

Trollspeak: who do Russian trolls tweet like?

Martin Eberl (Ludwig-Maximilians-Universität München)

martin.eberl@lmu.de

Abstract

The present study is concerned with the link between political affiliation and linguistic variation on lexical as well as sublexical levels on Twitter, with a focus on tweets regarding the 2016 US elections. Linguistic variation depends on a variety of factors and reasons, including social identities. As previous research has shown, political affiliation and associated sociolinguistic variation can be both reflected in speech and in writing. Variables in the latter may even include sublexical features such as punctuation. This study examines two datasets of tweets, one from supporters of the presidential candidates Hillary Clinton and Donald Trump, and the other from Russian operatives engaging in a misinformation campaign, as to their lexical variation in the text, variation in the use of emojis and hashtags as potentially metatextual¹ features and variation in the use of punctuation as a sublexical feature. The results show that the Russian tweets have their own, distinct features, while tweets from Trump and Clinton supporters are remarkably alike.

Keywords: *sociolinguistics, social media, twitter, variation, metatextual features*

1 Introductory remarks

Social media has become increasingly political since its inception. Twitter, for example, "has emerged as a key platform on which anyone with a smartphone can engage in political discourse" (Nguyen 2017). This increasing political aspect of Twitter and other platforms involves not only the offices of political parties and individual candidates running specific social media campaigns, but also discussions between normal users of the platform sharing their political opinions. As in other spheres, they seem to seek out those that share their opinions, leading to a polarization of the userbase (Conover 2013). But does this polarization show itself only in who the users talk to and what or whom they talk about? Or does it also reveal itself in how we talk, i.e. in sociolinguistic patterns?

In Labov's (2010) framework, for instance, political affiliation may reveal itself in sociolinguistic patterns, or, conversely, it may be presupposed by the hearer due to the speech of the interlocutor. However, such variation is not limited to speech, but may also show itself in written discourse. Schnoebelen (2012) has shown that stylistic variation may occur even in metatextual features such as Twitter emoticons. In a similar vein, Tatman and Paullada (2017) have shown, for instance, that such variation extends even to sublexical features such as punctuation. In their data, supporters of the Black Lives Matter movement tend to use less punctuation with relatively uniform distribution across tweets, whereas supporters of the Blue Lives Matter movement use more punctuation, which is mostly centered around the beginning and end of tweets. In addition, they found distinguishing lexical features such as the 'fist' emoji being used exclusively by Black Lives Matter tweeters, while the 'white star' emoji was used solely by those in favour of

¹ For a debate on whether emojis and hashtags are textual or metatextual features, see, among others, Zappavigna 2015 and Rambukanna (2015). I follow the latter in arguing that they should be treated "as both text and metatext simultaneously" (2015, 161). As such, even when hashtags are included in running text rather than attached separately, they serve functions beyond the utterance itself, such as topic marking, the reference to and activation of interpersonal relationships and structuring the text (Zappavigna 2015, 274). Therefore, hashtags, as used on Twitter and other social media platforms, cannot be simply textual.

Blue Lives Matter. Further research (Golbeck et al. (2011), Hu et al. (2013)) has shown that personality traits may be predicted from Twitter data and that language on Twitter shows distinct features when compared to other computer-mediated communication. Taken together, these findings indicate that sociolinguistic studies of Twitter data, as far as it is available, are a promising endeavour.

As luck would have it, a large trove of Twitter data from a specific politically motivated group was made available recently. On February 14th, 2018, NBC news published a database of 203,451 tweets that Twitter claimed to have engaged in 'malicious activity' during the 2016 presidential election. Said malicious activity was the undue influence of the presidential election. The users behind these tweets, then, comprise one political group with a specific agenda and should show similar sociolinguistic patterns.

I examine the tweet content for patterns of variations on three distinct levels: lexical variation in the text, variation in the use of emojis and hashtags as metatextual features and variation in the use of punctuation as a sublexical feature. The same analysis is then carried out on a set of tweets by users who have been identified as politically conservative and politically liberal from roughly the same time period. Taken together, these three groups should be identifiable through linguistic patterns.

In analysing these three datasets, the present study therefore seek to answer this central question: is political affiliation a strong enough factor to influence sociolinguistic variables on social media platforms such as Twitter?

