and our analysis suggests directions for future research on the rise of populist candidates and how they communicate on social media.

Keywords: Twitter · Trump

1 Introduction

Donald Trump was a prolific Twitter user at @realDonaldTrump before he became a candidate with a serious chance at the US Presidency. Throughout the 2016 Presidential primaries, his formal candidacy, his transition, and his early presidency, people questioned whether he would change his brusk style to better fit with what might be expected of someone in his position. He has continued to use the platform in a way that certainly breaks from how high-ranking officials typically conduct themselves online, and while some appreciate his style as "refreshing", others find it improper and disrespectful to the institution of the Presidency.

The Trump tweets that garner the most media attention tend to be the more blustery or controversial ones, but there is a lot of nuance to be found in the Trump Twitter stream. As has been widely reported and analyzed, Trump tweets are authored by multiple people. Having staff-authored tweets is a common practice, but in this case, the author of the tweet appears to be identifiable based on the platform. Trump posted (at least in the time we were collecting our data) from an Android and his staff posted from

G.L. Ciampaglia et al. (Eds.): SocInfo 2017, Part I, LNCS 10539, pp. 377–390, 2017.
DOI: 10.1007/978-3-319-67217-5_23

iPhone. This original analysis was posted online in non-peer reviewed analysis[1], but the text and social media analysis community online has supplied a lot of supportive evidence for this theory. Comments from staff also suggest it is true and scholarly literature has begun accepted this insight (e.g. Ott 2017). While we recognize platform may not be a perfect proxy for authorship, it appears to be a solid theory and we rely on it in this paper. Our prediction results discussed later add strength to the theory as well.

Knowing with high probability who is authoring any given tweet at @realDonaldTrump supports a content-based analysis of the tweets. As one of the most successful candidates among the populist movements that have arisen since 2016, Trump's communications in general are important. And since Twitter is Trump's main communication platform, understanding the nuances between Trump's own tweets and those of his staff can provide insight into their different types of language and interaction patterns and how those are perceived by the public. Thus, we undertake this characterization study with two major research questions:

1. RQ 1: What are the characteristics of iPhone (Staff) vs. Android (Trump) tweets, in terms of their social connections, timing, perceived appropriateness, and linguistic style attributes?
2. RQ 2: Are the social media characteristics of these tweets predictive of the source of the tweet? Can the text of a tweet identify its author?

Our findings suggest that Trump's tweets are more focused on the media, are more negative, and are perceived by citizens as less appropriate than those posted by his staff. The linguistic differences are so stark that a word vector approach can be used to classify a tweet as Trump-authored or staff-authored with $\sim$87% accuracy. We discuss the implications of this work for more deeply analyzing the Trump Twitter phenomenon and for understanding communication of populist figures like Trump.

2 Related Work

Though Trump's official political career and presidency have just begun, there is much research about how politicians – both from the U.S. and other countries – use Twitter as a platform for communication. The literature related to our case study and analysis looks at the interactivity of politicians on Twitter and the social networks of politicians on Twitter. Perhaps more relevant, however, are a few more recent pieces of scholarship that examine Donald Trump's use of the platform. This is a small but growing body of scholarship, which our research aims to contribute to.

2.1 Politicians on Twitter: Promotional Platform or Interactive Communication Tool?

Though social media is lauded as a democratizing, interactive, accessible communication tool, some scholars suggest that politicians on Twitter are not communicating

[1] http://varianceexplained.org/r/trump-tweets/.

often, or well, with their audiences and constituents. On one side of the argument is Gunn Enli and her article on how Donald Trump and Hillary Clinton used social media throughout their presidential bids. Enli states that political campaigns fail to use social media as a way to interact with voters or encourage dialogue, rather they use the platforms as channels for political marketing (Enli, p. 53). The 2016 campaigns were no different and Enli notes that campaign websites in 2016 lacked comments sections altogether, which meant the campaigns were clearly using social media platforms as "channels to promote candidates and mobilise voters, not to engage with the public" (Enli citing Pew Research 2016, p. 54).

The findings from more research, from Golbeck, et al., from 2009, found a similar trend: politicians, particularly Congresspeople in the United States, are using Twitter for information-sharing purposes rather than as a direct communication tool. All of this analysis suggests that politicians may not be using the platform to directly communicate with other users, but rather they see it as a one-way communication tool to share information and details about their event attendance.

Standing on the other side of the argument are scholars Graham et al. who examine the use of Twitter by British and Dutch politicians in general election campaigns. The authors suggest that social media (Twitter, in particular) allows for connections between ordinary people and the "popular, powerful and influential" (Graham et al. 2016, p. 766). They also call it an "interesting tool to reach out to voters" (Graham et al. 2016, p. 767). As of December 2012, 87% of democratic countries had a leader using Twitter (Graham et al. 2016, citing Digital Daya, 2012, p. 767).

Several studies have found that microblogging by politicians is used as a one-way broadcasting tool and for self-promotion, information dissemination, negative campaigning, party mobilization and impression management, rather than used for conversation and collaboration (Graham et al. 2016, p. 768). They also cite studies (Bruns and Highfield 2013; Burgess and Bruns 2012; Grant et al. 2010; Larsson and Moe 2011, 2013; Verweij 2012) that suggest politicians mainly interact with politicians, journalists and activists (p. 768). Other research (Ausserhofer and Maireder 2013) suggests that Twitter is in fact a place that can be joined by outsiders (p. 768). This is a lively debate, and is one that our case study and analysis of Trump's Twitter use hopes to contribute to.

The analysis done by Graham, Jackson and Broersma examines Twitter use by politicians (members of Parliament) in the United Kingdom and the Netherlands in 2010. Their research examined, among other things, the most common type of tweet and with whom candidates were interacting. They found that both British and Dutch politicians used @-replies frequently, 31.8% and 47.4% of tweets respectively. They also found that the politicians interacted most frequently with members of the public, 59.1% and 61.8% of tweets, respectively (Graham et al. 2016, pp. 774–775). Though this research is not based on US politicians, it lays important groundwork for further examining how politicians use of Twitter, especially when considering who their tweets are directed towards.

Other authors agree that conversations trump politics on Twitter. In 2012, Hemphill, Shapiro and Otterbacher examined how politicians in Chicago, Illinois use Twitter as a communication tool. The authors found that "politicians in Chicago are using Twitter (and potentially other social media) to engage in social conversations rather than formal politicking," (Hemphill et al. 2012, p. 3). The authors, did not look into who the politicians were communicating with and they did not analyze the communities and connections in their network analysis, but their findings related to politicians using the platform for conversations is relevant to our work.

2.2 @realDonaldTrump on Twitter

While it is important to examine how politicians use Twitter more broadly, in the United States and abroad, it is also critical to explore how other scholars have approached Donald Trump's use of Twitter – especially in his new role as politician and now, President of the United States – directly. Because Donald Trump is relatively new to the political sphere and has only been in office as president for a short time, the scholarship is limited, though this is an area ripe for research.

Enli's research, which focused on the 2016 presidential campaign cycle, looked closely at the interactions of the two candidates on Twitter. Through her analysis, she found that Donald Trump retweeted more frequently than Hillary Clinton (Enli 2017, p. 54). Approximately 25% of his tweets were retweets. He also engaged more with "ordinary users," as 78% of his retweets were from the general public (Enli 2017, p. 54). This suggests, according to Enli, that the Trump camp was more willing to engage with the general public and willing to "take the risk of retweeting content it did not control" (Enli 2017, p. 54). Enli also discusses the issue of authorship of tweets. She states that Donald Trump was more involved in social media strategy, tweeting from his @realDonaldTrump account, as though to "underline that the tweets came directly from Trump himself and were not managed and crafted solely by his campaign" (Enli 2017, p. 57).