2 The data

The first dataset comprises the tweets from so-called "Russian trolls", i.e. those tweets removed by Twitter on suspicion of having been written and disseminated by Russian operatives trying to sow discord before, during and after the election of 2016. This dataset was collected by the American news company NBC and made publicly available for research purposes. NBC describes the Russian operation as thus:

[T]hese accounts, working in concert as part of large networks, pushed hundreds of thousands of inflammatory tweets, from fictitious tales of Democrats practicing witchcraft to hardline posts from users masquerading as Black Lives Matter activists. Investigators have traced the accounts to a Kremlin-linked propaganda outfit founded in 2013 known as the Internet Research Agency (IRA). (Popken, 2018)

After Twitter had deleted the tweets in question, along with the accounts who had written them, "three sources familiar with Twitter's data systems cross-referenced the list of names released by Congress, excluding any account that Twitter later restored, to create a partial database of tweets that could be recovered from the suspended accounts". The restored dataset comprises 203,541 tweets from a total of 453 different accounts. For the purposes of this study, however any retweets were excluded, reducing the number to 55,889 analysable tweets. The tweets are dated from June 14th, 2014 to July 26th, 2017.

The second dataset has been sampled directly from Twitter using the public Twitter API and the *rtweet* package for R by Michael Kearney (2017). The goal was to identify political supporters of either Donald Trump or Hillary Clinton who had tweeted in favour of their candidate during the election process. This proved to be a significant problem, given that they could not be identified automatically, for instance through their use of hashtags, so as not to affect the later results of the linguistic and metalinguistic analysis. Thus, the individual

accounts had to be identified manually, before their Twitter timeline was sampled through the *rtweet* package. The criteria included their vocal support for one candidate, having tweeted multiple times in favour of their candidate, and activity throughout the election process. In addition, members of any of the various political campaigns as well as politicians and journalists were deliberately excluded from the dataset, as were accounts used by more than one person. This manual selection process identified 536 total accounts of interest. 76,878 tweets were sampled from these accounts. Excluding retweets, this left 39,152 tweets for analysis. Of these, 20,775 were made by Clinton supporters and 18,377 were made by Trump supporters. The maximum amount of tweets coming from a single user was set at 160 tweets, so as not to give a few users the ability to dominate the dataset. The median amount of tweets were 73 for the Trump supporters and 76 for Clinton supporters. The tweets were written between September 1st, 2010 and July 27th, 2017.

The figure below shows some exemplary tweets from the first datasets. Note that the tweeters' usernames, their Twitter handles and the time of the tweets have been withheld for these examples due to privacy concerns.²

Once gathered, the individual tweets were tokenized for the following analysis by using the *stringr* package for R (Wickham 2019). An additional pre-processing step was not necessary due to the output format of *rtweet*. Tweets were tokenized into individual words for analysis of lexical and metatextual features first, with a second tokenization step into individual characters to later determine the position of sublexical features.

Using the *str_split* feature of the *stringr* package allowed for exact tokenization on the basis of new lines and spaces as well as punctuation. In rare cases, it failed to properly extract words when written without spaces (either due to space issues or due to typos). The additional *str_extract_all* function was used to extract all strings beginning with a #, which allowed for the listing of hashtags. Emojis had to be identified using different methods for the Russian dataset when compared to the Trump and Clinton datasets. The former included encoding of emojis in 4 byte strings (starting with D), whereas *rtweet* gave the unicode for emojis (starting with <U+, e.g. <U+0001F602>). Given that they were encoded consistently throughout both datasets themselves, however, identifying them could also be achieved through the *str_extract_all* function.

² I am aware that there is a debate on how best to handle reproducing tweets in a scholarly context. Some researchers prefer giving the full account information for copyright reasons, while others prefer removing any information through which the authors can be identified due to privacy concerns, going so far as to replace lexical items in the tweets themselves. In my opinion, privacy concerns should take precedent as far as possible, but the linguistic examples themselves should not be altered. I have therefore opted for the middle ground to preserve the linguistic structure of the tweets themselves while making the authors not immediately identifiable.



Figure 1: Exemplary Tweets

3 Lexical variation

The first factor investigated in regard to the lexical variation of the three sets of tweets was their type-token-ratio. As the table below shows, there is very little difference between the ratios of the individual sets, with Russian tweets showing a very slightly lower type-token-ratio overall than that of Trump supporters, who in turn show a very slightly lower type-token-ratio than Clinton supporters.³ In a sample of the Russian dataset roughly matching the size of the other two datasets, the TTR was 0.11 as given in the parentheses.