Authors Lee and Lim examined the use of Twitter by both Donald Trump and Hillary Clinton during the campaign season in their own 2016 article. They examined 295 tweets from President Trump and 228 tweets from Clinton from two time frames in 2015 (based around dates of primary debates). They analyzed the tweets for the following traits: feminine traits, masculine traits, tweet type, tweet content, use of multimedia and civility (Lee and Lim 2016, p. 851). In examining the use of masculine and feminine language, they found that among 91 gendered trait words, Donald Trump used 38: 23 masculine (60.5%) and 15 feminine (39.5%) (Lee and Lim 2016, p. 852). As for tweet type, their analysis found that only 42.4% of Trump's tweets were original (i.e. not retweets). They also found that Trump is keen to retweet citizens. They found that almost half of Trump's tweets were retweets of constituents' accounts and over 80% of his total retweets were from citizen accounts (Lee and Lim 2016, p. 852). Eleven percent of his retweets were media organizations, almost 3% were his campaign staff and 2.5% of retweets were other public figures (Lee and Lim 2016, p. 852). As for the content of his tweets, 25% of his tweets were supportive comments and endorsements from others and another 25% of his tweets were found to be criticisms or attacks of others (Lee and Lim 2016, p. 852). The researchers also evaluated the civility of Trump's tweets and found that 10% of Trump's tweets included uncivil wording or attacks on other candidates (Lee and Lim 2016, p. 852).

Another recent article by Ahmadian et al. looked at the communication styles of Republican candidates in the U.S. presidential election, with an emphasis on Donald Trump's communication style. In order to conduct their research, they combined data from candidates' campaign speeches and their Twitter accounts (Ahmadian et al. 2017, p. 51). Though this study is not entirely focused on his use of Twitter, it gives insights into his conversation style, both on- and off-line. Most relevant to this research, the authors found that Trump's conversation style was rated highest in grandiosity—identified by the use of first person pronouns, or I-talk—when compared to other Republican candidates (Ahmadian et al. 2017, p. 51). Trump's was also rated highly informal—categorized by four variables: analytical thinking, formality, words per sentence and words with less than letters—when compared to other Republican candidates (Ahmadian et al. 2017, p. 51).

3 Data Collection

3.1 Tweet Collection and Platform Identification

Using the Twitter API, we collected all the tweets from the @realDonaldTrump Twitter account that were posted between January 1, 2016 and February 5, 2017. This time period covered his candidacy, transition, and the first few weeks of his Presidency. There were 4,590 total tweets.

Embedded in the tweet is the platform from which it was posted. Of the over 4,500 tweets, the Android and iPhone platforms were by far the most common sources. There were 291 tweets from the web client, especially leading up to the election, and 88 tweets from other platforms (ads, Instagram, etc.). However, 4,287 were from one of the two main mobile platforms.

Conventional wisdom, media reports, and comments from Trump's social media director all suggest that, during the period we are addressing, Trump was tweeting from the Android and the staff was tweeting from the iPhone. This appears to have changed – Trump may now be tweeting from an iPhone[2]. However, as we will discuss later on, classification algorithms can easily distinguish a Trump tweet from a staff tweet, so this platform distinction is not critical in research going forward.

3.2 Appropriateness Ratings

Trump's use of Twitter is unusual for someone in his position. They are often personal in nature, including personal attacks, they stray into topics not typically addressed by a candidate for President, and may appear unfiltered. The strongly partisan political environment in the US ensures that many people will have strong feelings about the politics reflected in the tweets. However, there is a different question that we found especially interesting in his case: which tweets are *appropriate*.

[2] https://twitter.com/DanScavino/status/846918912793083904.

By appropriateness, we consider whether the content is something that a citizen believes befits the role of Presidential candidate or the President of the United States, regardless of whether they agree with it politically. To determine appropriateness, we polled American workers on Mechanical Turk.

For this part of our analysis, we eliminated retweets and quoted tweets so we only considered posts written by someone using the account. This gave us 3,666 total tweets.

Next, we asked Americans to rate the tweets. We created a pool of about 175 people on Mechanical Turk. We restricted the HIT to workers based in the US, listed a requirement that workers be American citizens, and everyone in our pool passed a qualification test that posed several American cultural questions. We also had workers describe their political beliefs on a liberal-conservative scale. Each of the 3,666 tweets was rated by 10 self-described liberals and 10 self-described conservatives, randomly selected from the pool. Raters were asked if the tweet was "Appropriate", "Neutral", or "Inappropriate".

Instructions made it explicit that workers should not to rate a tweet based on whether they agreed *politically* with the statement but rather to rate if it was an appropriate statement that a Presidential candidate or the President should be making. We also included an attention check in the instructions:

> Please rate whether you personally believe these are appropriate tweets for the President of the United States or a candidate for President to be posting.
>
> This does not necessarily mean you agree with the politics of the tweet. For example, a tweet that says "Today I signed a bill that cuts taxes by 1% in all income brackets." may not reflect what you personally believe is good policy, but it is an appropriate tweet for the President to make. You would mark a tweet like that as Appropriate. To make sure you read instructions, always mark tweet 3 as neutral. However, a tweet that said "My hands are HUGE and anyone who says otherwise is spreading FAKE NEWS!" may be marked either way depending on whether you think this is something the President should be saying on Twitter.

We assigned perceived "appropriate" tweets a score of 1, "inappropriate" a score of −1, and "neutral" a score of 0. We averaged the rating of all 20 raters to get a score for each tweet. An average of 0.9 would mean 19 people labeled the tweet "appropriate" (1) and 1 person labeled it "inappropriate" (−1). An average score of −1 would mean all 20 people labeled the tweet "inappropriate".

3.3 Mention and Retweet Coding

President Trump reaches a wide audience with his 140-characters, with over 32.1M followers on the platform. And although he keeps his following list tight-knit (following just 45 accounts) he interacts quite frequently on the platform with other users. And rather than looking at the language used in the interactions, we can look at the type of accounts the tweets mention to understand behavioral differences between the platforms (Table 1).

From our original dataset, we extracted 550 retweets and 774 tweets that contained one or more account mentions. Some tweets mentioned multiple accounts, yielding 874 mentioned accounts in total. Using an open coding approach, we established a set of

Table 1. Codes for account types in Trump mentions

Category	Code	Details
Family	fam	Twitter handles of people related to Trump (including children and wife)
Media	med	Twitter handles of news media organizations and journalists
Politicians	pol	Twitter handles of politicians, including VP, Senators, House Reps and local politicians (governors, mayors, etc.)
Constituents/citizens	cit	Twitter handles of average Twitter users, including constituents
Official businesses	bus	Twitter handles of non-government businesses and companies
Government entities	gov	Twitter handles of government organizations, departments, branches
Celebrity	cel	Twitter handles of celebrities, well-known individuals
Trump admin	adm	Trump administration/campaign handles
Union	uni	Twitter handles for local unions

categories in order to better understand the populations within Trump's retweet and mentions network. The categories included: family, media, politicians, citizens, businesses, government entities, celebrities and members of the Trump administration.

4 RQ1: Content Characterization

4.1 Mentions and Retweets

Nearly all of the retweets on the @realDonaldTrump account came from the Android device. Of the 550 retweets, only 11 were from iPhone, including 3 administration members, 3 media retweets, and 2 citizens.

The Android retweets were dominated by retweets of citizens: 429 of the 536 (81.2%). Another 10% were retweets of media. Some of the retweeted accounts were no longer accessible: 13 were suspended and 19 were completely deleted.

Mentions follow a much different pattern. From Android, the vast majority of Trump's mentions are of media (81%) with another 10% addressing politicians. iPhone-based tweets have a more diverse set of mentioned account types, as shown in Fig. 1.

4.2 Appropriateness

Using the ratings from our Mechanical Turk workers, we can analyze the appropriateness of the @realDonaldTrump tweets. Recall 0 is neutral, 1 is appropriate, and −1 is inappropriate. The average tweet scored a 0.51, halfway between neutral and appropriate. 780 (23%) of tweets had an average negative score, putting them on the inappropriate end of the scale.