Table 1: Type-token ratios of all three datasets

Russian	Pro-C	Pro-T
0.09 (*0.11)	0.11	0.10

The uniformity of the type-token-ratio may suggest that this datapoint is varied less according to the individual users or their political stances, but is more dependent on the limitations of Twitter as a messaging platform, with the text being capped at 140 characters. Previous studies on TTR have shown the length of the text to affect the average type-token ratio (Caruso et al. 2014, 139), so it would seem logical to assume that tweets show similar type-token ratios. The tweet length of the various datasets is summed up in figure 2 below. It shows a density plot of all three datasets across the length of tweets in characters, meaning that higher values of density show a higher number of tweets with that particular length.⁴

³ TTR is calculated for each dataset as a whole rather than individual tweets.

⁴ Note that due to technical issues with the readout of tweets (such as encoded emojis accounting for more characters than their graphical equivalents), tweets in the datasets can be up to 150 characters. Since this affects all three datasets equally, the ratios between the datasets remain unaffected.

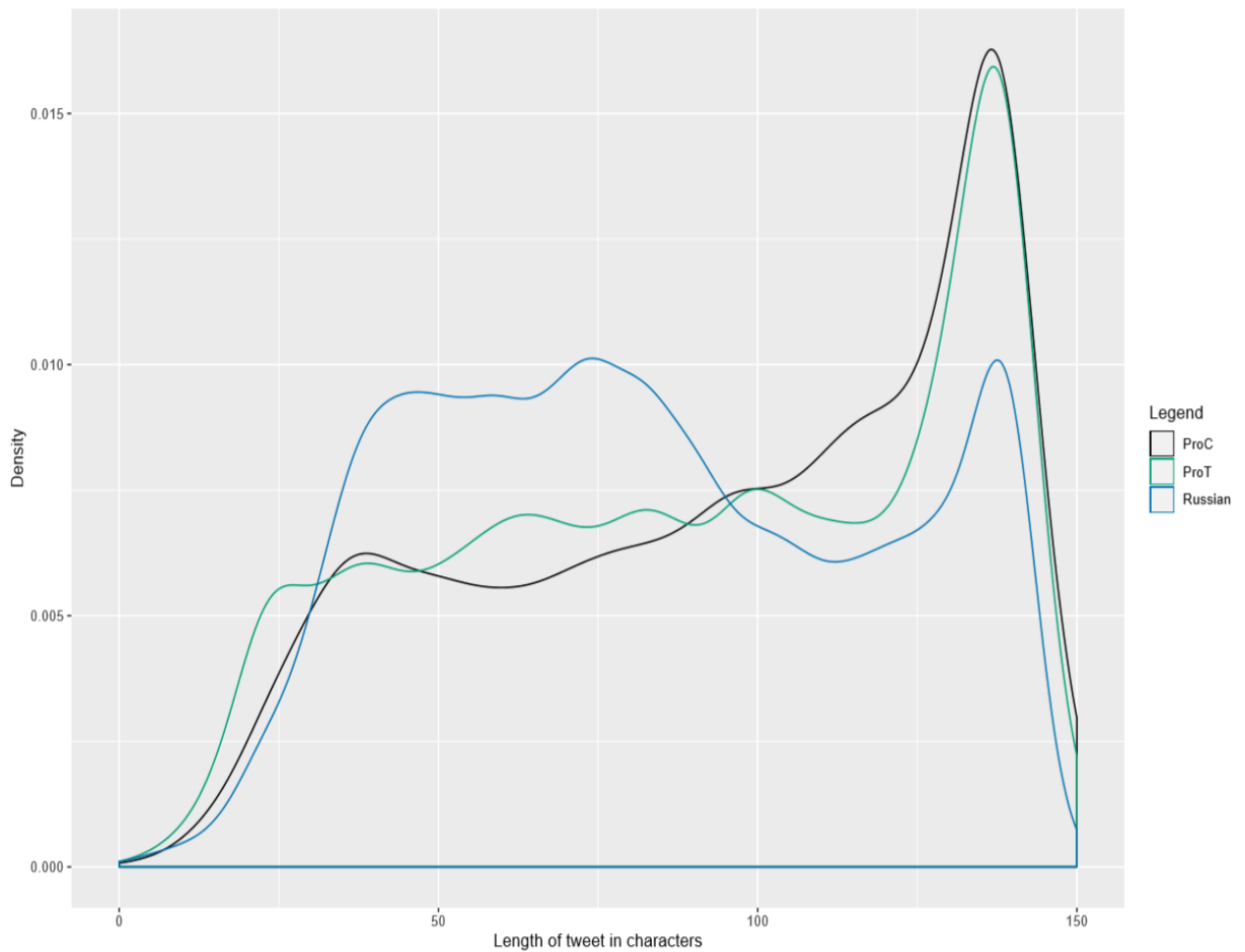


Figure 2: Length of the tweets in characters

As is evident, the character length of Trump and Clinton supporters is very similar, with a peak towards the maximum tweet length of 140 characters, which suggests that the users adjusted an originally longer message to the platform’s character limit. The Russian tweeters seem to have hit this limit far less often, and show a similar peak towards shorter messages between 35 and 75 characters. This may suggest that their tweets are less spur-of-the-moment and may have been pre-planned. It also explains their lower type-token ratio as explained above, however.

A second variable to be explored were the most frequently used nouns and adjectives in the three datasets. These were identified through a manual check of the word frequency lists generated for each dataset. The six most common ones for each are shown in the table below, with their absolute frequency:

Table 2: Frequent Nouns

Russian	Pro-C	Pro-T	Russian	Pro-C	Pro-T
trump (9,399)	hillary (1,171)	trump (4,552)	black (889)	good (457)	new (842)
clinton (4,345)	trump (928)	people (1,064)	new (889)	great (437)	great (620)
politics (3,709)	day (680)	hillary (738)	american (734)	new (429)	good (396)
hillary (3,561)	time (587)	don (680)	good (682)	happy (269)	happy (292)
obama (2,540)	president (516)	trumtrain (653)	great (586)	real (227)	best (224)
midnight (2,342)	people (472)	potus (617)	white (575)	proud (175)	free (213)

Table 3: Frequent Adjectives