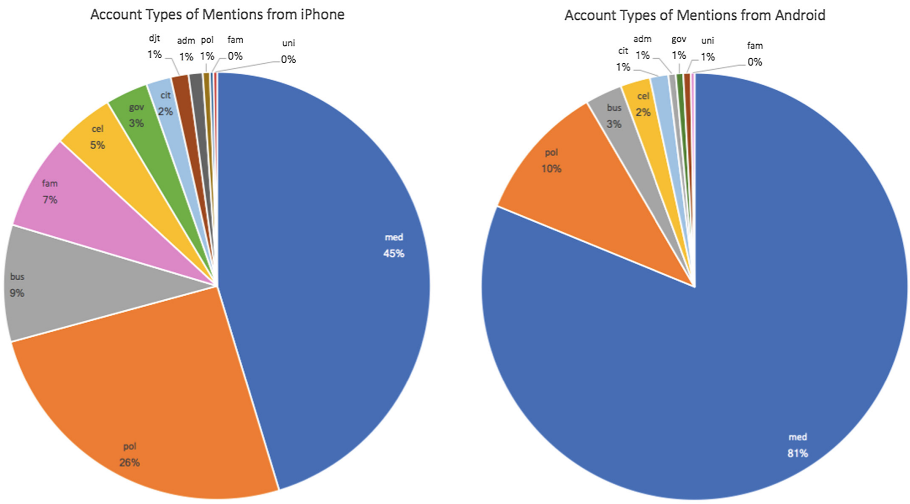

Fig. 1. Distribution of mention types from iPhone (left) and Android (right)

Not surprisingly, liberals and conservatives differ in what they find appropriate. The average rating from liberals is 0.38 while conservatives gave an average of 0.64. Liberals gave 1,064 (29%) tweets an average negative score, indicating they felt a sizable minority of the tweets were not befitting someone in Trump's position.

Breaking down the tweets by device also shows large differences in appropriateness ratings. Android-originating tweets were rated significantly lower than iPhone tweets. Among the 780 "inappropriate" tweets with average scores less than 0, 482 (61.9%) were posted from Android. Further dividing the ratings according to the rater's political leanings, we find dramatic results. Liberals give Trump's Android tweets an average score of only 0.1. Nearly half (49%) of Trump's tweets averaged neutral to inappropriate scores when rated by liberals. On the other hand, conservatives give Trump much higher marks and the staff tweets have quite a high rating. These results are shown in Fig. 2.

There are many tweets (1,297) that all of our raters agreed were neutral to appropriate (i.e. no one said they were "inappropriate"); 647 tweets were ranked totally appropriate by all 20 raters. Most of these came from iPhones, but some (118 of the 647, or 18%) came from Trump's Android device. These were typical campaign tweets for the most part, e.g. "It was so great being in Nebraska last week Today is the big day —get out and vote!" (Fig. 3).

However, there were some tweets that were considered inappropriate by liberals and conservatives alike. Trump's tweet threatening to "spill the beans" on Ted Cruz's wife received the lowest score; it was ranked inappropriate by everyone except one conservative (an average score of −0.9).

Some additional tweets received very low average ratings from all sides. The following tweets were the lowest rated, with overall average scores of −0.75 or lower:

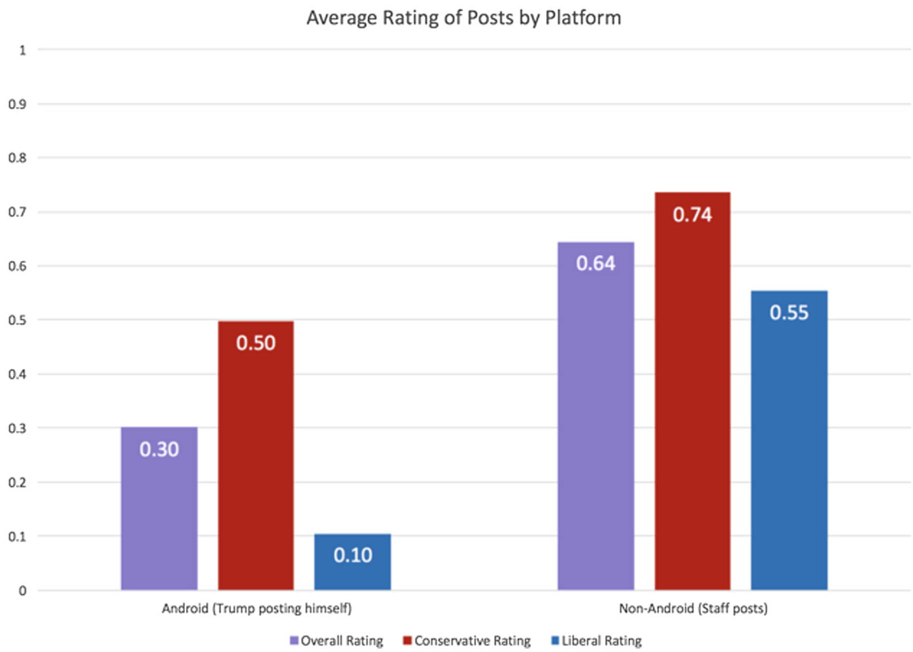

Fig. 2. Appropriateness ratings by platform and rater's political leanings

Fig. 3. Trump's most inappropriate tweet according to our raters with an average score of −0.9. The tweet originates from the Android platform, indicating Trump himself authored it.

> *Wall Street paid for ad is a fraud, just like Crooked Hillary! Their main line had nothing to do with women, and they knew it. Apologize?* (−0.85)

> *Explain how the women on The View, which is a total disaster since the great Barbara Walters left, ever got their jobs. @abc is wasting time* (−0.85)

HILLARY FAILED ALL OVER THE WORLD. #BigLeagueTruth ❌LIBYA ❌SYRIA ❌IRAN ❌IRAQ ❌ASIA PIVOT ❌RUSSIAN RESET ❌BENGHAZI… https://t.co/H1UH0svtt2 (-0.75)

.@FoxNews is so biased it is disgusting. They do not want Trump to win. All negative! (−0.75)

There are 53 tweets that liberals unanimously deemed inappropriate. These include the tweets mentioned below, criticisms of Megyn Kelly, John McCain, Elizabeth Warren, and Mitt Romney, and comments on pop culture. Below are four representative tweets from this pool:

Pocahontas is at it again! Goofy Elizabeth Warren, one of the least productive U.S. Senators, has a nasty mouth. Hope she is V.P. choice.

Really dumb @CheriJacobus. Begged my people for a job. Turned her down twice and she went hostile. Major loser, zero credibility!

.@NBCNews is bad but Saturday Night Live is the worst of NBC. Not funny, cast is terrible, always a complete hit job. Really bad television!

Mitt Romney, who was one of the dumbest and worst candidates in the history of Republican politics, is now pushing me on tax returns. Dope!

There were also some tweets where liberals and conservatives had major disagreements, where the majority of liberals said they were inappropriate and the majority of conservatives said they were appropriate. The tweets below all had a difference of 1.5 between the average liberal rating and average conservative rating, except for the first which had a difference of 1.6.

Taxpayers are paying a fortune for the use of Air Force One on the campaign trail by President Obama and Crooked Hillary. A total disgrace!

Goofy Senator Elizabeth Warren @elizabethforma has done less in the U.S. Senate than practically any other senator. All talk, no action!

Watching John Kasich being interviewed - acting so innocent and like such a nice guy. Remember him in second debate, until I put him down.

What Barbara Res does not say is that she would call my company endlessly, and for years, trying to come back. I said no.

People forget, it was Club for Growth that asked me for $1 million. I said no & they went negative. Extortion! https://t.co/oq8jmoep7i

Hillary Clinton has been involved in corruption for most of her professional life!

#CrookedHillary is outspending me by a combined 31 to 1 in Florida, Ohio, & Pennsylvania. I haven't started yet! https://t.co/BcoPrwqFMe

These tweets are particularly notable for those interested in the political divisions in the United States, as they reflect not just a difference in politics but a stark difference in how people believe political discourse should take place.