Quite unsurprisingly, the candidate’s names are among the most frequent nouns, with the candidate the tweeter identified with leading the count and the opposing candidate closely behind. Russian tweeters also mentioned the sitting president, Barack Obama, frequently. *Time*, in the tweets of Pro-Clinton tweeters, seemed to mostly refer to the time for a female president having arrived. In the end, however, the frequent nouns offer little insight into sociolinguistic tendencies.

The frequent adjectives are slightly more interesting. Note that they are exclusively positive for both Pro-Clinton and Pro-Trump tweets, and very similar among these two groups. One might have thought for *great* to be more favoured among Pro-Trump tweeters, but the Pro-Clinton tweeters seem to have countered the idea of ‘making America great again’ with the idea that ‘America is already great’. Russian tweeters, aside from the same positive adjectives *new*, *good* and *great* that the other groups use, quite ironically make frequent use of *american* as well. This is usually connected with ‘American values’ that the voter should keep in mind. Also note that they frequently use both *black* and *white*, which may reflect an effort to stoke racial tensions ahead of the election. Once again, the differences are more pronounced between the Russian tweets and the other two sets than they are between Clinton and Trump supporters.

4 Use of hashtags and emojis

Aside from the lexical variation, I also examined the variation in metatextual features such as hashtags and emojis. Their patterns reflected some of the tendencies already observed above. The following table shows the most frequent hashtags among the three datasets:

Table 4: Most frequent hashtags among the three datasets

Russian	Pro-Clinton	Pro-Trump
#politics	#ImWithHer	#Trump2016
#news	#StrongerTogether	#MAGA
#Merkelmussbleiben	#quote	#MakeAmericaGreatAgain
#TrumpForPresident	#DemConvention	#TrumpTrain
#IslamKills	#Hillary2016	#quote
#PJNET	#GOPdebate	#Trump
#tcot	#mapoli	#POTUS
#Brussels	#Election2016	#NeverHillary
#StopIslam	#HillYes	#maga

As can be gleaned from the table, variation between highly frequent tokens of Clinton and Trump supporters is, as has been the case with the nouns above, mostly limited to the candidates’ names and their slogans. While Pro-Trump tweeters prefer #Trump2016 and #MAGA, Pro-Clinton tweeters opt for her slogans of #ImWithHer and #StrongerTogether. It may be surprising to some observers that only one of the most frequently used hashtags is a negative hashtag oriented at the political opponent, namely #NeverHillary, which is, of course, an adaptation of the #NeverTrump slogan that Republicans used to show their party-internal opposition to nominating Donald trump as their candidate. Pro-Clinton supporters also seemed to be more interested in electoral events such as the Democratic national convention