4.3 Linguistic Style

As part of RQ1, we hope to characterize linguistic differences between the Staff and Trump tweets as well as differences in the appropriate and inappropriate tweets. We are focused on two categories of linguistic differences: pronoun usage and affective words. Pronouns are likely to capture differences in who the tweets are talking about. With a focus on affective processes, we anticipate differences in positive vs. negative emotional words and language relating to anger and anxiety.

We removed mentions and retweets from this analysis, using 3,287 total tweets: 1,842 iPhone tweets and 1,445 Android tweets. We used LIWC2007 (Pennebaker et al. 2001) to analyze the text.

When comparing between the devices, we found significant differences in all the hypothesized areas. Statistics are calculated with Student's t-tests with a Bonferroni correction and are significant for $p < 0.01$.

Android (i.e. Trump) tweets used "I", "he/she", and "they" significantly more often than iPhone (i.e. Staff) tweets. iPhone tweets, on the other hand, used "you" significantly more than Android tweets, at 5.98 times the rate. These indicate that Trump uses his tweets to talk about himself and other people while the staffers address the audience.

There was no significant difference in the overall number of affective words used based on device, but there were large differences in the types of affective words. iPhone tweets used significantly more positive emotion words and Android tweets used significantly more negative emotion words, anxiety words, and anger words.

These differences were mimicked in the analysis of appropriate vs. inappropriate tweets. We separated tweets into Fully Appropriate (received scores of appropriate from all raters—647 tweets, 82% of which were non-Android) and Inappropriate (had an average negative score—780 tweets, 62% of which were Android posts). Since the prevalence of Android vs. iPhone tweets is related to the appropriateness ratings, we expect some overlap in findings. That did occur, but some results were more extreme than what we found in the platform analysis. All results reported are significant for $p < 0.01$ with the same tests and corrections as described above.

Inappropriate tweets use "he/she" over 16 times more often than appropriate tweets. Appropriate tweets use "we" 3.3 times as often than inappropriate tweets; there was no significant difference in the use of "we" based on platform. Appropriate tweets also used "you" more than 17 times as often as inappropriate tweets.

Emotional content also differs in that inappropriate tweets are more negative. The inappropriate tweets used words related to negative emotions more than 8 times as often as the appropriate tweets; words associated with anger 6.6 times as frequently; and words associated with anxiety 7.4 times as frequently. Appropriate tweets use words related to positive emotions at 3.3 times the rate of the inappropriate tweets.

5 RQ2: Prediction

Our content analysis echoes discussions in the media about who is authoring tweets on the @realDonaldTrump account. Trump representatives appear to have confirmed that, until recently, the platform was a fairly accurate distinguishing source but this appears to no longer be the case with all recent tweets to the account coming from an iPhone. Thus, being able to automatically connect a tweet to its author becomes more important as a tool for Trump Twitter Analysts.

We use the words of the tweets themselves as features to predict authorship. Each tweet is represented as a word vector and it is labeled with its source (Android or iPhone) as a class. We used standard 10-fold cross validation for evaluation. We trained the Weka SimpleLogistic classifier, the best performing of five different types of classifiers we tested. We report precision, recall, F-measures, ROC AUC, and accuracy.

Results are shown in Table 2. We are able to predict the source, and thus the authorship, of tweets in our dataset with 86.7% accuracy and a ROC AUC of 0.924.

Table 2. Results for classification with SimpleLogistic algorithm for classifying tweets

TP Rate	FP Rate	Prec.	Recall	F-Meas	MCC	ROC area	PRC	Class
0.927	0.193	0.827	0.927	0.874	0.74	0.924	0.911	Android
0.807	0.073	0.918	0.807	0.859	0.74	0.924	0.935	iPhone
0.867	0.133	0.872	0.867	0.867	0.74	0.924	0.923	Weighted average

Given that device is connected to authorship for the tweets in our dataset, this classifier will allow prediction of authorship on new tweets just with the language of the tweet, regardless of what device is used.