(#Dem-Convention) and the Republican debates (#GOPdebate). The variation in Russian lexical items and hashtags hints at a broader array of topics, including highly controversial debates and islamophobic concepts. Also note that it is not limited to discussions surrounding the U.S. elections, but also includes other international political hashtags such as the German #Merkelmussbleiben. In addition, Russian users also use more hashtags overall than the other two groups: whereas Clinton supporters use 0.51 hashtags per tweet and Trump supporters use 0.64, the Russian tweets contained 0.86 hashtags per tweet on average⁵. Once more, we find that Russian tweeters seem to be concerned with a variety of political topics, whereas Pro-Clinton and Pro-Trump users show little variation.

Emojis vary a bit more between the datasets. In general, Russian emoji use is much lower, clocking in at only 0.06 emojis per tweet, whereas Pro-Clinton users employed 0.29 emojis per tweet and pro-Trump users employed 0.33 emojis per tweet⁶.

Table 5: Emojis per tweet

Russian	Pro-C	Pro-T
0.06	0.29	0.33

In qualitative terms, there are also significant differences between the emojis pro-candidate tweeters used versus those used by the Russian tweeters. The ever-present 'tears of joy'-emoji, for instance, was the most common emoji among both pro-Trump and pro-Clinton tweeters. Russian tweeters, however, have a stronger preference for the 'sad' emoji. The table below sums up the six most common emojis of the various datasets:

Table 6: Frequency of emojis

Russian	Pro-C	Pro-T
sad (269)	tears of joy (399)	tears of joy (575)
vehicle (108)	black heart (192)	dash symbol (369)
chick (79)	fire (162)	steam loc (173)
flag (76)	clapping (132)	right arrow (160)
pointdown (65)	cryface (115)	rightward hand (123)
elephant (55)	camera (91)	fire (111)

In addition to the difference between 'tears of joy' and 'sad', there is almost no overlap between the frequent emojis. The 'steam loc' emoji, for instance, is only used by Trump supporters in parallel to the *trumptrain* being one of the most common nouns. On the other hand, only Pro-Clinton users seemed to have been on 'fire' and 'clapping' frequently. It is here, then, that we may observe the sociolinguistic patterns that differ between the latter two groups. There is also a possible argument to be made about the 'sad' emoji being used, once more, to emphasize

⁵ The total numbers are 10,567 hashtags for Pro-Clinton tweeters, 11,773 for Pro-Trump tweeters and 48,216 for the Russian tweeters.

⁶ The total numbers of emojis identified in the datasets were 1,836 for the Russian dataset, 6,866 for the Pro-Trump tweeters and 5,405 emojis in the Pro-Clinton dataset.

negative emotions and connotations with the state of the Union, as it were, to further sow discord among the electorate.

5 Sublexical variation

As a final variable in this study, the use and position of both punctuation marks and hashtags was examined. The first category included exclamation marks, question marks, quotation marks, periods and commas across the length of the tweets⁷. The three plots in figure 3 below show the general punctuation patterns of these signs across the three datasets:

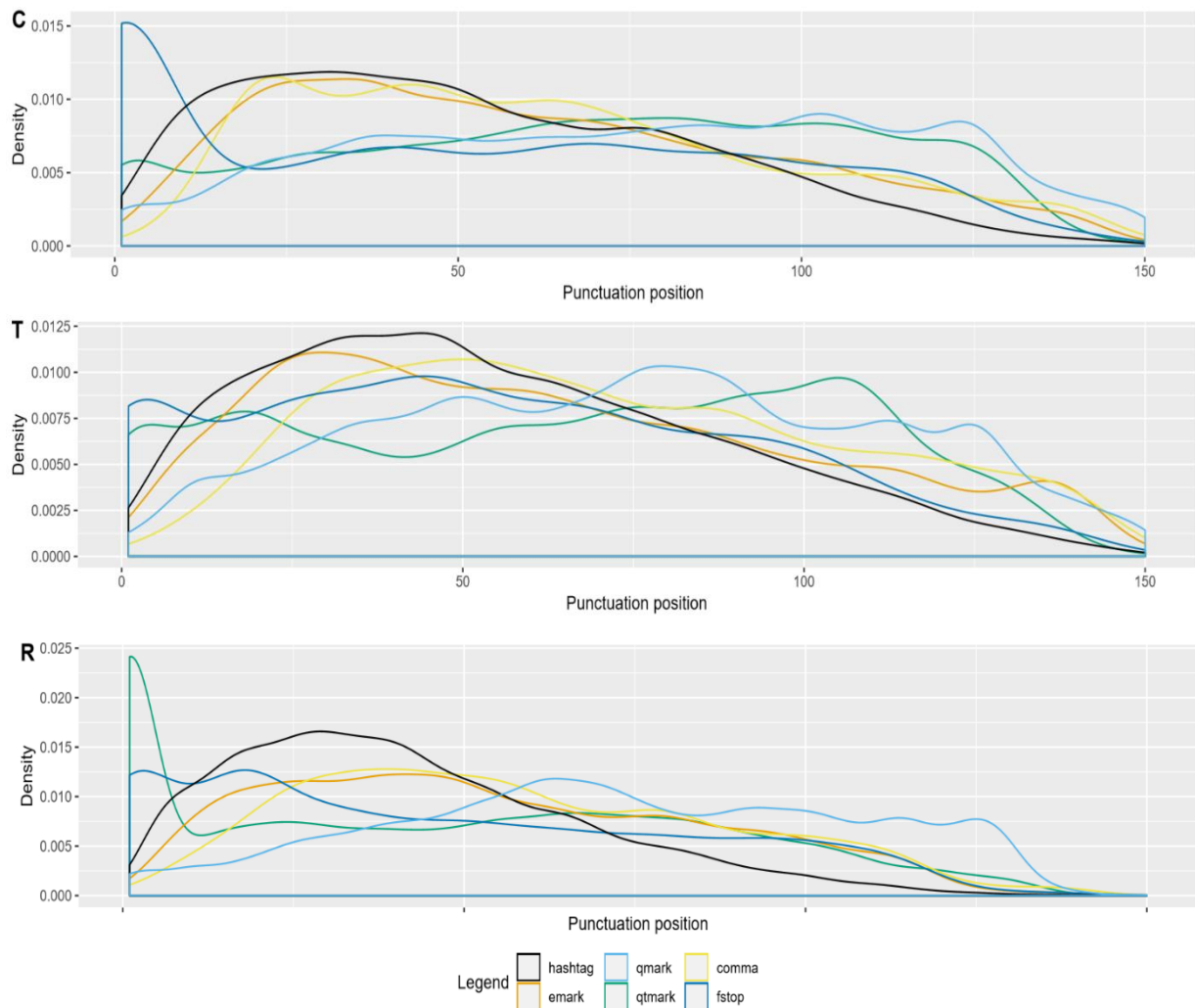


Figure 3: Position of various punctuation marks across tweets

⁷ Due to technical processes in the sampling of the tweets through Twitter's API and the way that the *rtweet* package works, the maximum length of the tweets in the graphs below is 150 characters instead of 140 characters. This does not, however, affect the general patterns observed.

From this data, it seems that the general position of punctuation marks and hashtags is similarly spread out across pro-Trump and pro-Clinton tweeters as well as the Russian tweeters. Figure 4 below focuses more closely on the position of individual signs as they appear in Russian tweets, pro-Clinton tweets and pro-Trump tweets.