We achieve similar results using the Word Vector approach with the SimpleLogistic algorithm to classify tweets as “Appropriate” or “Inappropriate”. Tweets were classified as Inappropriate if their average was < 0 and “Appropriate” otherwise. The classifier achieved and 87.3% accuracy rate with ROC AUC of 0.917.

6 Discussion

6.1 Identifying Tweet Authorship and Impact

Our predictive model shows that tweet authorship can be accurately predicted with a word vector. This means that even as indicators like platform change, the language of the tweet itself can identify the author. Thus, going forward, such models will allow differentiation of content for analysis.

We were also able to classify perceived appropriateness with high accuracy. While there is some overlap in authorship and appropriateness, it is only partial. Being able to predict bi-partisan disapproval of tweet content ahead of time can be especially useful for communications professionals. It is possible that Trump and even members of his team are unconcerned with citizens’ perception of what is appropriate. However, there are certainly candidates and legislators who are concerned with this, and the success of the classifier suggests promise in future work that can provide algorithmic feedback about social media posts before they are made public.

6.2 The Language of Populism

This characterization study was not designed to describe the entire populist movement. However, as an outspoken and successful populist, Trump serves as an interesting case study. His staff's tweets from the period we studied are often more traditional in their online communication approach when compared with Trump's own comments. Our results that show a focus on himself and others – both through his linguistic patterns and heavy use of media mentions – as well as tweets rooted in negative emotions, anger, and anxiety.

These results suggest an approach to analyzing the language of populists on Twitter may yield interesting linguistic insights about the movement itself and how it differentiates itself from mainstream political communication online.

7 Conclusions

Trump is changing what we expect in public statements from our leaders. In this paper, we leveraged the observation that the platform from which a tweet was posted to @realDonaldTrump can identify its author as Trump or Staff. Our results show that Trump himself tends to more actively tweet about the media, speak about himself or others (rather than the audience), and to communicate about anger, anxiety, and negative emotions when compared with staff tweets. We also found that citizens' perceive Trump's tweets as less appropriate for someone in his position to be posting. Using language features, we were able to develop predictive models for tweet authorship and for tweet appropriateness that perform with high accuracy. Our findings have implications for the study of Trump specifically and for understanding the online language and behaviors of the populist movement.

References

Ahmadian, S., Azarshahi, S., Paulhus, D.L.: Explaining Donald Trump via communication style: grandiosity, informality, and dynamism. Pers. Individ. Differ. **107**, 49–53 (2017). doi:10.1016/j.paid.2016.11.018

Enli, G.: Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US presidential election. Eur. J. Commun. **32**(1), 50–61 (2017). doi:10.1177/0267323116682802

Golbeck, J., et al.: Twitter use by the U.S. congress. J. Am. Soc. Inf. Sci. Technol. **61**(8), 1612–1621 (2010). doi:10.1002/asi.21344

Graham, T., Jackson, D., Broersma, M.: New platform, old habits? Candidates' use of Twitter during the 2010 British and Dutch general election campaigns. N. Media Soc. **18**(5), 765–783 (2016)

Hemphill, L., Shapiro, M.A., Otterbacher, J.: Chicago Politicians on Twitter (2012)

Hsu, C.L., Park, H.W.: Mapping online social networks of Korean politicians. Gov. Inf. Q. **29**(2), 169–181 (2012)

Lee, J., Lim, Y.: Gendered campaign tweets: the cases of Hillary Clinton and Donald Trump. Publ. Relat. Rev. **42**(5), 849–855 (2016). doi:10.1016/j.pubrev.2016.07.004

Ott, B.L.: The age of Twitter: Donald J. Trump and the politics of debasement. Crit. Stud. Media Commun. **34**(1), 59–68 (2017)

Pennebaker, J.W., Francis, M.E., Booth, R.J.: Linguistic Inquiry and Word Count (LIWC): A Computerized Text Analysis Program, vol. 7. Lawrence Erlbaum, Mahwah (2001)