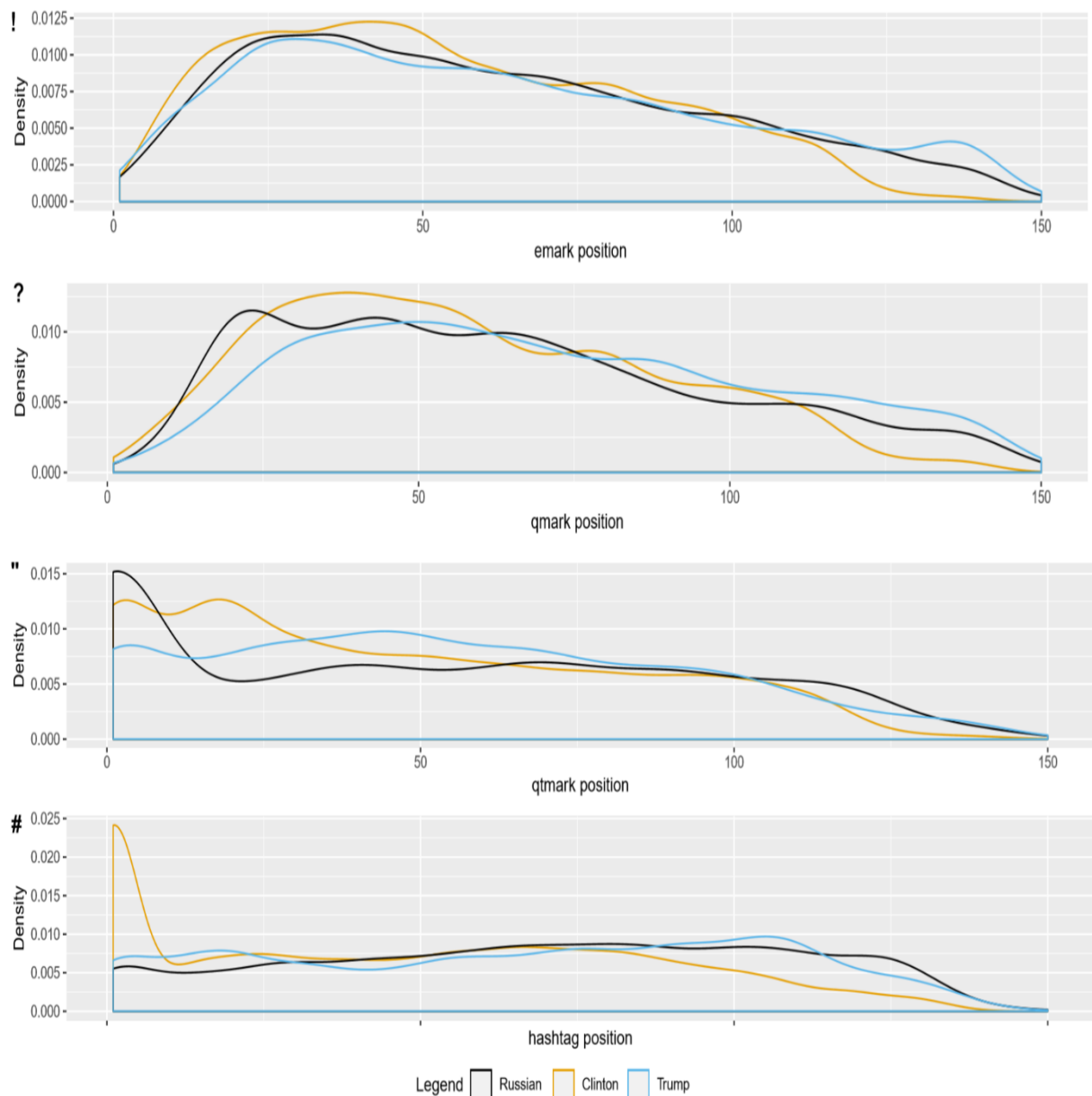


Figure 4: Position of specific punctuation marks across the datasets

As is evident, the use and position of both exclamation marks and question marks match closely across all datasets. Deviations in more discourse-related punctuation include the prominent fronting of hashtags in the Russian tweets. This adds up with the lower overall use of hashtags by Russian tweeters to suggest a different use of the metatextual feature as such.

In addition, there seems to have been a propensity by Clinton supporters for longer quotes (mean character count 53, mean word count 9.5) over Trump supporters (means 33 and 55.7), who are also slightly more likely to use quotation marks for single noun phrases. Russian quotes have a mean of 45 characters and 8 words) and thus lie between.

6 Conclusion

The results of earlier studies pointing to different use of sociolinguistic variants could only be partly replicated. While there is some variation on the lexical level, the tweets of Clinton and Trump supporters are remarkably alike on a structural level, including the use of metatextual hashtags and punctuation. As a result, the patterns observed by the Russian troll tweeters do not match either of the groups, but show some of their own, distinct features.

These patterns paint a similar picture across the three categories of features surveyed. In regard to the lexical variation, Russian tweets used slightly different adjectives, playing up racial tensions between 'black' and 'white'. They also referred back to the sitting president Obama far more than the other two groups. The variety of their lexical items shows a broader picture less concerned with the individual candidates and more concerned with American and international politics at large. In terms of metatextual features, they were more likely to use hashtags and showed a higher propensity for fronting them in their tweets, both of which are likely intended to get hashtags trending and influence current discourse topics on Twitter. On the sublexical level of punctuation, there is once again little variation: exclamation marks, question marks and commas closely match, suggesting that these variables are influenced more by the structural environment and structural limitations of Twitter rather than sociolinguistic factors and influences.

The usage of emojis was the only variable in which all three groups differed significantly, and, maybe more importantly, a variable in which the pro-Clinton and pro-Trump groups showed some variance. This may suggest that emojis are a feature that more closely reflect socio-cultural group identity than surface text features do.

The central question of whether political affiliation is enough of a social factor to drive sociolinguistic variation on Twitter could thus not be answered in the affirmative. If it were a strong factor, we would expect to find more variation between the tweets of Trump and Clinton supporters. Whether or not the differences between the US-based tweets and the Russian tweets may be said to be due to political affiliation on a larger scale or on other extralinguistic factors remains up for debate. More analysis will have to be done on these and other similar features. Given that Nate Silver's platform, FiveThirtyEight, has, more recently, released an even bigger dataset of Russian tweets, the data is available to do so.

References

- Caruso, Assunta/ Folino, Antonietta/ Parisi, Francesca/ Trunfio, Roberto. 2014. A statistical method for minimum corpus size determination. In: Née, Emilie (ed.), *JADT 2014: Proceedings of the Twelfth International Conference on Textual Data Statistical Analysis, June 3-6, 2014*, 135–146. <http://lexicometrica.univ-paris3.fr/jadt/jadt2014/01-ACTES/11-JADT2014.pdf> [accessed 2019-12-17]
- Conover, Michael D./ Ratkiewicz, Jacob/ Francisco, Matthew R./ Gonçalves, B./ Menczer Filippo/ Flammini, Alessandro. 2011. Political polarization on twitter. In: Adamic, Lada A./ Baeza-Yates, R./ Counts, Scott (eds.), *Proceedings of the Fifth International Conference on Weblogs and Social Media, Barcelona, Catalonia, Spain, July 17-21, 2011*. Palo Alto, CA: The AAAI Press.

- Golbeck, Jennifer/ Robles, Christina/ Edmondson, Michon/ Turner, Karen. 2011. Predicting personality from twitter. In: *Proceedings of 2011 IEEE International Conference on Privacy, Security and Trust (PASSAT) and 2011 IEEE Third International Conference on Social Computing (SocialCom)*, 149–156. Los Alamitos, CA, The Institute of Electrical and Electronics Engineers, Inc.
- Hu, Yuheng/ Talamadupula, Kartik/ Kambhampati, Subbarao. 2013. Dude, srsly? the surprisingly formal nature of twitter’s language. In *Proceedings of the Seventh International AAI Conference on Weblogs and Social Media, Cambridge, Massachusetts, USA, 8 - 11 July 2013*. Palo Alto, CA: AAI Press.
- Kearney, Michael W. 2017. rtweet: Collecting twitter data. <https://cran.r-project.org/package=rtweet>. [accessed 2019-12-17]
- Labov, William. 2010. *Principles of linguistic change, cognitive and cultural factors*, vol. 3. Malden, MA: John Wiley and Sons.
- Nguyen, Michelle. 2017. Twitter’s role in politics.
URL: <https://northwesternbusinessreview.org/twitters-role-in-politics-b3ed620465c9>.
- Popken, Ben. 2018. Database of 200,000 russian troll tweets.
URL: <https://www.nbcnews.com/tech/social-media/n844731>.
- Rambukanna, Nathan. 2015. From #racefail to #ferguson: The digital intimacies of race-activist hashtag publics. *The Fibreculture Journal*, 26. 159-188. <http://fibreculturejournal.org/wp-content/pdfs/FCJ-194Rambukanna.pdf> [accessed 2019-12-17]
- Schnoebelen, Tyler. 2012. Do you smile with your nose: Stylistic variation in twitter emoticons. *University of Pennsylvania Working Papers in Linguistics*, 18(2):14.
- Tatman, Rachael/ Paulada, Amandalynne. 2017. Social identity and punctuation variation in the #bluelivesmatter and #blacklivesmatter twitter communities.
URL: http://www.rctatman.com/files/Tatman_2017_SocialIdentityAndPunctuation.pdf.
- Wickham, Hadley. 2019. stringr: Simple, consistent wrappers for common string operations.
URL: <https://cran.r-project.org/package=stringr>.
- Zappavigna, Michele. 2015. Searchable talk: the linguistic functions of hashtags. *Social Semiotics*, 25 (3):274–291. ISSN 1035-0330. doi: 10.1080/10350330.2014.996948.